Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *03/07/2019* **par :**
ELIANE BALAA

Study of the interaction with a virtual 3D environment displayed on a Smartphone

JURY

FRANCK POIRIER	PR - Univ. Bretagne Sud	Rapporteur
BENOÎT MARTIN	PR - Univ. Lorraine	Rapporteur
VÉRONIQUE GAILDRAT	PR - Univ. Toulouse	Examinateur
EMMANUEL DUBOIS	PR - Univ. Toulouse	Directeur de thèse

École doctorale et spécialité :
  MITT : Image, Information, Hypermédia
Unité de Recherche :
  Institut de Recherche en Informatique de Toulouse (UMR 5505)
Directeur(s) de Thèse :
  Emmanuel DUBOIS et *Mathieu* RAYNAL
Rapporteurs :
  Benoît MARTIN et *Franck* POIRIER

Remerciements

After paying my deep acknowledgement to God Almighty, The Holy Spirit, Saint Mary, Saint Elie, Saint Maroun and Saint Charbel, many people have earned my gratitude for their contribution to my research. More specifically, I would like to thank four groups of people, without whom this thesis would not have been possible : my advisors, my lab mates, my colleagues and coworkers, and my family.

First, I am appreciative to my thesis director, M. Emmanuel Dubois. Since my first day at IRIT, M. Dubois believed in me like nobody else and gave me endless support. It all started in fall 2011 when he offered me such a great opportunity to join the lab. He provided the direction and proper orientation, guided me closely specifically at the beginnings when all seemed so dark. Under his supervision, I learned how to define a research problem, find a solution to it, explore, review and analyze previous studies in the field, and finally publish the results.

I am equally grateful to my advisor, M. Mathieu Raynal who accompanied me in the nitty-gritty details of the research. He was present with his inspiring ideas and ways to solve and defeat the problems with a clear smile on his face. He is so ingenious when it comes to analyzing data statistically, interpreting the substantial info, and inferring the conclusions. In difficult times when we had to take the difficult decisions, he was present and confident of taking the right ones. With the development of the research, he proved to be right, efficiently. On a personal level, M. Raynal inspired me by his perseverance, endurance, hardworking and passionate attitude.

To summarize, I would give both M. Dubois and M. Raynal most of the credit for becoming the kind of scientist I am today.

I would like to thank my lab mates for their continued support and for making

Table des matières

Table des figures

It always seems impossible until it's done.

Nelson Mandela

Introduction

A three-dimensional Virtual Environment (3D VE) is a computer-based environment that simulates a wide range of worlds. They include the real world [de Souza et al., 2010, Rohs et al., 2007, Vemavarapu and Borst, 2013], the science fiction [Piazza et al., 2013, Watts et al., 2010], super heroes [Jankowski and Hachet, 2013] and historical locations [Rodríguez and León, 2015, Lehtinen et al., 2012, Riboldi and Maciel, 2010]. Moreover, 3D technologies have seen increased popularity and they are recently being implemented within different applications and areas. Such applications include : desktop computing and CAD [Liang and Green, 1993, Medeiros et al., 2013], 3D games [Daiber et al., 2012, Zaman et al., 2010], messaging tools, video conference [Billinghurst et al., 2002] as shown in Figure 1 - left, Virtual Environments [Stuerzlinger et al., 2006] as shown in Figure 1 - center, Augmented Reality [Gordon et al., 2002] as shown in Figure 1 - right, ubiquitous computing, aeronautic [Raynal et al., 2013] or tele-operation [Zhai et al., 1997, Bowman et al., 2001c].

Nowadays, Smartphones provide new computing capability and connectivity far beyond the traditional voice communication [Robbins et al., 2004]. The variety of mobile input/output (IO) devices such as built-in sensors like, the gyroscopic sensor [Ahrens et al., 2008], built-in camera, accelerometer, built-in touchpad and touchscreen, and the Internet connectivity exploded the boom in Smartphones application, usage and interaction techniques [Teather and Stuerzlinger, 2008]. The improvement of Smartphones hardware performance [Scoditti et al., 2011] enabled mobile devices to exploit more innovative 3D applications and develop the capability to interact on these applications. Thus, 3D Virtual Environments became portable and accessible

FIGURE 1 – left) Video Conference : 3D live interface - center) Virtual Environments : The CoViD system in action - right) Augmented Reality

on mobile devices.

Smartphones operate a multitude of daily services. In addition to providing the traditional voice communication, they can perform office work, connect to social web and can be used as portable calendar, radio, camera, voice recorder, flash disk, internet browser, television, entertainment tool and much more [Garner et al., 2005] as shown in Figure 2. For example, employees of an organization connect through Remote Access to its network remotely and use the office facilities as if they are at the office. Thus, it can replace the traditional computer or laptop. It is a unified device and has the advantage of having a small size. Moreover, business imposes more requirements in terms of mobility, portability and telecommuting. The advancement in communication technologies and business applications made the realization of these requirements possible. Again, Smartphones gain advantage over the laptops and desktops in this area due to their small size and mobility.

Due to increased processing power and improved screen capability, 3D applications run on smart devices. 3D games [Daiber et al., 2012, Zaman et al., 2010], Pokémon Go [Serino et al., 2016] as shown in Figure 3 - left, 3D games for entertainment, GPS to be located on maps and navigation and 3D representations of archaeological site and interactions with visitors and tourism [Riboldi and Maciel, 2010] as shown
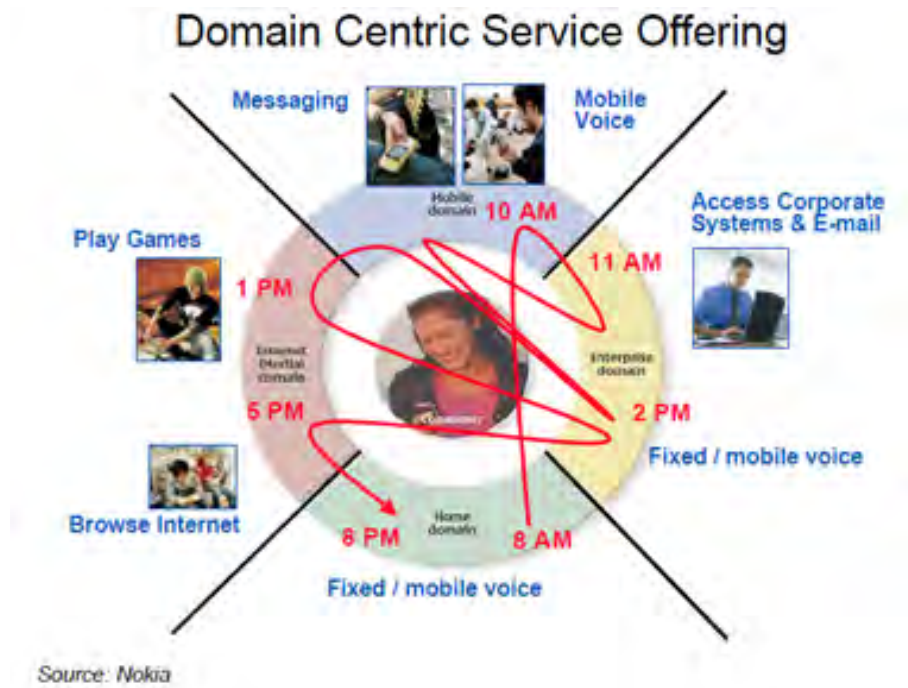
FIGURE 2 – Smartphones operate a multitude of daily services [Garner et al., 2005]

in Figure 3 - right are some applications provided by the Smartphones.

Even though the 3D VE is a virtual simulated environment, users need an interaction technique to employ the 3D VE. Consequently, designing an efficient, accurate and quick interaction technique is necessary to exploit a 3D application promptly and actively. Accordingly, 3D VE requires real-time actions and responses. Consequently, researchers in 3D virtual environment have developed new techniques and metaphors for 3D interaction on mobile devices [Dang et al., 2003, Steed, 2006]. This was done to improve 3D applications usability, to support the selection task, and to face the problems or factors affecting the selection performance [Wobbrock et al., 2008, Hwang and Jeon, 2009].

The variety of input and output devices on Smartphone extends the horizon to envision many 3D interaction techniques. Multiple studies described in literature show different ways of interaction using the input/output devices (IO devices) [Stuerz-

FIGURE 3 – left) Pokémon Go ; right) Historical heritage via an augmented reality interactive system [Riboldi and Maciel, 2010]

linger and Wingrave, 2011, Hachet, 2010]. The vast majority of mobile devices are equipped with a keyboard and function key (buttons). Therefore, it is important to study how to interact with a 3D environment using these devices access tools. For example, a proposed interaction technique in literature [Hynninen, 2012, Jankowski and Hachet, 2013] allows users to interact with a 3D scene using the navigation (Up and Down) keys to move around in a 3D environment. In fact, this is a common technique commonly used in video games.

Likewise, smartphones allow their users to interact with a 3D scene using a touchscreen that represents a real advantage compared to the keys for introducing continuous actions compared to the discrete actions only. The touchscreen then provides the ability to manipulate a cursor continuously. When used in a 3D scene, a multi-touchscreen enables detecting the existence of several points of contact and their positions allowing the use of several fingers simultaneously to interact with an application.

Today, a large majority of Smartphones is equipped with a video camera having a high resolution and potentially close to professional digital cameras performance. The possibilities to using the mobile device camera to interact within a 3D scene are large. It is possible, for example, to detect and track over time the position of specific points of the scene for analysis in the video stream as shown in Figure 4 [Çapın et al.,

2006]. However, this is not trivial and can be costly in terms of computing time.

Figure 4 – Snapshots of the views acquired from Li's and Cheng's third-person camera control module in a Stonehenge-like ring of standing stones [Jankowski and Hachet, 2013]

Furthermore, some mobile devices are equipped with special sensors, such as accelerometer, gyroscope, positioning system and magnetometer, which can be very useful for interacting with a 3D environment [Pietroszek et al., 2015, Bergé et al., 2015]. These sensors allow to add more dimensions in interacting with the device and to provide a smoother and more intuitive interaction. Various mobile applications take advantage of accelerometers. For example, in 3D games to drive a car by rotating the device to the left or right as shown in Figure 5 - left [Du et al., 2011, Daiber et al., 2012], in an audio player to switch to the next track, in the playlist by shaking the device to the right [Ahrens et al., 2008], or in a photo viewer to rotate a photo by changing the orientations of the mobile as shown in Figure 5 - right [Rodríguez and León, 2015].

It is also possible to add the information provided by a positioning system such as GPS, for example, in order to place the user in a 3D environment or in an augmented
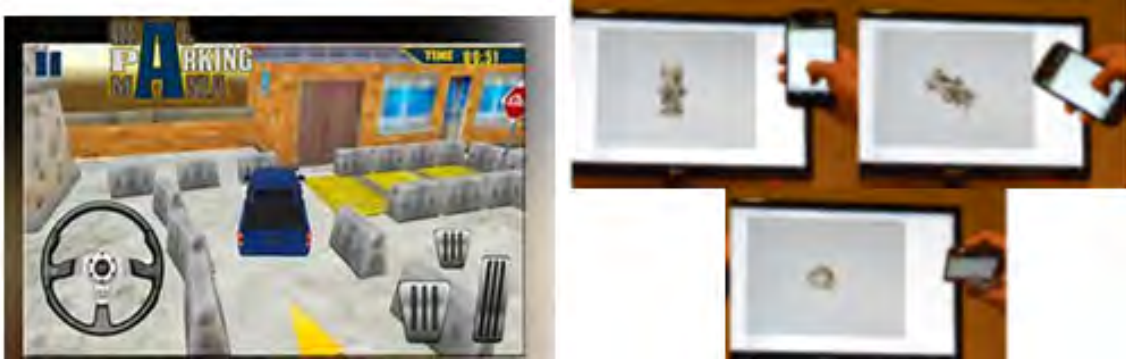
FIGURE 5 – left : Parking mania mobile application : the tilting gesture enables the car rotation to left or right ; at right : A remotely rotation of a photo using the mobile orientation [Rodríguez and León, 2015]

reality. For example, the Pokémon-Go game uses the mobile location found from GPS and the augmented reality (AR) technology to introduce in the 3D environment the fictive object (a Pokémon) to chase. Consequently, the most suitable interaction techniques to the Smartphone combine the use of hardware capability (touchscreen, physical button, camera, etc.) with sensors (accelerometer, gyroscope) integrated on the Smartphone.

However, small devices are not largely adopting 3D VE applications except few of them such as video games and augmented reality applications that involve a limited number of items in the virtual scene [Marzo et al., 2014]. This is due to the lack of availability of interaction techniques to manage a dense 3D VE composed of many objects close to each other and displayed on a small and flat screen. Recently, technologies allow the complete display of 3D real-time scenes on Smartphones, but the problem remains on how to control and interact within these scenes.

For this thesis work, the analyst chose to focus on the selection task. Indeed, this task is one of the most used elementary tasks during the manipulation of 3D VE [Bowman et al., 2004]. In a mobile device, selecting an object in 3D crowded VE holding a high density of objects faces the following constraints in interacting with this VE : the environment density, the depth of targets and the occlusion. This work presents three contributions. First, a technical contribution residing in develo-

ping a 3D interaction technique : DichotoZoom, that is both powerful and easy to learn and use. The originality of this approach is particularly to eliminate the 3D VE constraints (density, depth, and occlusion) in the context of the selection task. Second, an experimental contribution consisting in designing and comparing four interaction modalities, for the DichotoZoom technique, based on the various capabilities and sensors presented on smartphones. This approach shows the huge capacities provided by smartphone devices to improve technology and services. Third, a theoretical contribution proposing a new classification for the interaction techniques, since this classification aims to propose guidelines used to evaluate an existing interaction technique or to design a new one that responds to the environment properties and limitations.

The first chapter of this thesis presents a state of the art of the existing selection techniques in 3D VE and the selection techniques on smartphones. It exposes the selection techniques in 3D VE structured around the selection subtasks : navigation, pointing and validation. Moreover, it describes disambiguation techniques providing the selection of a target from a set of pre-selected objects. Afterward, it exposes some interaction techniques described in literature and designed for implementation on Smartphone. These techniques are divided into two groups : techniques performing two-dimensional selection tasks on smartphones, and techniques performing three-dimensional selection tasks on smartphones. Finally, we expose techniques that used the smartphone as an input device.

The second chapter discusses the problematic of selecting in 3D VE displayed on a Smartphone. It exposes the three identified selection problems : the environment density, the depth of targets and the occlusion. Afterward, it establishes the enhancement offered by each existing technique in solving the selection problems. It analysis the assets proposed by different techniques, the way they eliminates the problems, their advantages and their inconvenient. Furthermore, it illustrates the classification of the selection techniques for 3D VE according to the three discussed problems (density, depth and occlusion) affecting the selection performance in a dense 3D VE.

The third chapter focuses on defining and describing the fruit of this study : DichotoZoom interaction technique. It compares and evaluates the proposed technique

to the Circulation technique, suggested by the literature. The comparative analysis shows the effectiveness of DichotoZoom technique compared to its counterpart.

The fourth chapter evaluates the use of DichotoZoom in different modalities of interaction available on Smartphones. It reports on the performance of the proposed selection technique based on the following four interaction modalities : using physical buttons, using graphical buttons, using gestural interactions via touchscreen or moving the device itself.

Finally, the fifth chapter lists our contributions to the field of 3D interaction techniques used in a dense 3D virtual environment displayed on small screens and proposes some future works.

Chapitre I

Selection Techniques

The first chapter of this thesis presents a state of art of the existing interaction techniques. The first part focuses on the analysis of the selection interaction task in 3D VE. It explores the three basic subtasks identifying the selection task in 3D environment (the navigation, the pointing and the validation subtask) [Bowman et al., 2004]. Next, we concentrate on the exposition of the interaction technique used in the pointing subtask. These techniques are divided into two categories depending on the number of target selected ; interaction techniques used to point a single object and interaction techniques used to point a subset of objects. In case of the selection of a subset of objects, a selection technique must be used to eliminate the ambiguity of the user selection. Moreover, this chapter will describe the different strategies used to reduce the ambiguity then to point the target.

Afterward, the second part of this chapter exposes a variety of interaction techniques described in literature and designed for implementation on Smartphone. These techniques are divided into two groups of techniques performing two-dimensional selection tasks on smartphones or three-dimensional selection tasks on smartphones.

Finally, the third part shows techniques that used the smartphone as an input device.

We will start this chapter by defining the performance criteria used to evaluate the interaction techniques performance.

1 Performance metrics

Bowman [Bowman, 2002] defines performance metrics as an indicator allowing the quantification of the performance of a technique, a comparison tool between studied techniques and a measure of application interaction requirements.

Performance metric can be divided into two groups : the objective performance metric deduced from an objective evaluation and objective results and the subjective performance metric deduced from a subjective evaluation and subjective results. Some of the common performance metrics for interaction techniques in virtual environment are listed below :

Two aspects are considered for the objective evaluation : task completion time and error rate.

The task completion time refers to the time of the task execution. Measuring the task completion time allows comparing the speed of execution for various techniques and determines the ranking of technique in terms of speed. The task completion time is the time calculated from the display of the 3D VE on the screen until the users click on the screen to validate the selection of the target then it is the time calculated from the start of a trial until the confirmation of the selection [Rohs et al., 2007].

Error rate consists of measuring the number of wrong selection during the task. Error rate evaluate the techniques accuracy by measuring its precision or exactness. The task error rate is the number of times participants performed a wrong or inaccurate selection regarding the total number of task to achieve. The task's validation error occurs once the participants validated the selection for an object different from the desired target [Teather and Stuerzlinger, 2013].

Three aspects are considered for the subjective evaluation : usability, attractiveness and user preferences.

Usability measures the user acceptance and buy-in to the technique. This metric is usually got through personal feedbacks. We measured the usability of each technique via the SUS questionnaire System [Brooke et al., 1996] (SUS : the System Usability Scale Questionnaire). The SUS questionnaire System consists of 10 items, with odd-numbered items worded positively and even-numbered items worded negatively, the

participant respond to an item by assigning a number. The SUS questionnaire System consists of 5-point scales numbered from 1 (anchored with "Strongly disagree") to 5 (anchored with "Strongly agree") [Lewis and Sauro, 2009]. After each session, participants were asked to complete the SUS survey, a SUS score was computed for each interaction technique : given the score the usability the interaction technique is rated.

Attractiveness expresses the user experience and measures the technique appeal. It is usually measured by personal reports and questionnaires. To evaluate and compare the attractiveness of an interaction technique, the AttrakDiff evaluation method [Hassenzahl, 2004] is used. AttrakDiff evaluates the user experience and measures the attractiveness of a product. AttrakDiff is a questionnaire based method used to evaluate mobile device applications. Answers to the questions are multistage rating scales. Each of the four dimensions has items representing features of the system. Within each item a word-pair extent a rating scale and it consists of two extremes and seven stages in between. It has four dimensions for evaluating a system. The first dimension is the pragmatic quality (PQ) : it measures how much the system characteristics are useful and usable and how they contribute on achieving the objectives and the goals correctly. The second dimension is the hedonic quality - identity (HQ-I) : it measures the product's ability to help user represent their identity to the others. This value is used to indicate the degree to which the user self-aligns to the product. The identification helps the user to transmit his professionalism, style and innovation through the system. The third dimension is the hedonic quality - stimulation (HQ-S) : it represents novelty, inspiration, stimulation, interesting features, and motivation that the system offers. This stimulation captures the user's attention, motivates him, and engages him to the accomplishment of his work and to use that product. Finally, the fourth dimension is the attractiveness (ATT) : it describes the global value of the product and the ability to give the user a good impression of it. Hedonic and Pragmatic qualities are not related to each other, and provide equally to the evaluation of attractiveness.

Participants were introduced and trained on using the different interaction techniques, they filled in the AttrakDiff questionnaire then the evaluation is displayed

into three diagrams : the profile of the Word-pairs, the diagram of average values and the portfolio presentation.

The profile of the Word-pairs displays a detailed list of all word-pairs presenting the items in AttrakDiff and shows the mean values of each word-pair for the studied interaction techniques. The extreme values of the mean values of each word-pair are very interesting for analysis because they show which are either critical or well-resolved.

The diagram of average values shows the average values of the questionnaire items on the level of the dimension. The mean values of the four dimensions are visible. The study of the mean values shows the characteristics of every interaction technique then leads to compare and evaluate them.

The portfolio presentation displays the overall evaluation of the AttrakDiff questionnaire and classifies the application into a "character-regions". In the portfolio presentation the values of the hedonic quality are represented on the vertical axis (the bottom level has the lowest value) and the values of pragmatic quality are represented on the horizontal axis (the left value has the lowest values). Each product is classified into one or more "character-regions" depending on the confidence rectangle. The confidence rectangle gives information about the evaluation's reliability and certainty. A small confidence rectangle shows that the results are more reliable and less coincidental. The bigger the confidence rectangle shows that the user's results are more variable from one to another.

User Preferences determines the interaction technique effectiveness perceived by users and gives the personal opinion of the user in ranking the techniques. This factor is deduced from indications by users and determined in comparative studies. Where, at the end of the sessions, participants were asked to rank the different interaction techniques from the best to the worst according to their preference and give comments on each technique. The ranking often depends on the amount of fatigue, discomfort and satisfaction they experienced while using the interaction technique. Using diagrams, the rank and the comments evaluation can be established to show the effectiveness of each interaction technique.

2   Selection task in 3D VE

Various interaction techniques have been developed by researchers to interact in interactive 3D environments. For 3D VE applications, interaction can be categorized in terms of three universal interaction tasks : Navigation, Selection and Manipulation, and System Control [Bowman et al., 2004].

Navigation task refers to exploring the contents of the 3D virtual environment in different ways to search for the target and visualize it. This task confers to motor and cognitive components. It is composed of two subtasks : travel and wayfinding [Jankowski and Hachet, 2013]. Travel is the motor component of navigation ; it designates the physical movement of moving a user's viewpoint through an environment from place to place to look for a specific target. Wayfinding is the cognitive component ; it designates the navigator's ability to conceptualize the space for making the right decision to avoid getting "lost-in-space".

Selection consists of specifying and choosing an object or a set of objects for a certain goal, then manipulating it. Manipulation consists of modifying object properties such as : position, orientation, scale, shape, color, texture and behavior. 3D manipulation interaction techniques can be found in three fundamental tasks : object translation (positioning), object rotation, and object scaling [Jankowski and Hachet, 2013]. Selection and manipulation are often combined together, but selection may be considered a separate task. Finally, System Control designates communication between a user and a system. This task is independent of the virtual environment. It refers to a task, in which a command is executed to change either the state of the system or the mode of interaction [Bowman et al., 2001b].

In this study, the focus is on the analysis of the selection interaction task, since, in complex interactive virtual environments, selection tasks are frequently performed. The objectives of the selection task are to request information about the object, to move it to another location, to indicate an action, or to configure it.

2.1 Navigation subtask

Navigation is the process of getting around a virtual environment, to discover its constituents and to seek a specific target in it. Therefore, the goal is to find the target, to make it visible and to travel towards its direction. The user may change his viewpoint through the environment or move the scene to visualize the target. In this phase, researchers define four types of navigation : the general movement, the targeted movement, the specified coordinate movement and the specified trajectory movement [Jankowski and Hachet, 2013].

The general movement is an exploratory movement ; its objective is to gain knowledge of the environment. The principal possible actions used are rotating, panning, and zooming [Phillips and Badler, 1988]. Rotate is defined by the orbiting of the camera around a central point in any direction ; Pan is defined by the translation of the camera along x and y axes ; and Zoom is defined by the translation of the camera along its line of sight as shown in Figure 6.

FIGURE 6 – Rotating, panning, and zooming are the primary camera movements used in almost every 3D modelling environments

The other way of navigating in a 3D VE is simply to let the user walk along the ground, drive a vehicle or fly a plane [Bowman and Hodges, 1997] proposed by Walking/Driving/Flying techniques ? Finger Walking in Place (FWIP) technique [Kim et al., 2008], uses the multi-touch technology to travel in a virtual world. The user can move forward and backward and rotate in 3D VE while their fingers, representing human legs, slide on multi-touch sensitive surface. Figure 7 illustrates the way of using two-handed fingers and the way of using of one-handed fingers.

The world-in-miniature (WIM) technique [Stoakley et al., 1995] uses a small

28
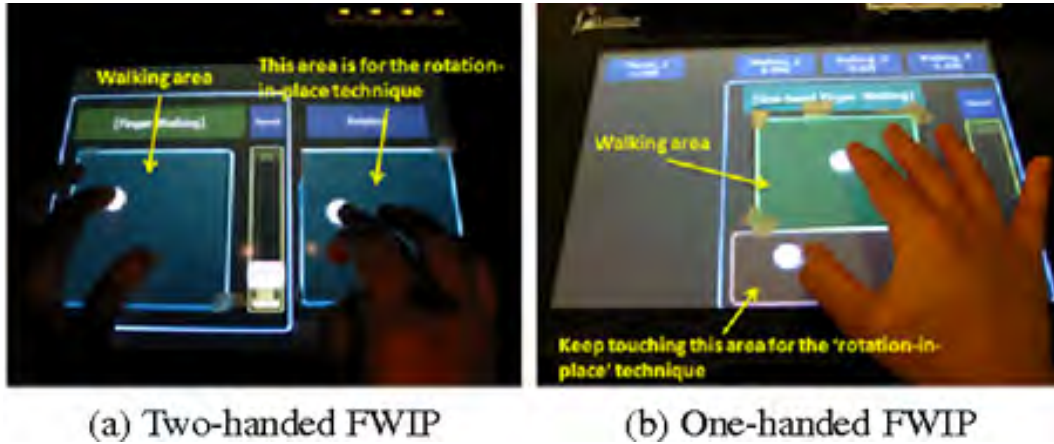
Figure 7 – Finger Walking in Place (FWIP) technique : a) Two-handed FWIP b) One-handed FWIP

representation of the 3D environment to manipulate and interact with its objects. The user sees through a head-tracked display the 3D VE and holds in his dominant hand a buttonball used for selection and special action and in his non-dominant hand a clipboard used to display a miniature copy of the world as shown in Figure 8 - right. He uses a simple virtual hand for indirect manipulations of the full-scale objects by interacting with their icon in the small version. The user manipulates the miniature objects and then affects the full-scale objects as shown in Figure 8 - left. By moving the clipboard outside the visual field, the user can only see the 3D VE without the small copy. By moving the clipboard or flying the object representing the user, we can navigate the VE following different perspectives. By pressing the button of the buttonball, the user can select the object in the VE or its corresponding in the miniature copy. To move the object in the 3D VE, the user moves its corresponding in the WIM environment. To display the WIM environment through several scales, the user presses a button of the buttonball. This technique enables different actions : the rotation of an environment object or of the whole WIM (the rotation angle limit is 30 degrees, and a complete rotation demand repeated rotations), the view of alternate representations, the different types of updates and the display of multiple WIMS to visually compare different scales, different locations, and/or different time. This

technique is accurate and the user uses both hands.

FIGURE 8 – The world-in-miniature (WIM) Technique

This interaction technique is a point cursor technique, used to select a single target. The cursor is used to select a 3D visible and deep target. Whereas, the WIM technique [Pausch et al., 1995] is used for pointing and navigation. It combines the pointing and the navigation subtasks to perform the selection task. This interaction uses two copies of the virtual environment, two hands, several views and several scales. Circulating in Depth is performed by the pointing subtask. Whereas circulating a pointer or the navigation subtask is done using a zoom. The user employs a separate flying mechanism then points the target in the real 3D VE or in its miniature copy. If the target position in depth is beyond than the hand reach or difficult to touch, the user makes a zoom or changes the point view in the navigation subtask to ease the selection. This interaction technique uses zoom to reduce the density of the environment and to simplify the selection task.

The Targeted movement is a searching movement ; its objective is to move in to examine a detail of an object or a part of the environment specified by the user. The main possible actions used are listed next. The Path Drawing [Igarashi et al., 1998] consists of drawing a path for the camera as shown by the sketch-based navigation example in Figure 9 - left. The Point of Interest Movement (Go-To/Fly-To) [Mackinlay et al., 1990] is about jumping the camera to the goal position that the user had specified. Hyperlinks and Bookmarked Views [Ruddle et al., 2000] permit

instant navigation and displacement to the target. This kind of technique is fast but causes cognitive difficulties and disorientation. Finally, the Navigation by query [van Ballegooij and Eliëns, 2001] allows users to navigate the environment based on prior assumption or knowledge of the world. This technique is difficult to apply without prior information about the 3D VE and the target position. This kind of navigation is a guided navigation.

The Specified trajectory movement refers to a movement along a position and orientation trajectory. It is a guided navigation that limits the user's freedom while travelling through the 3D VE. The principal possible proposed action is the Guided/Constrained Navigation which consists of creating constrained path to help the user navigates the environment and avoids disorientation as in StyleCam experience illustrated in Figure 9 - right [Burtnyk et al., 2002] : the top of the Figure 9- right shows the user input and the bottom of in Figure 9 - right shows what the user sees. The Automatic Viewpoint/Path/Motion Planning helps the navigation of the user in a 3D VE. This is done by creating a first constrained path based on an analysis of the required tasks in a specific environment, for example. Salomon et al. [Salomon et al., 2003] and Nieuwenhuisen and Overmars [Nieuwenhuisen and Overmars,] present techniques based on an algorithm that computes a global roadmap of the environment for interactive navigation. Lastly, the Assisted Navigation techniques [Chittaro et al., 2003] leads the user on a guided tour of the virtual environment often used in games (e.g., Star Wars : Knights of the Old Republic).

The Specified coordinate movement refers to a movement to a precise position and orientation. The user has to precise the exact position and orientation of the target by entering a triplet of (x, y, z) coordinates for the destination positions and the orientation using a keyboard. Used in CAD and 3D editing software, this type of navigation precise the position of the camera.

Research's recommend efficient wayfinding to avoid user getting lost in space, being unable to relocate recently visited places, and wandering how to find a place for the first time. Efficient wayfinding is based on one of the following tools : Maps that showed to be a priceless tool for acquiring orientation and position in a 3D VE [Darken and Sibert, 1996], guidelines for landmarks proposes by Vinson [Vinson,
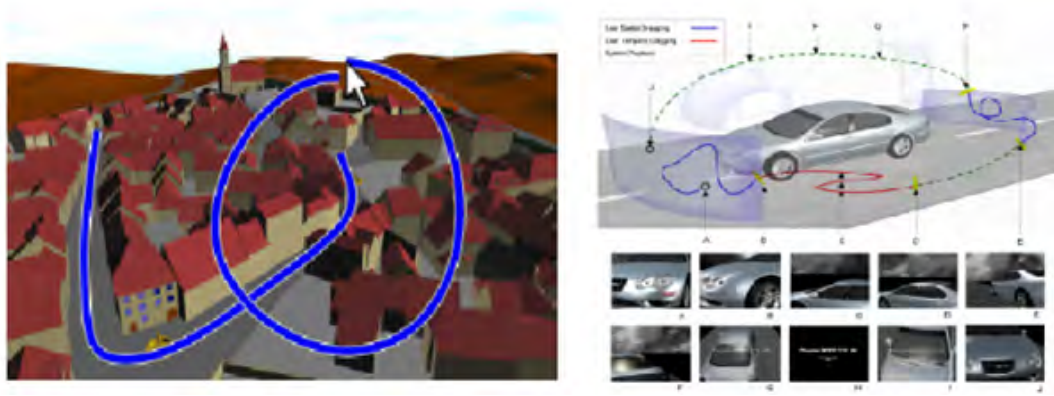
FIGURE 9 – At left, example of a sketch-based navigation [Hagedorn and Döllner, 2008] ; at right, example of StyleCam experience. On the top : system components and their reaction to user input. On the bottom : what the user sees

1999], semitransparency as a navigation utility in Virtual Environments [Chittaro and Scagnetto, 2001], 3D arrows as a navigation tool permitting users to easily locate objects inside large-scale environments [Chittaro and Burigat, 2004], or ViewCube as a 3D orientation indicator and controller for CAD users [Khan et al., 2008].

2.2   Pointing subtask

The pointing task supposes that the scene is fixed and the desired target is visible. It refers to moving a pointer in the 3D VE to reach the target [Grossman and Balakrishnan, 2006]. Accordingly and in general, the pointing action consists of designating from the scene : the target from a spare environment or a set of objects containing the target from a dense environment and then performing a refinement pointing step to catch the desired object. Accordingly, based on the number of selected items, the discussed interaction techniques are classified into two broad categories : the immediate selection techniques and the progressive refinement techniques.

In one hand, the immediate selection technique allows the direct selection of a single object. On the other hand, the progressive refinement technique works in steps : first, it selects a set of objects. Second, it needs a phase of selection to point to the

desired object among the pre-selected ones. Coupled to the selection technique, a disambiguation technique helps finding the desired object.

The immediate selection technique needs a single high-precision selection without refinement because it selects one target in one-step. Contrarily, the progressive refinement technique consists of two phases of selection one targeting a set of pre-selected objects and a second one pointing to the specific object in the set. The progressive refinement selection techniques takes into consideration a performance compromise between the techniques using one precise action to select an object and the progressive refinement techniques requiring very low selection precision repeated multiple times. The immediate techniques execute a single selection but the selection is slow and has a high error probability. Conversely, progressive techniques execute multiple selections but every step is fast and accurate.

Thus, the main possible interactions for pointing an object in a 3D VE are either a direct or an indirect interaction.

Direct interaction is often possible using a touchscreen. In this case, the user can directly point to the desired object by touching the screen. For example, Daiber [Daiber et al., 2012] proposes sensor-based interactions for mobile 3D gaming. Rotation and Flipping Gestures make the navigation possible. The combination of a Direct Touch with Tilting and a Rotation of the mobile allows the object manipulation. Object selection is realized by a simple click where the uppermost pointed object will be selected. However, these techniques are not useful in case of a dense VE to point partially or fully occluded objects. In case of an indirect interaction, the user moves a pointer or a ray in the 3D VE by means of a pointing device. Generally, a transfer function makes it possible to transform the movements made with the pointing device into a movement of the virtual pointer in the scene.

The following paragraphs expose the different existing solutions for these cases. Interaction techniques are categorized into two sets : techniques used to point immediately a single object and techniques used to point a subset of objects.

2.2.1 Pointing a single object

The immediate selection technique points only one object and uses one precise action without refinement. The simplest way is to consider that the first element touched by a point cursor or by a ray is the selected item. Accordingly, researchers proposed the point cursor techniques and the ray cursor techniques. However, this strategy can only be used in sparse environments. In the case of high-density environments and occluded objects, the immediate techniques execute a slow and a high error probability selection.

2.2.1.a Point Cursor based techniques   The techniques based on a Point Cursor displacement are analogue to the classical 2D pointing. These interaction techniques are used to select one target in case of a visible target or expected target position [Schmitt et al., 2012]. These types of techniques consist of moving a 3D cursor in the 3D VE toward the desired object to point it [Schmitt et al., 2011]. The common 3D cursor used is a small cursor [Hinckley et al., 1994] or a virtual hand [Poupyrev et al., 1996]. To do this, either the users interact with a specific pointing device such as a 3D mouse [Raynal et al., 2013] or they may use a tracking system for user movements, usually the user's hand movement, and associate it to the pointer displacement in the 3D scene [Bowman et al., 1997].

The simplest pointing technique consists of moving a pointer in the 3D VE. A pointing device controls the pointer movement. The 3D Point Cursor [Grossman and Balakrishnan, 2006] is an interaction technique where a cursor moves in the virtual environment (cf. Figure 10). When the users press a button, the 3D position and orientation of their input device are linearly transformed into coordinates in the VE and a cursor (the cross-hair of the example in Figure 10) shows at this position. To reach the target, the users move their input device gradually. Consequently, the cursor moves according to this displacement. If the cursor touches an object, this object is highlighted (the red object of Figure 10). Thus, the selection is confirmed by pressing a button of the input device. To point a target precisely, the users have to make fine displacement of the cursor. However, this task is complicated since it requires more

accurate movements and multiple adjustments in depth. Studies showed that a cursor based techniques face two major problems. First, these techniques are very slow, especially when the distance between the cursor and the target is large [Bowman et al., 2001b]. This delay is due to lack of precision of the cursor size. This slowness results from a decrease in precision associated with the size of the cursor according to the depth. Indeed, the more the cursor moves in depth, the smaller its size goes on the screen. It, then, becomes more difficult for the users to select precisely the desired object. Second, in addition to the problems resulting from the size of the target the users make fine displacement of the cross-hair to point the target precisely. Consequently, the response time increases to make many displacement of the input device to catch the target. This technique depends on the target size and becomes slower as the size decreases. This technique presents a shaking hand problem and on the other hand this technique is accurate. As a result, the selection becomes even more complicated if the 3D VE is dense.
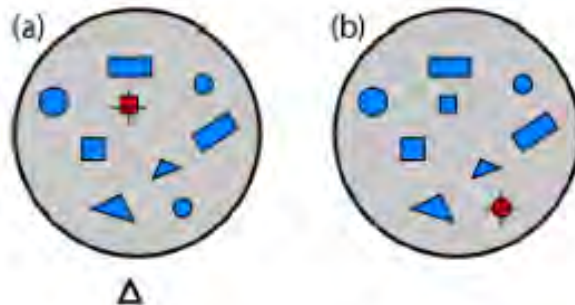
FIGURE 10 – The 3D Point Cursor Technique. A pointer is represented by a cross-hair and controlled by the pointing device. The red object is the target pointed by the user using the pointing device.

2.2.1.b Ray cursor based techniques Another type of techniques consists of emitting a ray in the 3D VE toward the desired object to point it [Liang and Green, 1994]. The user emits a virtual ray through the target direction to point it more easily in depth. This ray may intersect multiple objects. Then, the first object

intersecting with the virtual ray is selected. These techniques are based on the Ray cursor technique [Bowman and Hodges, 1999]. To avoid the problem of crossing a large number of objects by a ray, researchers suggest curving the ray.

The Ray cursor technique [Bowman and Hodges, 1999] is an interaction technique that consists of emitting an infinite straight ray is in the 3D VE. In ray cursor technique, the users control the start point and the orientation of a virtual ray by adjusting the input device position and orientation. After detecting the tracker, a virtual ray is emitted from the user through the target direction. This ray can intersect multiple objects. In this case, the virtual ray intersects with the first object that is highlighted and the users confirm the selection by pressing a button on the input device as shown in Figure 11. This technique reduces the response time ; it is faster than the 3D point cursor pointing a single target. The Ray cursor technique is accurate, independent of the target size even if the performance decreases while the size decreases. This interaction technique presents some limitations in case of a small target (since many displacements of the position of the input device are made to catch the target) and it presents a handshake problem.
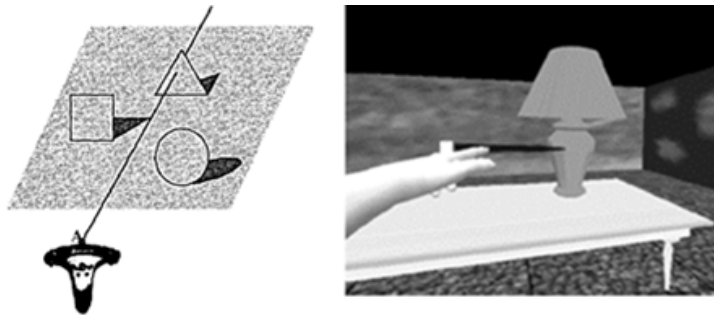
FIGURE 11 – Grabbing an object via The Ray Cursor Technique

An infinite ray emitted in the 3D VE may intersect with multiple objects. The Ray cursor technique permits to select the first intersected object. A problem occurs, if the target is placed behind a number of objects. To avoid the problem of a ray crossing a large number of elements and thus having to use a second step of selection, Curve techniques offer a possible solution. They allow bending of the ray launched

by the user and thus bypassing the objects that would hide the desired one.

The Curve with fixed length technique [Dang et al., 2003] is an interaction technique where a fixed length curve is moving within the virtual environment to select a visible target. The source and destination of the Curve define its end-points. The source is the input device (wand) position (point A in Figure 12). The destination (point B in Figure 12), points to the target in the VE. After drawing the curve and, in order to select the target with precision, the user moves the position of the wand, and consequently, the fixed curve moves, as shown in Figure 12. Once the destination endpoint reaches the target, it is highlighted and then selected by clicking on the input device button. Thus, the pointing task is completed. The Curve techniques are based on the transparency of the environment. This technique presents a displacement problem due to the curve travel in a non-transparent environment. Using a curved ray has the advantage of selecting a deep target without passing through the environment objects. Accordingly, the user can detect the position of the target and make the necessary adjustment to reach it. This task is complicated since it requires more accurate movements and multiple adjustments in depth. The user has to move multiple times the wand position in order to select a target. This interaction technique presents some limitations resulting from a shaking hand or a small sized target since the response time increases with the number of displacements of the input device until catching the target. This technique depends on the target size and becomes slower as its size decreases.

2.2.2 Pointing a subset of objects

In a high density 3D VE, pointing an object with the previous techniques leads to the selection of several elements of the scene at the same time. Therefore, the pointing task is difficult and inaccurate and it requires multiple adjustments. To ease the selection of an item in a 3D VE, many techniques based on finite- or infinite-volume ray or point cursors are proposed. Therefore, the pointing volume is then larger and all the objects that are inside (or in contact with) the volume can then be selected. In order to finalize the selection task, a refinement-pointing phase of these
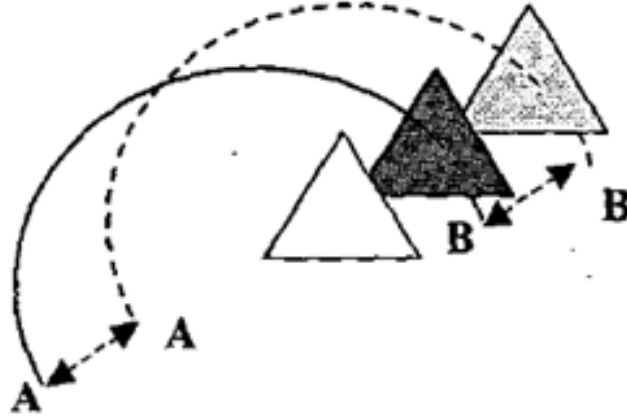
FIGURE 12 – The Curve with fixed length technique

preselected elements is a necessity.

2.2.2.a Defined volume based techniques This section presents a number of interaction techniques using finite volume cursors for subset selection. The volume contours have well-known shapes : a rectangular volume cursor (used in the Silk cursor technique), a transparent sphere (found in the transparent sphere technique) and a starfish-shaped volume (seen in the Starfish technique).

Certain types of 3D VEs, such as cloud of structured particle [Yu et al., 2012] or cloud of non-structured particle [Hamelin and Dubois, 2015] can contain multiple elements. In an environment constituted of cloud of non-structured particle, rather than directly selecting a particular element, subset selection techniques are used. To select the desired subset, Naito et al. [Naito et al., 2009] use a sphere ; the user can increase or decrease the size of that sphere to select more or less elements. To make a more precise selection of the subset of points, Hamelin and Dubois [Hamelin and Dubois, 2015] propose to surround the subset of points by a hand. In other clouds of structured particle, it is easier to select a subset according to its density [Yu et al., 2012] or its context [Yu et al., 2015]. Accordingly, tablet tools are proposed by Yu et

al. [Yu et al., 2012, Yu et al., 2015] to surround the desired elements on the 2D plane, and then, an algorithm selects the particles based on the density or the context of the cloud.

The solution proposed to resolve these problems consists of using a volume selection area rather than a small cursor. In the Silk cursor technique [Zhai et al., 1994], the cursor is replaced by a semi-transparent rectangular volume cursor. The semi-transparent rectangular volume cursor movement is controlled by the user's hand. The 3D position and orientation of the user's hand are transformed into coordinates in the virtual environment where a volume cursor is displayed. The user moves his hand to reach the target. Consequently, the coordinates of the cursor change. When the target is inside the rectangular volume, the user closes his hand and the selection is concluded by pressing the spacebar. The rectangular cursor facilitates the selection of small objects since it requires less selection precision. The almost transparent texture (like a "silk stocking") gives information about the position in depth of the object in relation to the volume cursor. This allows the user to know whether the cursor is in front of, on, or behind the object as shown in Figure 13. This technique presents a solution for the shaking hand problem and the finger occlusion problem.
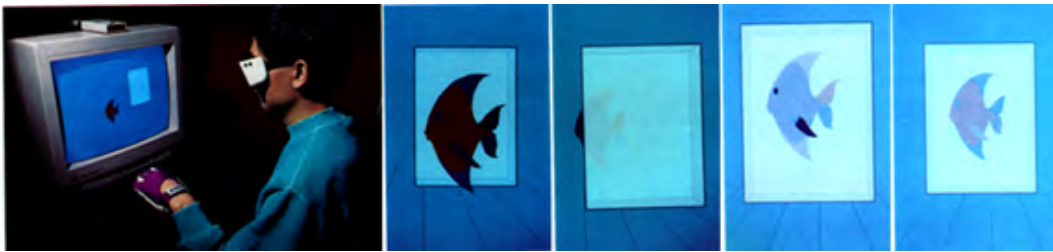
FIGURE 13 – Selecting a fish by using the Silk Cursor Technique (from left to right : The Silk Cursor Technique, A fish in front of the cursor, a fish behind the cursor, a fish partially inside the cursor, a fish completely inside the cursor)

With the transparent sphere technique [Dang et al., 2003], the crosshair is replaced by a transparent sphere. A pointing device (Wand) controls the movement of the transparent sphere. The 3D position and orientation of the wand are transformed into coordinates in the virtual environment where a transparent sphere is drawn.

To reach the target, the user moves the input device. Consequently, the coordinates of the sphere change and the target is one of the N objects inside the sphere. In Figure 14 the target is one of the four objects selected by the sphere. The next section describes the possible disambiguation techniques used in order to finalize the pointing task and to point the desired object from the N preselected objects. This technique selects accurately the target. It is independent of the target size, but it requests a large response time due to having to choose the target from a set of already pre-selected ones. On the other hand, this technique is precise. It does not require accurate movements or multiple adjustments in depth of the input device position to catch the target. And, it does not present a shaking hand problem.
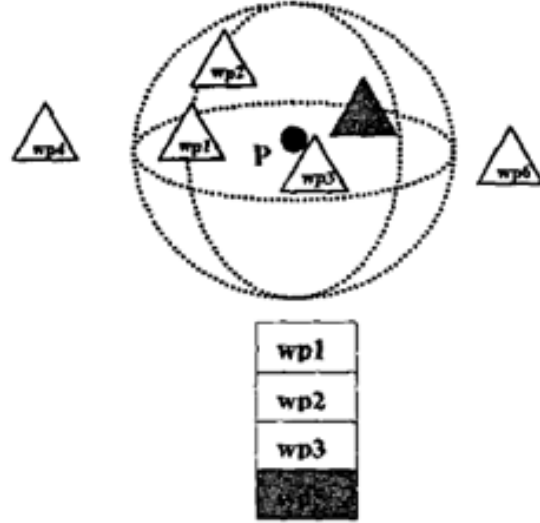
FIGURE 14 – The Transparent Sphere Technique

The Starfish technique [Wonner et al., 2012] is a precise interaction technique to select targets in high-density 3D virtual reality environments. Based on volumetric meshes and a preselecting subset of targets, Starfish provides a solution for occluded targets. The user controls a 3D pointer delimited by a starfish-shaped surface via a tracker device equipped with two buttons : the Move button and the Select button. When the Move button is pressed, the position of the device is represented by the

40

position of the Starfish-head in the 3D scene. At any given moment and according to the position of the head, a starfish-shaped volume is dynamically rebuilt to capture targets near the head as shown in Figure 15. Once the target is preselected by a branch, the user presses the Select button to lock the shape and move the head inside its volume. The head is controlled by a pointer representing the position of the user's hand. The head slides down the desired branch reaches and selects the corresponding target. Starfish cannot preselect all the nearest targets and presents some restrictions in the following cases : if the Euclidean distance between the pointer and the target is larger than a parameter R_{max}, if the angle between two branches is too small, if the number of preselected targets near the head is greater than N_{max}, or in case of a wrong choice, the captive pointer should never jump from branch to branch and the user should have to follow the branch direction to return to the center and to make another choice.
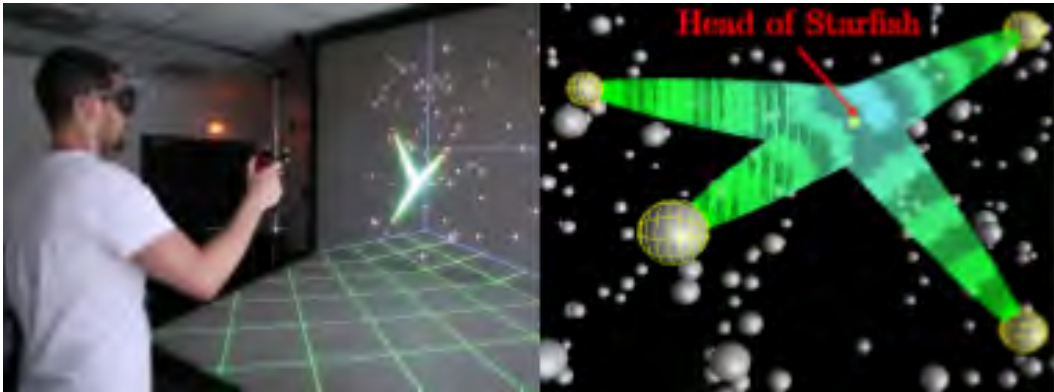
FIGURE 15 – The Starfish selection Technique

2.2.2.b   Infinite volume based techniques   Although, the selection is done using 3D point cursor or ray cursor, the position of the target on the screen has a significant impact on accuracy. In fact, an object placed on the edges of the user's field of vision is more difficult to select than an object placed in front of him. In this case, the user is required to direct the beam with a relatively large angle from its point of view, and hence, the pointing accuracy decreases. Therefore, volumetric rays have

been proposed, such as using a cylinder having the axis as the main ray manipulated by the user [Dang et al., 2003], or using a cone suggested by the "Spotlight" technique [Liang and Green, 1993]. In this case, the vertex of the cone is the end manipulated by the user and the cone is extended in the direction desired by the user. Consequently, the further is the surface from the vertex of the cone, the bigger is the selection area, and then the deepest objects are the easier to be selected.

The transparent cylinder interaction technique [Dang et al., 2003] combines a ray cursor technique and a transparent cylinder. After the detection of the 3D position and orientation of the input device, a ray is emitted. A transparent cylinder is displayed around the ray. All objects contained in the cylinder are selected as shown in Figure 16. The target is one of the N objects inside the cylinder. Then, the target is selected using a suitable refinement pointing technique. Considering a dynamic activation area [Grossman and Balakrishnan, 2005] augments the chances to select the target at the first attempt. Accordingly, the technique performs a volume selection using a cylinder .The transparent cylinder interaction technique does not require higher selection precision to point to the target. It is an accurate interaction technique and presents solutions to the problems resulting in case of small target and the shaking hand. The limitation of this technique is in the huge number of objects inside the cylinder. Consequently, the response time needed by the refinement pointing technique is immense and the performance degrades. The response time can be reduced by minimizing the number of objects selected by the ray by adjusting the cylinder volume and the number of objects inside. Thus, the user can detect the position of the target and make the adjustments to reach it. But this task is complicated since it requires more accurate movements and multiple adjustments of the ray.

The Spotlight [Liang and Green, 1993] interaction technique combines a ray cursor technique to a dynamic cone. After the detection of the 3D position and orientation of a bat, a ray is emitted. A transparent cone is displayed around the ray. Multiple objects inside the cone are hit and selected as shown in Figure 17. An anisotropic distance metric S is calculated to define the target. Rotating the bat by the user fingers along the x-axis makes adjustment of the cone's angle. Consequently, if the target was inside the cone and highlighted, by clicking the wand button, the selection
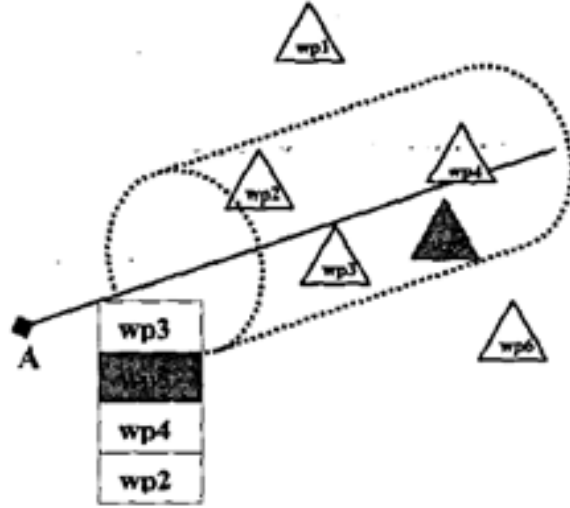
FIGURE 16 – The transparent cylinder interaction technique

is ended. Spotlight combines a ray cursor technique to a dynamic selection area technique. It does not require a higher selection precision to point the target. Spotlight is accurate and presents solutions to the problems resulting in case of small target, hand shake and the big number of selected objects in a dense environment.

2.3 Disambiguation Techniques for Multiple Selections

In case of the presence of several objects selected, a selection technique must be used to eliminate the ambiguity of the user selection.

The progressive refinement selection techniques are able to select any object regardless how small it is and how crowded the scene is. This can be done by allowing users to make refinements in a relaxed and imprecise manner. They have very high precision. These techniques allow the selection of objects inside a set of already pre-selected ones. The progressive refinement techniques require very low selection precision but multiple steps. The progressive techniques execute multiple selections but every step is fast and accurate.

The selection by progressive refinement consists of reducing, in every step, the
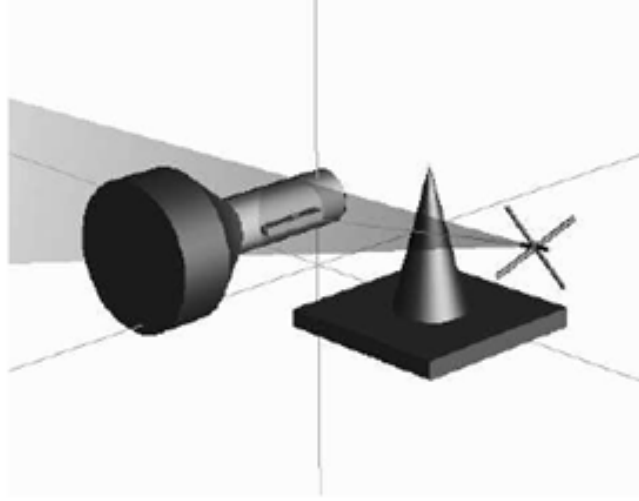
FIGURE 17 – The Spotlight menu interaction Technique

set of selectable objects until only the target remains. The type of the selection by progressive refinement depends on the criteria used to reduce the set of selectable objects. The criteria can be defined by a specific space (volume, surface, direction or position), a specific attributes (color, size or shape) or a specific subset of objects (from a list or menu instead of the environment).

Different strategies have been studied and different existing progressive refinement techniques are described in the literature.

2.3.1 Interaction based techniques

In order to perform a second step of selection to remove the ambiguities, different techniques based on interactions are proposed to identify and point the correct target from the group of objects.

Researches proposed an arm extension technique while using a non-linear transformation by the Go-Go technique [Poupyrev et al., 1996] or a linear transfer function by the HOMER technique [Bowman and Hodges, 1997].

The ray casting is used in multiple interactive techniques that have added a number of improvements and expansions to its basic definition. The Go-Go interaction

technique is a ray casting metaphor with nonlinear transformation, based on the metaphor of changing the length of the hand as shown in Figure 18 - right. The Ray casting is more effective for simple selection [Angus and Sowizral, 1995, Song and Norman, 1993, Mackinlay et al., 1990]. In the Go-Go technique [Poupyrev et al., 1996], the virtual ray is replaced by a virtual hand. Go-Go is an arm extension technique using a non-linear transformation. The virtual ray begins from the torso of the user and points in the direction of the user's hand. The user's hand is transformed into a virtual hand in the virtual environment as shown in Figure 18 - left. It is maintained on the virtual ray and can move on the entire radius depending on the distance of the physical hand from the user's torso. Go-Go is an interaction technique aiming to control the ray and move the target in a nonlinear transformation. This virtual hand can be used to navigate the target in depth and pinpoint the inaccessible target position. The Go-Go interaction technique is a natural action for touching nearby and far inaccessible target using the same technique. This technique needs multiple adjustments to point the target and the ray cursor requires a higher selection precision. Consequently, in case of a dense environment, this technique is not accurate to select an object. This technique presents a hand shake problem. The selection time depends on the target size : The selection speed decreases when the target size decreases because the user needs to make adjustments to accurately reach the target.

The Go-Go interaction technique therefore has some limitations for small targets and because of the limited length of the arm. In the latter case, it is difficult to reach a remote target in depth from the point of view of the user. The problem of the limited length of the arm is solved by a new version of GO-GO, the "Stretch Go-Go" technique or by the HOMER technique [Bowman and Hodges, 1997].

HOMER [Pierce and Pausch, 2002] is an interaction technique used to select objects any distance by using the ray casting metaphor and a linear transformation and to manipulate objects using a hand centered handling. The HOMER (Hand-centered Object Manipulation Extending Ray-casting) technique [Bowman and Hodges, 1997] combines the ray casting technique with the hand-centered manipulation technique. The user selects the object using the ray cursor technique : a ray is emitted, a vir-

FIGURE 18 – Left : The Go-Go Technique, Right : The vector R_r represents the real hand position and the vector R_v represents the virtual hand position

tual hand appears at the position of the target touched by the ray and the object is attached to the hand as shown in Figure 19. Then the target is selected. The second step consists of manipulating the target. The virtual hand moves the position of the object using a one to S linear transfer function between hand motion and object motion, where S is a scale factor. Therefore, the object can move in a sphere of radius less than S times the length of the physical hand. When the object is released, the hand returns to its position. HOMER can select objects at any position with no restrictions and perform easy selection and manipulation. Go-Go is limited by the real hand length and the difficulty of grabbing ; these limitations are solved by the HOMER interaction technique. HOMER presents two versions : direct and indirect HOMER. The direct version presents limitation for hand length. The indirect HOMER based on the button manipulation can select arbitrary positions not limited by the length of the arm and does not present restrictions. HOMER is more precise than the virtual hand in Go-Go. The ray casting selection is easier and HOMER is more precise than the virtual hand in Go-Go.

Unlike the Ray cursor and the curve with fixed length techniques where the user has to move the input device position many time in order to select a target, researchers propose variable length techniques. The Ray with variable length technique and the Curve with variable length technique move the destination end-point along the three dimensions in order to select its target and to minimize the response time.

FIGURE 19 – The HOMER (Hand-centered Object Manipulation Extending Ray-casting) Technique

Therefore, the pointing task became faster.

The Ray with variable length technique [Dang et al., 2003] is an interaction technique used in case of a target object placed after N aligned objects that the ray should pass across. The Ray with variable length technique identifies and points the correct target from the group of the aligned objects. This technique aims to change the ray's length by changing the position of the ray's endpoint (changing the position of the point B as shown in Figure 20). The closest objet to the endpoint is highlighted and selected when the user pushes a tracked wand button. In the ray with variable length technique, the user controls the origin position and the orientation of a virtual ray. He adjusts the input device position and the ray length, and therefore the endpoint position. After the detection of the 3D position and orientation of a tracker with a button, a virtual ray is emitted from the user through the target direction. This ray intersects multiple objects and the target. The last object intersected by the virtual ray is highlighted ; if the highlighted objet is the desired target, then the selection is confirmed by clicking the input device button. Otherwise, the user changes the ray length. Thus the endpoint position changes and the highlighted object changes to reach the target. This technique presents a hand shake problem.

The Curve with variable length technique [Dang et al., 2003] consists of displaying a curve with variable length in the 3D VE. The length of the variable curve is modified in the virtual environment to select a visible target. The curve endpoints are : the source, which is the input device (wand) position, and the destination, which should
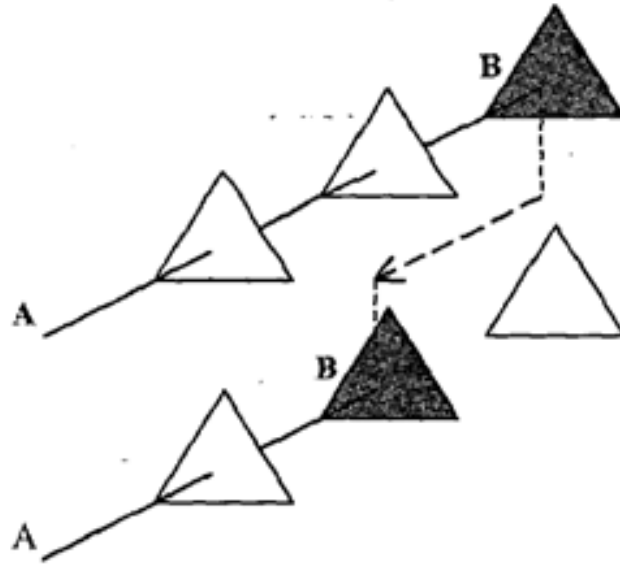
FIGURE 20 – The Ray with variable length Technique

point to the target in the VE, as shown in Figure 21. After displaying the curve, by pressing the wand button, the user moves the destination end-point along the three dimensions in order to select its target. Consequently, the curve length changes to touch the target with precision. Unlike the curve with fixed length, where the user has to move many time the wand position in order to select a target, a curve with variable length uses a joystick or a wand button to move along the three dimensions. Once the target is reached by the destination endpoint, it is highlighted and then selected by clicking on the input device button ; the selection is ended.

The Ray with variable length and the Curve with variable length techniques present some limitations in terms of small target since we have to make many displacements of the position of the input device to catch the target. Consequently, the response time is immense and the performance decreases. These techniques depend on the target size and become slower as the size decreases and present a hand shake problem.

Curve with fixed length and Curve with variable length travel in depth and face the challenge of avoiding objects placed in front of the desired target. In a crowded
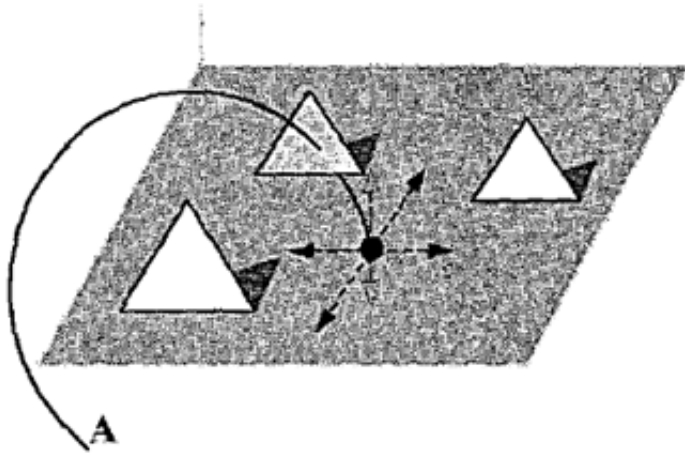
FIGURE 21 – The Curve with variable length technique

environment, since the curve techniques are based on the transparency of the environment, the user can detect the position of the target and make multiples adjustments and curvatures to reach it. Using a curved ray has the advantage of selecting a deep target without passing through the environment objects. The target is selected due to the flexibility of the curve in the case of the flexible pointer technique.

The Flexible pointer [Feiner, 2003, Olwal and Feiner, 2004] technique is visualized by a flexible curve. After the detection of the 3D position and the curvature of the user's hands a pointer is displayed. Two sensors placed on the hands of the user are able to detect the curvature of his hands. The flexible pointer direction is determined by vector formed by the two hands. The amount of curvature is defined by the orientation of each the hand as described by Figure 22. It is used to point visible, partially or fully occluded objects and to point around objects. An adjustment of parameters is made to make the end point of the pointer touching the target. Therefore, the target is selected and highlighted then the selection is ended. Accordingly, the flexible pointer technique circulates in depth and the final curve does not have necessary a well-known shape. Using an adjustable and non-fixed curve shape helps avoid targets in front of the desired one. This interaction technique presents some limitations in terms of small targets and the hand shake. On the other hand,

49

this technique allows an adjustment of the pointer to point the target efficiently, by changing the pointer parameters. Consequently, the response time is immense and the performance decreases. This interaction technique presents some limitations while using two hands and a displacement problem due to the curve travel in a non-transparent environment. On the other hand, it's an advantage because it solves the finger occlusion problem. In addition, this technique is confronted with the same problem as the traditional ray casting technique where several objects can still be traversed by the ray. Finally, as for point cursor technique, this ray casting technique suffers from a problem of precision. Indeed, it is complicated to precisely point an object that is far from the point of view of the user.
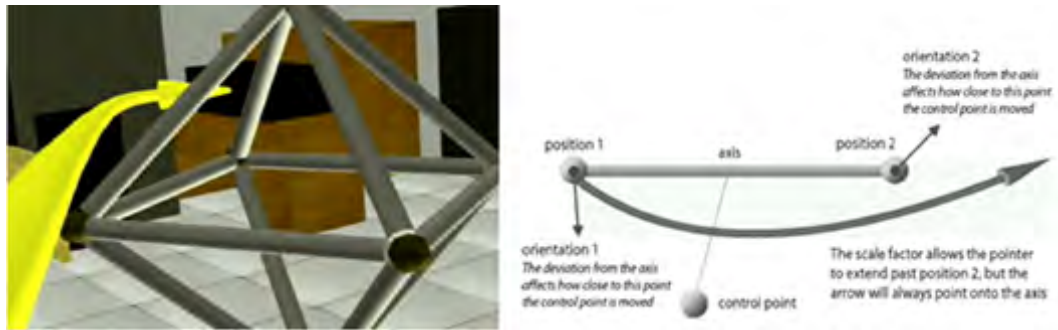
FIGURE 22 – The Flexible pointer Technique

Other type of strategies proposed to remove the selection ambiguity by using an additional interaction while using the pointing device to manipulate the ray or the cursor. The most common interaction is to browse the pre-selected objects : by circulating the objects and going directly from one to the other [Dang et al., 2003], or by manipulating a further pointer on the ray [Grossman and Balakrishnan, 2006] or by adding multiple actions. In the last case, PORT [Lucas, 2005] selects multiple objects and uses multiple movements and resizing actions to define the targets.

Another technique is the Depth ray [Grossman and Balakrishnan, 2006]. When the ray passes through multiple targets, different techniques are proposed to identify and point the correct target from the group of the aligned objects. The Depth ray [Grossman and Balakrishnan, 2006] interaction technique combines a ray cursor tech-

nique and a depth cursor. After the detection of the 3D position and orientation of a wireless mouse with a button, a ray is emitted and hits a number of aligned objects in the 3D VE. Selected objects are highlighted and a spherical cursor is displayed, which is a depth marker. The depth marker travels between the objects following the input device displacements. Moving the mouse toward and away from the 3D screen is transformed into cursor displacement to reach the target as described in Figure 23. Each time it comes close to an intersected object, the object becomes red. When the target becomes red, the selection is ended by clicking the input device button. The depth marker profit to identify and point the correct target from the group of the N aligned objects. Using a further pointer in the depth-ray technique [Grossman and Balakrishnan, 2006], requires multiple actions for adjusting a depth cursor to specify the target and select occluded objects. In this case, the pointing accuracy problems mentioned for a cursor based techniques will be identical. In fact, the 3D VE will have a depth circumstances and a large number of objects will be intersected by the ray, hence, it will be difficult for the user to accurately point the desired object.
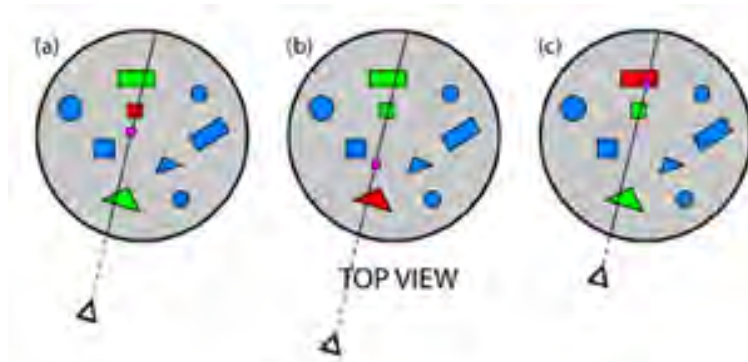
FIGURE 23 – The Depth ray interaction Technique : a) a pink depth marker is displayed ; the closer object is selected (the red square), b) moving the input device backwards selects the red triangle, c) moving the input device forwards selects the red rectangle

However, the Depth ray uses a continuous displacement of the input device that can cause modifications of the ray position. This limitation is solved by the Lock ray technique. The latter is a Depth ray version aiming of locking the ray then the depth

marker appears. This helps avoid confusion between selection and disambiguation phases [Grossman and Balakrishnan, 2004, Guiard et al., 2004].

The Lock Ray technique [Grossman and Balakrishnan, 2006] is an improved version of the Depth ray. It consists of using the Depth ray and a lock. After the emission of the ray, by clicking and holding the mouse button, the ray is locked and fixed. The 3D orientation and position of the beam will not be modified and the depth cursor is displayed at the center of the ray (cf. Figure 24). The lock provides a separation between the ray emission phase and the cursor adjustment phase. The lock eliminates the hand shake problem. Putting the second user's hand in a direction perpendicular to the ray removes the ray. The lock provides a separation between the ray emission phase and the cursor adjustment phase. The Lock ray uses a continuous displacement of the input device without making modifications of the ray position.
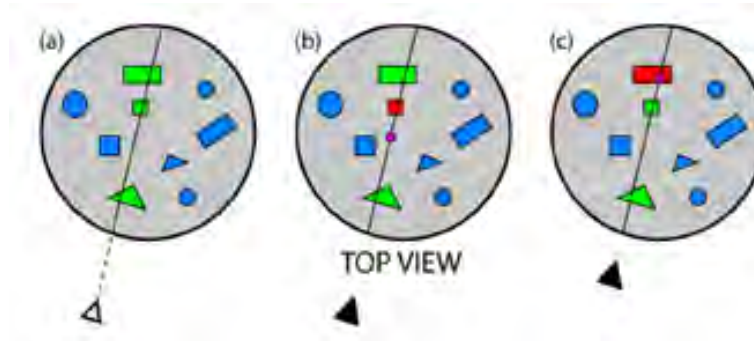
FIGURE 24 – The Lock Ray technique. Left : intersected targets are highlighted, center : the ray is locked ; the depth marker is displayed at its center, right : the depth marker is controlled by the input device than the closed target is selected

In the circulation technique, the N objects inside the set of pre-selected items are highlighted one after the other (cf. Figure 25 - right). When the target is highlighted, the selection is confirmed by clicking on the input device button. In the case of the Circulation technique [Dang et al., 2003], the act of circulating the objects and passing from object to another directly by pressing a button of the input device can limit the accuracy problems. In fact, the cursor displacement is discreet rather than continuous. This technique does not depend on the target size but requests a large

response time because we have to wait for the target to be highlighted when using the circulation technique. The circulation technique is used while attempting to select a target from a set of pre-selected items. The target is one of the N objects inside the set. Then, the target is selected using a suitable refinement pointing technique. The transparent cylinder interaction technique [Dang et al., 2003] and the transparent sphere technique [Dang et al., 2003] use the circulation technique (cf. Figure 25). The limitation of these combined techniques is in the huge number of objects of the pre-selected items. Thus, the response time is immense and the performance decreases. The response time can be reduced by reducing the number of pre-selected objects by adjusting the selected volume and the number of objects inside. The user can detect the position of the target and make the adjustment to reach it, but, this task is complicated since it requires more accurate movements and multiple adjustments. On the other hand this technique is accurate and does not require accurate movements or multiple adjustments in depth of the input device position to catch the target. Moreover, this technique does not present a hand shake problem.
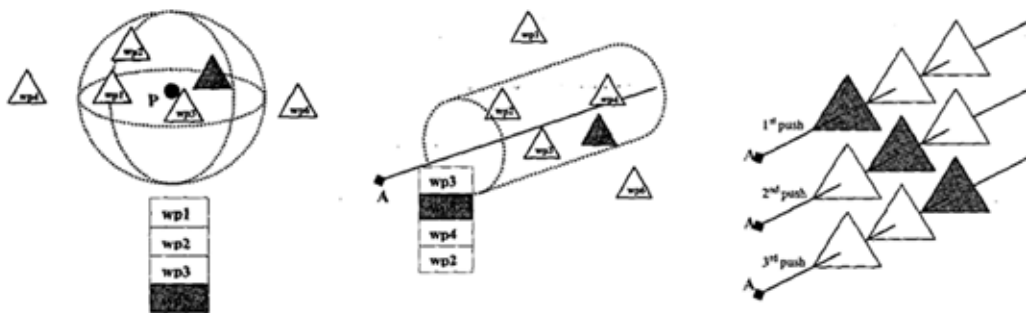
FIGURE 25 – left : the circulation inside a sphere, center : the circulation inside a cylinder, right : the circulation of aligned objects

Most of the presented techniques are only concerned with the selection subtask. The navigation subtask is not supported by the proposed techniques and requires different approaches. This complicates the interaction task in the 3D VE. The following techniques cover both the selection and the navigation subtasks.

Further strategies have been proposed to eliminate the ambiguity of these mul-

53

tiple selections. Wyss and al. [Wyss et al., 2006] proposed to add a navigation phase. For example the two ray selection technique and the shadow cone-casting technique [Steed and Parker, 2004] use two beams to select the desired object. They use continuous movement with the ray or a cone for casting to eliminate ambiguity and reach an accurate selection. The target is the object intersected by the two beams. However, this technique requires the use of two pointing devices. For an effective use of this technique, the user must have two views of the 3D VE away enough so that the intersection of the two rays will be one and unique object. Another way of implementing the navigation technique is by the Zoom used by Shift and Jump and Refine interaction techniques. Shift [Vogel and Baudisch, 2007] uses a refinement step for fine-tuning small target selection. Shift solves two problems : the selection ambiguity resulting from the finger's contact large area and the target occlusion by the finger. Shift is designed to solve the fat finger occlusion problem not to enhance pointing accuracy. Jump and Refine interaction technique [Hachet et al., 2007] is used for selection of sparse objects in a 3D environment on mobile phones. This technique is defined by two levels of cursor displacement while using the thumb-devices of the mobile phones. This technique is faster than moving a cursor due to the jump movement.

The Two-ray selection technique [Dang et al., 2003] proposes to emit two rays in different viewpoints. After the detection of the 3D position and orientation of a Tracked wand with buttons, a first virtual ray is emitted in the direction of the target. Multiple targets could be intersected in the 3D VE. Selected objects are highlighted. To select the target from those that have been intersected by the first ray, the user changes his point of view and emits a new ray. The target is located at the intersection of the two virtual rays as shown in Figure 26. Then the selection is confirmed by clicking the input device button. The Two-ray selection technique is a combination of two ray casting interaction techniques. The navigation phase changes the viewpoint to access the target. An algorithm is needed to memorize the objects crossed by the first ray. In case of a crowded environment, this technique is not accurate to select an object. The ray cursor requires a higher selection precision. Since, it requires more accurate movements and multiple adjustments of the ray

direction in depth. To solve this problem, researchers developed many techniques based on the ray casting technique by adding new feature like the dynamic selection area. By adding a cone, they define the shadow cone technique.
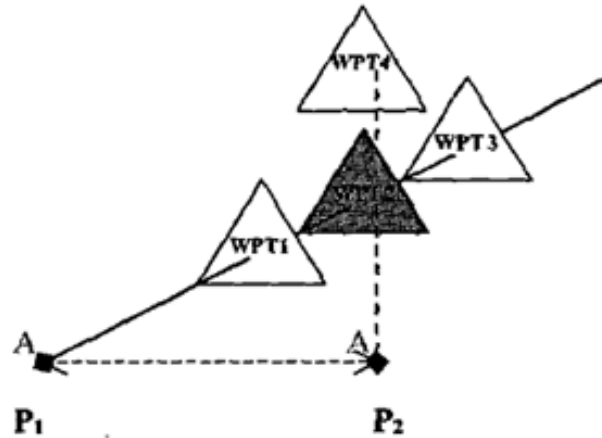
FIGURE 26 – The Two ray interaction technique

The shadow cone technique [Steed and Parker, 2004] proposes to emit two virtual cones in different viewpoints to select a group of objects, which are located at the area formed by the intersection of the two cone shape rays. The shadow cone technique has the same principle as the two-ray selection technique, but uses virtual cones instead of rays. After the detection of the 3D position and orientation of an input device, a first virtual cone is emitted in the direction of the target. Multiple targets could be intersected in the 3D VE. Selected objects are highlighted. To select the target from those that have been intersected by the first cone and highlighted, the user changes his point of view and emits a new virtual cone as shown in Figure 27. An algorithm is needed to memorize the objects crossed by the first cone. The target is located at the intersection of the two virtual cones. Then the selection is confirmed by clicking the input device button. The shadow cone technique is used to select a group of targets. The shadow cone technique is a combination of two ray casting

interaction techniques. The navigation phase changes the viewpoint. Using a cone instead of a ray reduces the hand shake problem. This interaction technique presents some limitations in case of small target since we have to make many displacement of the position of the input device to catch the target and the user will select a huge number of objects.
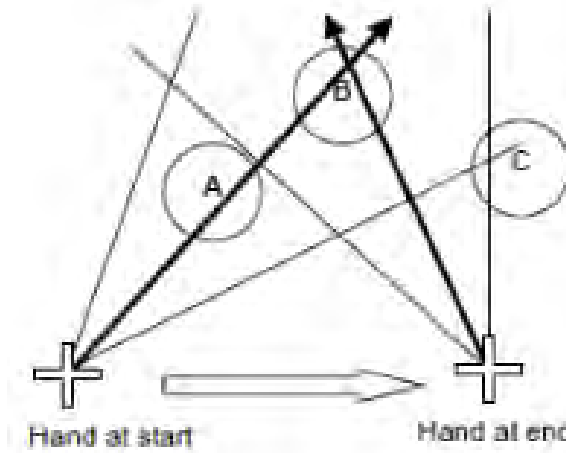
FIGURE 27 – The shadow cone interaction technique. A and B are highlighted at the start than A is dropped after the modification of the hand position

The Smart ray [Grossman and Balakrishnan, 2006] proposes to emit several rays in different viewpoints. After the detection of the 3D position and orientation of a wireless mouse with a button, a ray is emitted and hits a number of objects in the 3D VE. Selected objects are highlighted, their color became green and a spherical cursor is displayed at the center of the selected objects. This circle refers to a weight for each selected target. The target having the greater weight is in red. To adjust the weight of the target, the user has to change the position and the orientation of the mouse and send another ray in the 3DVE. For each ray emitted, the technique adds a weight to the selected objects and the weight are adjusted in function of two conditions : first, if the target is hit several times by the ray and second, if the ray is closer to the center of the selected object. The red target with the highest weight can be selected. Then the selection is confirmed by clicking the input device button

(cf. Figure 28). The Smart Ray is a combination of multiple rays and a weight. This technique is based on the user intervention and a prediction based on the display of a weight for each target, the prediction is based on an algorithm and the history of the ray movement and not on the metric criteria or the intersection of two rays. A constraint is presented in the possibility of error in case of errors in the weight values (lack of algorithm). Other limitations are presented in case of small target and the hand shake. The Smart ray has poor performance and not recommended with respect to response time and errors.
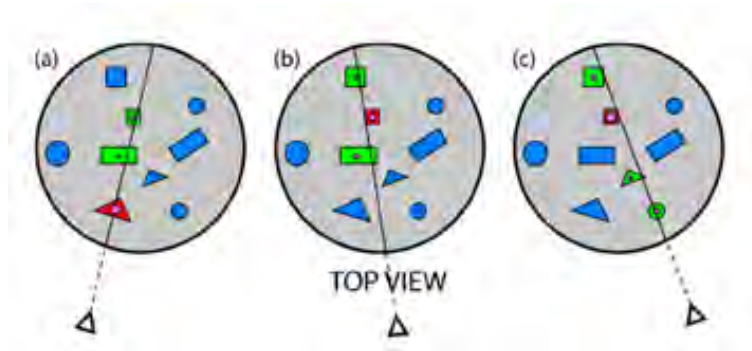
FIGURE 28 – The Smart Ray interaction Technique

Researchers introduce the dynamic selection area [Grossman and Balakrishnan, 2005] to reduce the number of elements of the pre-selected items. The dynamic selection area can have a specific well-known geometric form like the cylinder used by the transparent cylinder technique, a cone used by the Spotlight or a sphere used by the transparent sphere technique or a form that morphs in function of the environment density.

The Spotlight [Liang and Green, 1993] interaction technique combines a ray cursor technique to a dynamic cone. Rotating the bat by the user fingers along the x-axis makes adjustment of the cone's angle. The spread angle of the cone is adjusted to minimize the number of objects inside it. The area around the object is gradually shrunk to minimize the total number of the selected objects. Consequently, if the target satisfied a special numeric condition (an anisotropic distance metric S is defined to emulate the spotlight effect) then it will be inside the cone and highlighted.

Spotlight combines a ray cursor technique to a dynamic selection area technique. It does not require a higher selection precision to point the target. Spotlight is accurate and presents solutions to the problems resulting in case of small target, hand shake and the big number of selected objects in a crowded environment.

The hybrid technique formed by the combination of a ray cursor and a transparent sphere technique [Dang et al., 2003] adds a dynamic sphere volume at the ray endpoint to augment the selection area and minimize the pointing problem. In this technique, the user controls the origin position and the orientation of a virtual ray by adjusting the input device position. After the detection of the 3D position and orientation of a tracker with a button, a virtual ray is emitted from the user through the target direction. At the ray endpoint, a transparent sphere is displayed and the target is one of the N objects inside the sphere. Based on the number of objects inside the dynamic selection area, researchers propose a technique to reduce ambiguity and to select one target. On the other hand this technique is accurate and does not require accurate movements or multiple adjustments in depth of the input device position to catch the target. This technique does not present a hand shake problem. This technique does not depend on the target size. The large response time is solved by a selection algorithm.

The bubble cursor [Grossman and Balakrishnan, 2005] is a 2D technique that enhances the cursors area by dynamically resizing, growing or shrinking the activation area of a circular cursor such that it only contains one object to be selected. Depending on the surrounding targets, the bubble cursor morphs to cover and isolate one target and solves the problem of multi target selection. Vanacken et al. [Vanacken et al., 2009] present a 3D extension of the bubble cursor, by proposing the use of a sphere instead of a circle. The hybrid technique formed by the combination of a ray cursor and a bubble cursor technique [Grossman and Balakrishnan, 2006] provides a shape transformation and warp to select one target in a dense volume. In this technique, the user controls the origin position and the orientation of a virtual ray by adjusting the input device position. After the detection of the 3D position and orientation of a tracker with a button, a virtual ray is emitted from the user through the target direction. At the ray endpoint a bubble cursor is displayed. The bubble

cursor is a semitransparent disk having a crosshair cursor inside as shown in Figure 29 –a). To select a target, the user moved the crosshair cursor towards the target direction. Then the cursor radius changed and the bubble cursor is transformed : its size is dynamically changed to wrap the target when it is not fully contained in the main bubble. The bubble cursor technique is based on the dynamic resizing of the cursor area. The selection is ended by clicking the input device button. There is always exactly one target in the disk. The bubble cursor is dynamically transformed depending on the objects around to wrap the target and to select only one target as shown in Figure 29 –d). The two techniques reduce target acquisition times but perform worse in case of a hand shake in a crowded environment, since the small hand movement will cause the deviation of the ray and therefore the cursor. Consequently, the cursor quickly morphs to select a new target.
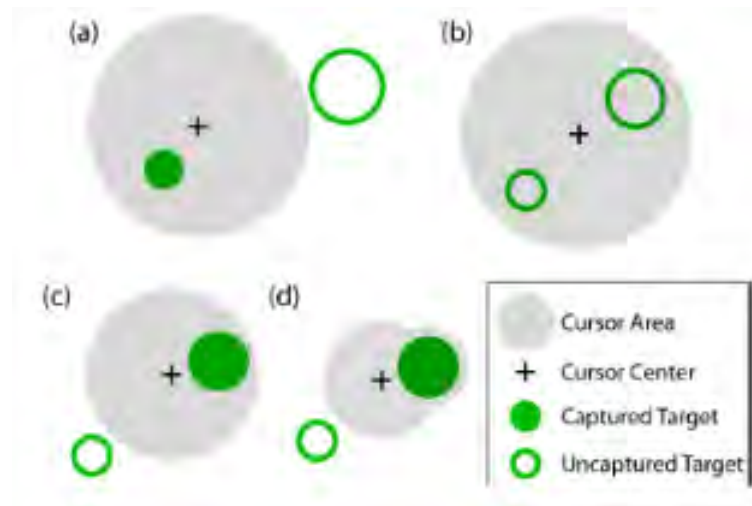
FIGURE 29 – The bubble cursor technique

2.3.2   Techniques based on the rearrangement of the object disposition

Other studies suggest the use of an additional component, usually a menu [Dang et al., 2003, Liang and Green, 1993] or directly re-arranged objects on the screen [Grossman and Balakrishnan, 2006], to select the desired item among different pre-

selected objects. All pre-selected objects are displayed as an item in a menu, and the user has to only select the target. For example, the flower ray technique [Grossman and Balakrishnan, 2006] uses a ray casting technique to select a set of targets and a menu, in the second phase, to display the set of selected targets. However, this technique has two major weaknesses : firstly, how to refer to an object of the scene as text in a menu [Dang et al., 2003] ? Secondly, this system can work with few items in a sparse environment. Once the number of pre-selected objects becomes too large, the menu takes up too much space on the screen and completely hides the 3D VE. It is even possible, given the small size of the smartphone's screen, that not all the selected objects can be represented.

Proposed techniques consist of using a menu to rearrange the preselected objects and then to select a target from the N pre-selected objects. When the ray passes through multiple targets or a set of targets is selected by a volume cursor, different techniques are proposed to identify and point the correct target. The menu technique consists of displaying objects of the set, and then the target is selected from the menu. Different techniques exist : a 2D flower menu in the Flower Ray technique (which is a 3D extension of the splatter technique) ; a virtual sphere menu in the Daisy menu technique ; a circular menu in the Ring menu technique or a list menu in the Floating menu technique or the Transparent Sphere Technique.

The Flower ray [Grossman and Balakrishnan, 2006] interaction technique combines a ray cursor technique and a 2D flower menu. After the detection of the 3D position and orientation of a wireless mouse with a button, a ray is emitted and hits a number of aligned objects in the 3D VE. By pressing and holding the mouse button the selected objects appear as a 2D Flower menu according to the depth criteria as shown in Figure 30 up b). To select the target, the user moves the 2D cursor to the target direction then the object becomes highlighted (the red object in Figure 30 up c)). The selection is ended by clicking the input device button. The Flower ray interaction technique is a 3D extension of the splatter technique [Ramos et al., 2006] (cf. Figure 30 down). This interaction technique is used in case of a target object placed after N aligned objects that the ray should pass across. The Flower ray technique identifies and points the correct target from the group of the aligned objects and

keeps information about the depth locations and positions that are not considered in other techniques. This interaction technique transforms the 3D problem into a 2D menu problem. In fact, searching for the target in the 2D menu minimizes errors but on the other hand it will lead to a delay and an immense response time. The response time can be reduced by reducing the number of objects selected by the ray. In addition, a problem results while pointing small targets increasing the response time. In fact, we have to make many displacements of the position of the input device to catch the target. Consequently, the response time is immense and the performance decreases.
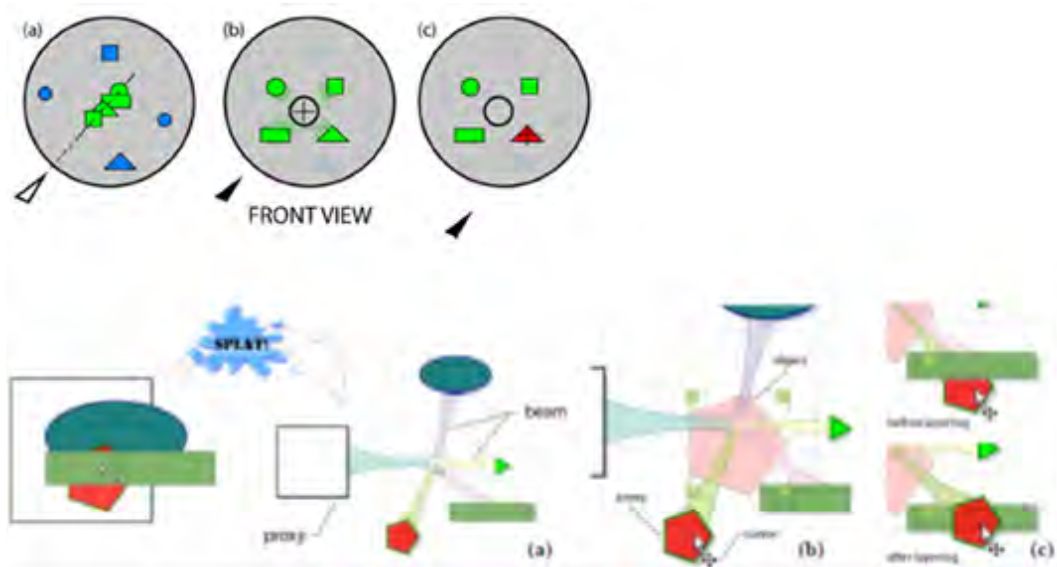
FIGURE 30 – up : The Flower ray interaction Technique, down : the splatter technique

The Daisy menu [Liang and Green, 1993] interaction technique combines a ray cursor technique and a virtual sphere menu. After the detection of the 3D position and orientation of a bat, a ray is emitted. A number of objects are hit and selected. A daisy menu is displayed. It is a spherical menu : the selected objects are displayed in a sphere form at the center, a cursor having the same orientation as the bat and a conic basket at the 3D cursor oriented towards the user. This is shown in Figure 31. Rotating the bat by the user fingers makes rotation of the daisy menu. Consequently,

61

the menu objects moves and passes through the basket. The object inside the basket is highlighted then the selection is ended by clicking the input device button. This interaction technique is used to select a target object from a set of N aligned objects. In addition to losing data about the depth locations and positions, this interaction technique has other limitations such as the selection of small sized targets and the user's hand shake. Not to mention, the complications caused by the rotation of bat wire : Selecting a target requires many displacements of the input device's position and consequently, the response time becomes immense and its performance decreases.
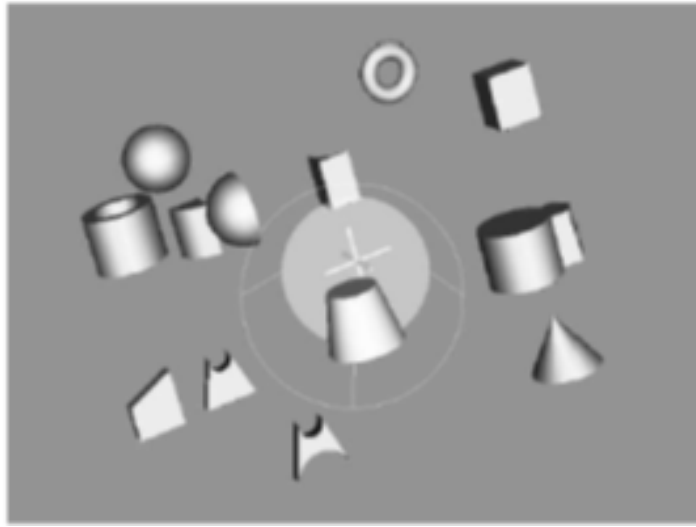
FIGURE 31 – The spherical menu of the Daisy menu interaction Technique

The Ring menu [Liang and Green, 1993] interaction technique combines a ray cursor technique and a circular menu. After the detection of the 3D position and orientation of a bat, a ray is emitted. A number of objects are hit and selected. A Ring menu is displayed as a circular menu where the selected objects are displayed along the contour of the circle. At the center, a 3D cursor and a transparent gap in the direction of the user are displayed as shown in Figure 32. Rotating the bat by the user fingers along the x-axis creates a rotation of the elements of the ring. Moreover, rotating the bat by the user fingers following 3 axes creates a rotation of the ring ; consequently the menu objects moves and passes through the gap. The

object inside the gap is highlighted and the selection is ended by clicking the input device button. The Ring menu is a ray cursor technique with a menu. This technique loses information about locations and depth positions. Furthermore, it can become inadequate when selecting a small target due to the user's hand shake. Another limitation is the bat wire rotation : the positions of the input device have to be altered many times prior to capturing a target, resulting in a longer response time and a decrease in the performance.

FIGURE 32 – The ring menu of The Ring menu interaction Technique

The Floating menu [Dang et al., 2003, Ramos et al., 2006] interaction technique combines a ray cursor technique and a list menu. After the detection of the 3D position and orientation of a Tracked wand with buttons, a ray is emitted. A number of objects are hit and selected. A list menu is displayed. It is an ordered list that displays all the selected objects within a rectangular frame. The objects are ordered vertically according to their z-coordinate in the 3D VE as shown in Figure 33. To adjust the z value of an object, the user has to click and drag the object window to another position in the list. By clicking the wand button, a menu item is highlighted then the selection is ended by a second click. The Floating menu is a familiar method based on the use of a ray cursor technique and a menu technique. However, the Floating menu technique also loses information about the depth locations and positions. It presents difficulties to distinguish between objects of similar colors, sizes and shapes, since, by displaying objects in a menu, the position and the depth values

will be lost. The selection of similar objects with different sizes will be also confusing, if the label of the object in the menu does not precise the object size.
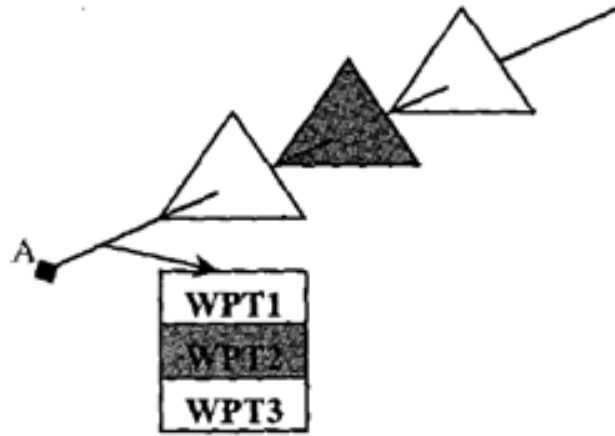
FIGURE 33 – The list menu of The Floating menu interaction Technique

The list menu is also used by the Transparent Sphere Technique. This list menu contains all the targets that are inside the sphere as shown in Figure 34. The user chooses the target from the menu by pressing the wand's button. This technique does not depend on the target size, but requests a large response time since we have to choose the target from a long list.

Another type of strategies, proposed to remove the selection ambiguity, suggests a progressive algorithm to reorganize objects disposition.

The Virtual pointer metaphor [Steinicke et al., 2006] is a ray casting metaphor, using a dynamic selection area. To refine the selection of the target from the pre-selected set, this technique performed a menu or a selection algorithm. Finally, the bendable ray pointing the target is visualized by a quadratic Bezier curve as shown in Figure 35 - left. This technique consists of releasing a ray in the target direction. A virtual ray is displayed in the 3D VE indicating the direction of the virtual pointer. The second step consists of calculating the distance between the environment objects and the virtual ray. The calculated distance is displayed in a menu. The goal of this technique is to find the active object by searching for the object having the minimum
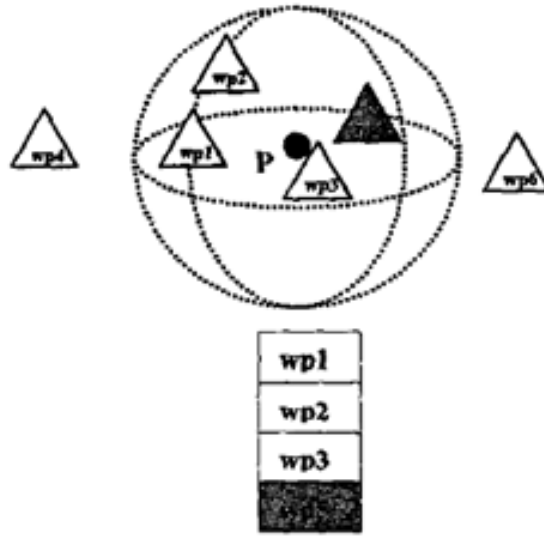
FIGURE 34 – The Transparent Sphere Technique using list menu

distance to the virtual ray. To reduce the number of object displayed on the menu, another method consists of searching for the active object inside geometric shapes like a cylinder or a cone and making adjustment for the selection area if necessary. After finding the target, a curve is displayed showing the virtual pointer hitting the active object as shown in Figure 35 -right. Then the selection is ended. It is used to select one target independently of its size or its position (near or far from the virtual ray). The selection depends on the criteria of selection and the selection algorithm. This technique presents solutions for the hand shake problem. However, these techniques face the same problem as the other rays in the case of crowded 3D VE since ; it becomes difficult to get around all objects to point a particular target.

A progressive refinement algorithm reduces the completion time by eliminating a range of selected items in each step as for the snapping technique. The snapping technique presented by Haan et al. [De Haan et al., 2005] estimates which object the user wants by using a selection volume to calculate and accumulate scores over time for each object.

Finally, the last way to facilitate the pointing of a target from the preselected

65
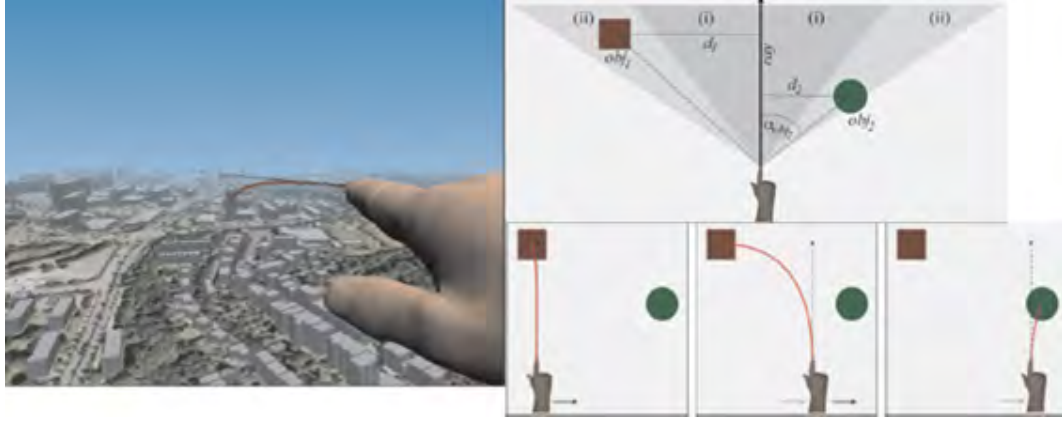
FIGURE 35 – The Virtual pointer metaphor

subset is by rearranging the elements on the screen as proposed by the SQUAD [Kopper et al., 2011], Expand [Cashion et al., 2012] and Semantic 3D pointing [Elmqvist and Fekete, 2008] Techniques. SQUAD and Expand techniques dispose on 2D plane all the preselected objects. This has the advantage of facilitating the pointing of elements and in particular on a touchscreen. However, if the number of items to display is too large, they cannot be all displayed at the same time on the screen. On the other hand, the rearrangement of the elements can disturb the user. It adds cognitive charges as the user should find again the desired item in this new organization of the space. Thus, the fact of rearranging the elements results in the loss of the original object position and in particular their positioning relative to other objects. It is important to maintain this positioning if the elements are dependent on each other.

SQUAD technique [Kopper et al., 2011] is designed by the combination of a ray cursor and a transparent sphere. It adds a dynamic sphere volume at the ray endpoint to augment the selection area in a crowded environment. In this technique, the user controls the origin position and the orientation of a virtual ray by adjusting the input device position. At the ray endpoint, a transparent sphere is displayed and the target is one of the N objects inside the sphere as shown in Figure 36 - left. A proposed technique consists of using progressive refinement technique to

select a target from the N selected objects. A four-item menu displays the visible and the occluded objects inside or intersected by a sphere as shown in Figure 36 -right. SQUAD proposes a new arrangement of the selected items to refine the initial selection by selecting a subset from the four-quad containing the target, until the target is the only one left. This hybrid technique is formed by the combination of the ray cursor, the transparent sphere, a menu and a refinement criterion to select a group of objects. SQUAD technique starts by displaying two objects on the screen : a yellow circle in the center and the red target. Using a wand with a button, the user clicks at the yellow object of the scene. Consequently, the yellow object disappeared and the screen will be filled with distractor objects. Then, the pointing task begins. The cursor on the 3D VE changed on a sphere and the objects inside or intersected by the sphere are highlighted (with a yellow border) for selection. By pressing and holding the trigger button, the selected objects are distributed among four quadrants on the screen on a quad-menu. Each quadrant has the same number of objects. To select the target, the user points the quadrant containing the target and clicks the button. A refining phase occurs ; the selected objects are distributed on a quad-menu ; the user repeats the pointing anywhere in the quadrant containing the target. The selection is ended by clicking the quadrant when the desired object is the only one left. This technique is accurate and does not require accurate movements or multiple adjustments in depth of the input device position to catch the target or at any moment during the selection task. It does not present a hand shake problem, and does not depend on the target size. Moreover, it is faster for small object selection. SQUAD is useful in case of interaction with very large data sets, such as astrophysical or atomic datasets, in addition to the supermarket task [Figueroa et al., 2010]. The limitation of these combined techniques is in the huge number of objects displayed in the menu. However, SQUAD uses two distinct refinement phases. In the first phase, the use of a volume selection area reduces the number of selected objects to the highlighted objects. In the second phase, SQUAD allows the selection a subset from the four-quad. It is a fast technique and reduces the number of task to $log_4 n$ where n is the initial number of items. This interaction technique transforms the 3D problem into a 2D menu problem, searching for the target in the 2D menu minimize errors but

67

will lead to a delay reduced by the refinement selection. SQUAD presents a limitation when selecting two or more similarly looking objects placed in different 3D positions : by transforming the 3D space into a 2D space the user will lose information about the target original position and will be confused in his choice. However, it does not keep information about the depth locations and positions and cannot be used in case of selection tasks that depend on object location. Consequently, SQUAD faces the problem of losing original context.
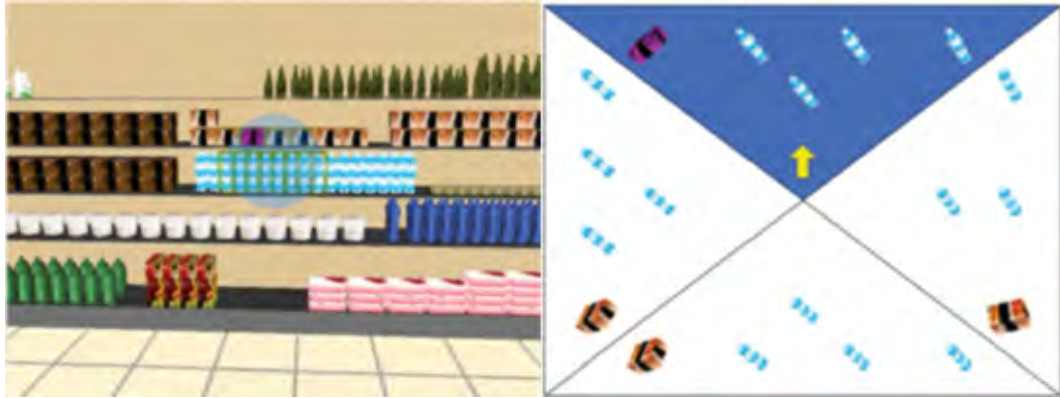
FIGURE 36 – The SQUAD interaction technique. Left : Sphere-casting, Right : Quad-menu

Inspired by SQUAD interaction technique, Expand technique [Cashion et al., 2012] has been developed as a 3D selection technique for dense and dynamic game-based VEs. The Expand technique is a variation of SQUAD that adds the possibility of zooming in on the selected targets than places the target objects in a grid for the user to choose from. The quadrant arrangement used in SQUAD was replaced by a virtual grid displayed on the screen and dynamically arranged depending on the number of selected objects as shown in Figure 37. Expand is faster than SQUAD, and is significantly faster with higher object density, but SQUAD has a lower error rate. However, during the refinement step, the grid representation of the selected objects does not correspond to the original 3D position. Then, a possible wrong and imprecise selection may occur while trying to select an object from a set of similar looking objects.
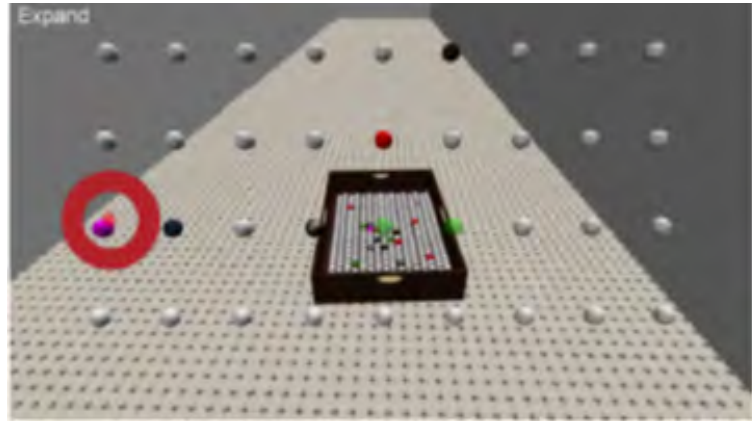
FIGURE 37 – The Expand Method

Based on quad-tree representation and spatial query algorithm, Semantic 3D pointing [Elmqvist and Fekete, 2008] is used to point an object in 3D VE. This technique consists of shrinking empty space and expanding target sizes on the screen by dynamically adjusting the ratio between the cursor movement on the screen and the movement of the input device. A non-linear scale function is used to control the ratio. The input of the scale function is the distance measured from the cursor to the closest target as shown in Figure 38 -right. In a dense area, the distance is small and the cursor movement is slowed down to allow for higher accuracy. In an empty space, the distance is big so the cursor movement is speeded up. Semantic 3D pointing improved accuracy but not the completion time due to the navigation phase. In case of a crowded environment and a set of aligned objects, the distance to the nearest object is computed (it is the distance to the first object) to reach the nth object. The algorithm has to calculate n distances and to circulate all the objects to reach the target because the cursor cannot jump directly to the target. This increased the task completion time.

The refinement techniques based on a ray emission are slow and error prone because they need a high precision on selection and ray emission. Due to the hand tremor the origin and the orientation of the ray are difficult to be controlled by the user and the problem is enlarged in case of smaller and farther targets. The refinement

69
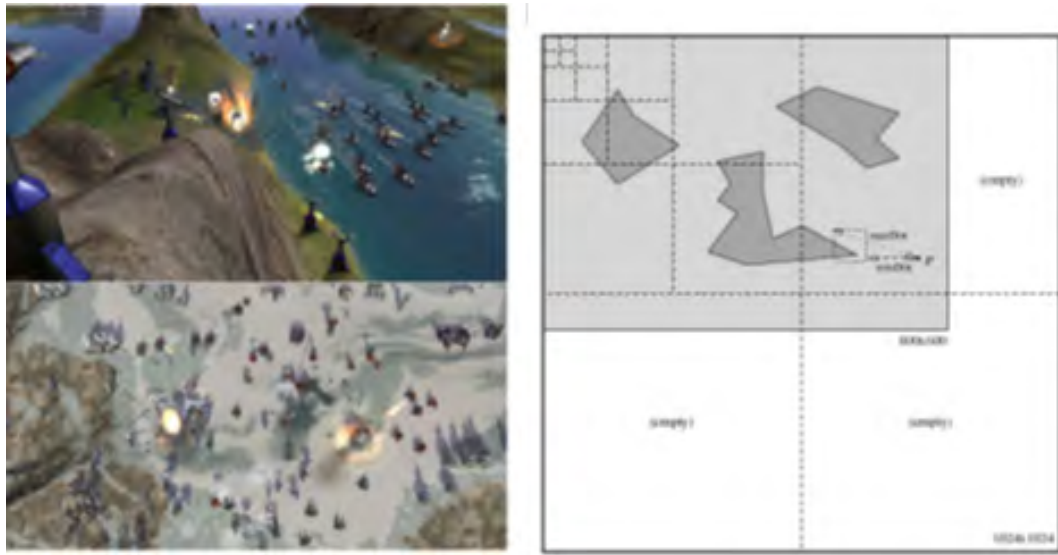
FIGURE 38 – The Semantic 3D pointing Technique

techniques based on volume selection are slow because they need a navigation phase or multiple accurate movements and calculations or a menu that can occlude the environment. The efficient way of using a refinement technique will be after using a dynamic activation area. A user experiment [Delamare et al., 2013] showed that the refinement techniques that require maintaining the users' focus on the physical world are performing the refinement techniques that require switching their focus to a digital representation of the physical objects. Due to the fact that the user needs to shift its focus of attention when a digital representation is involved than to concentrate on one and only space configuration, the completion time and error are augmented.

2.3.3 Validation

Validation consists of making an action to confirm and validate the operation of the 3D selection [Bowman et al., 2001c, Bowman, 2002]. When the pointer reaches accurately the target, the target is the only selected item ; therefore, the validation consists of giving commands to the system by the user to complete the 3D selection

task and to confirm the operation. After the validation subtask, the desired target is selected and can be handled and manipulated. Validation designates communication between a user and a system independently of the virtual environment [Hand, 1997]. It is a task consisting of applying a command to change the state of the system or the interaction process [Bowman et al., 2001a]. Therefore possible validations can be accomplished by using traditional WIMP graphical user interfaces for 2D applications, Post-WIMP user interfaces [Van Dam, 1997] or another techniques combining WIMP and Post-WIMP [Kurtenbach and Buxton, 1994, Jankowski and Hachet, 2013].

Validations are performed by using the traditional 2D "point-and-click" WIMP graphical user interfaces defined by interfaces based on windows, icons, menus, and a pointing device, typically a mouse. Since, WIMP interfaces were introduced in early 80's they are still the leading type of interaction style for the fact that giving a command in a WIMP interface is simple and became intimately familiar to users. The set of WIMP user interface modules such as : windows, icons, menus, and a mouse cursor are easy to use to perform a "point-and-click" and then a validation process. The WIMP interfaces are designed and generally used for 2D applications. For 3D applications, however, the navigation and the pointing subtasks are functions accessible through a 3D user interface ; the validation process can be controlled through a conventional 2D GUI.

A 2D pointing device aims to control 2D applications with their interface components using WIMP. However, controlling 3D applications with 2D control widgets was not so simple and natural [Van Dam, 1997]. Andries van Dam proposes the development of new user interfaces and defines "post-WIMP". Post-WIMP user interfaces was founded on Gesture-based interfaces, for example, gesture and speech recognition, eye/head/body tracking, etc. [Van Dam, 1997]. They contain at least one interaction technique different from the classical 2D widgets [Van Dam, 1997]. Gesture-based interfaces were also introduced to games. For example, in the Black&White game [Jankowski and Hachet, 2013], actions can be implemented by making hand gestures. Gesture based system control was applied in sketch-based 3D modeling interfaces of SKETCH [Zeleznik et al., 2007] and ShapeShop [Schmidt et al., 2006].

71

Other validations consist of combining WIMP and Post-WIMP [Jankowski and Hachet, 2013, Kurtenbach and Buxton, 1994] while using "marking menus" [Kurtenbach and Buxton, 1994]. To perform a menu selection, the user can either pop-up a radial/pie menu, or make a straight mark in the direction of the desired menu item without popping-up the menu. Autodesk Maya 3D modeling software adopts Marking menus.

3 Selection task on Smartphones

The majority of the 3D VE studies has focused on the use of traditional display surfaces (such as the computer screens and the projections on large surfaces) or the display on more appropriate 3D environments [Grossman and Balakrishnan, 2006] and more immersive. The display and the interaction on 3D VE using a Smartphone is more difficult because the displayed surface offered by the device's screen is much smaller. Thus, two solutions are offered to display a complex 3D VE on these small devices : either to display a part of the 3D VE to retain a rather large object size ; either to display the entire 3D VE and reduce the objects size.

In consequence, researchers in 3D virtual environment have developed new techniques and metaphors for 3D interaction to improve 3D application usability on mobile devices, to support the pointing task and to face the problems or factors affecting pointing performance.

Researchers have developed a variety of interaction techniques to point and manipulate a target in a 3D VE while using a Smartphone.

3.1 2D Selection task on Smartphone

This category of interaction techniques consists of using mobile devices as a "window into the virtual world" by providing a real time 3D rendering of dense virtual scenes or displaying maps for navigation. In this case, an Augmented Reality video or a map for navigation is seen through a hand-held mobile device. The mobile device provides a window into the VE or maps to display a part of the environment on a lar-

ger size to enlarge the view of a specific part of the total environment. However, the interaction in handheld AR environments is restricted to a 2D pointing and clicking via the device's touchscreen.

Different interaction techniques are described in the literature to propose an accurate technique for pointing a target in an environment displayed on a small screen of a mobile device. But the majority of these techniques are designed to solve the occlusion problem residing by pointing with a fat finger.

Different techniques are proposed; few examples are ZoneZoom [Robbins et al., 2004], Escape [Yatani et al., 2008], Shift [Vogel and Baudisch, 2007], Offset cursor [Potter et al., 1988], POI [Lehtinen et al., 2012] (cf. Figure 39), TapTap and MagStick [Roudaut et al., 2008]. These techniques solve the 2D density and the problems of occlusion and accuracy resulting from the fat finger touch but do not solve the 3D pointing problems. In such cases, navigation movements in the virtual environment will be necessary.

FIGURE 39 – POI interaction technique : left) the viewport is updated according to turning/rotation, and switched to the top-down view if turned to horizontal orientation, right) the user is operating the exocentric view by panning.

ZoneZoom [Robbins et al., 2004] is an easy and quick navigation technique, used

73

to navigate around a huge detailed two-dimensional information space on a smartphone. ZoneZoom has been improved to SmartPhlow [Horvitz et al., 2012] which is a traffic map monitoring application. Using ZoneZoom, the information space is divided up into nine sub-segments or nine non-regular divisions of zoom-levels. Each division or sub-segment is accessible via the number keypad. ZoneZoom adapts recursive view segmentation for map navigation for Smartphones. To zoom into one of these sub-sectors, the user clicks on the number key that corresponds to that sector of the screen (as shown in Figure 40) and performs a zoom back out to the parent view by pressing the same number key again. ZoneZoom is an interaction technique based on recursive zoom actions to navigate huge information on a smartphone by displaying a part of the 2D environment on the device's screen. For each iteration, a part of the environment is displayed on a bigger and clearer manner. This technique proposes a 2D navigation interaction to reduce the 2D density and the fat finger accuracy problems, but it is not useful for a 3D pointing task.

Figure 40 – ZoneZoom interaction technique

Shift interaction technique [Vogel and Baudisch, 2007] is designed to allow small dense target selection with finger error prone on mobile devices. Shift solves two problems : firstly, the selection ambiguity resulting from the finger's contact large area to define the selection point in case of a fat finger and secondly, the target occlusion by the finger. In case of ambiguous target selection, when the screen is

touched by the user, Shift creates a callout and places it in a non-occluded screen location. The callout displays a copy of the occluded screen area and shows a pointer representing the selection finger point. Keeping the finger on the screen, the user moves his finger to transport the pointer into the target then validate the selection by lifting his finger (as shown in Figure 41). Shift only creates the callout when necessary ; over large targets when occlusion is not a problem no callout is created and Shift acts like a simple touch action. Shift is designed to solve the fat finger occlusion problem not to enhance pointing accuracy. It uses a refinement step for fine-tuning small target selection. By using Shift in a 3D VE in depth direction, adjustment and accuracy problems will rise. Therefore, we will have the point cursor technique problem.
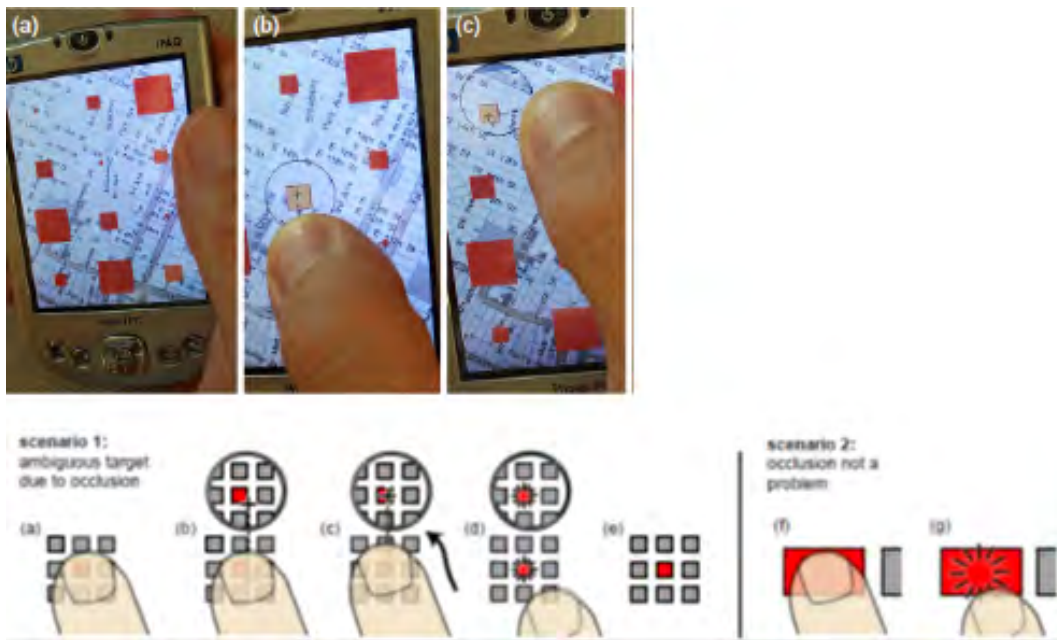
FIGURE 41 – Shift interaction technique and design

TapTap and MagStick [Roudaut et al., 2008] are 2D interaction techniques for target pointing on mobile devices with small touch-screens using the user thumb. These two techniques are designed to solve the problems raised by the thumb inter-action. The main concerns of this way of pointing are the screen accessibility, the

visual occlusion and the accuracy. TapTap [Roudaut et al., 2008] uses a two-step zooming strategy (cf. Figure 42). It is based on two taps ; the first tap defines the focus area on the screen. This area is magnified and displayed as a popup on the center of the screen. The second tap consists of selecting the target in the popup or canceling the selection by taping in an empty space. TapTap presents some constraints for example : the environment will be occluded by the popup, or the zooming factor could be inadequate with target size and the target still inaccessible precisely.

FIGURE 42 – From left to right : TapTap and MagStick interaction techniques, Tap-Tap Design, MagStick Design

MagStick depends on a telescopic stick that controls a "magnetized" cursor. To point a target, the user defines a reference point by pressing the screen and two-part stick by moving his thumb. The two-part stick derives from the reference point and has two end points : the thumb current position and the cursor location (cf. Figure 42). The cursor reaches the target by dragging the thumb in the opposite direction of the target this movement augments the two-part stick length and make the cursor close to the target. The target attracts the cursor as if it was "magnetized". Finally the user thumb is released to finalize the selection. The dynamic adjustment of the cursor avoids its fine positioning. However, the cursor movement in the opposite direction of the thumb is efficient for eliminating the finger occlusions but it is confusing to travel in the opposite direction of the desired one. Another constraint is to well position the reference point because the stick's two parts have the same length and grows in the same manner. Accordingly, the user has to take into consideration : the position of the reference point to touch the target correctly, the length of the stick and the dragging time of the thumb. The previous described techniques are developed for

mobile AR based on the use of the multi-touch abilities of the device. However, selecting a 3D object by using multi touch input induces problems resulting from the interaction through only one available hand (two handed selection techniques are not applicable) and from the inaccuracy of the user's finger touch.

3.2   3D selection task on Smartphone

This category of interaction techniques consists of displaying a part of the 3D VE on the Smartphones screen to retain a rather large object size [Bergé et al., 2014a]. This case offered a solution to the problems resulting from displaying a huge and complex 3D VE on these small devices, but presented some restrictions. If the 3D VE is partially displayed on the screen, the user does not have a whole perception of the scene. Consequently, the object to achieve is not always visible. Therefore the user is required to achieve a prior navigation phase to find the desired object and then to point it. This type of techniques is based on a navigation phase to display a part of the huge environment on the Smartphones display. Concentrating on a specific part of the total environment reduces the density and clarifies the view.

The majority of the interaction techniques of this category improve the navigation subtask of the pointing task to reduce the 3D density of the 3D VE [Bergé et al., 2014b]. Many researches were proposed in this scope. We cite the two mid-air navigation techniques : the Mid-Air Phone technique and the Mid-Air Hand technique [Bergé et al., 2014c] are based on free physical movement of the mobile device for 3D VE navigation as shown in Figure 43. Tracking a camera-equipped mobile phone provides the navigation subtask of a pointing task. The dynamic peephole method without visual context and the magic lens paradigm using external visual context [Rohs et al., 2007] are two methods for map navigation with mobile devices. These two techniques are based on free physical movement of the mobile device for map navigation as shown in Figure 44. [Çapın et al., 2006] present a camera-based technique to navigate virtual environments. Based on the estimated devices motion and the cameras live input images, a window of the 3D VE is displayed on the devices screen as shown in Figure 45. This technique provides interactive applications : na-

vigation and gesture-based interaction in interactive 3D virtual environments. The user moves the device up, down, left and right to look around the space. Mapping these motions into space direction to display the virtual environment creates the illusion of a small window into an environment that surrounds the user.

FIGURE 43 – the Mid-Air Phone technique and the Mid-Air Hand technique exploring 3D telescope visualization

Other interaction techniques of this category improve the navigation and the selection subtasks of the pointing task to reduce the 3D density of the 3D VE : Jump and Refine interaction technique [Hachet et al., 2007] is used for selection of sparse objects in a 3D environment on mobile phones. This technique is defined by two levels of cursor displacement while using the thumb-devices of the mobile phones. As shown in Figure 46, in the first level (Jump) a grid is displayed on the screen. This level is used for a fast displacement of the cursor from cells to cells into the targeted area (dedicated grid). In the second level (refine), the current cell is only displayed. This level consists of a refinement step, which allows an accurate movement and positioning of the cursor in the selected grid to point the target. This technique is faster than moving a cursor due to the jump movement. It is not efficient in a dense environment.

Another way of implementing the navigation and the selection subtasks is by the Touch Selection Interface interaction technique. Touch Selection Interface [Vemavarapu and Borst, 2013] presents a solution to overcome selection problems obtained from occluded distant or small targets by displaying a part of the total 3D VE on the Smartphone's screen. The position and orientation of the handheld touch device determines a displayed mapped sub region of the virtual environment as shown in
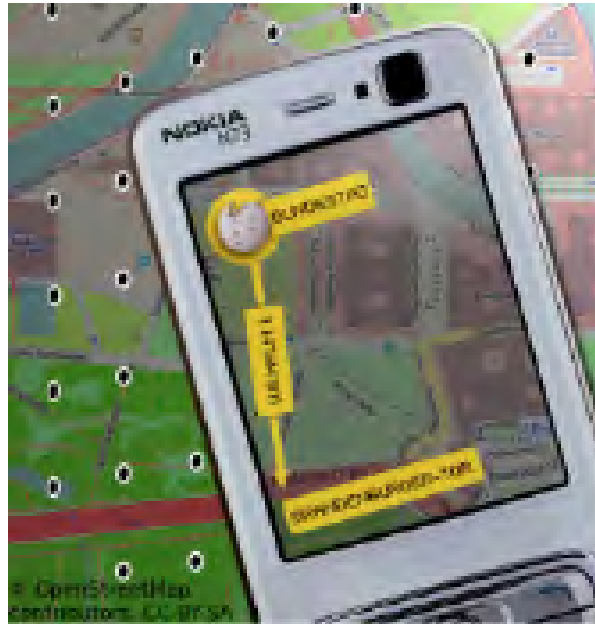
Figure 44 – Application example of the dynamic peephole method without visual context and the magic lens paradigm using external visual context techniques

Figure 47 - right. A 2D overview of a 3D sub region of the virtual environment is displayed on the mobile screen. 3D objects are presented by one size labeled circles. The refinement selection phase begins by touching the screen ; the user can move their thumb position for target selection. The closest object to the user's thumb is highlighted and selected when the user releases the touch as shown in Figure 47 - left. This technique consists of displaying on the mobile device a 2D sub region of the total environment. Touch Selection Interface presents a solution to overcome selection problems obtainable from occluded distant or small targets. A prior navigation phase is required to find the desired object, to reduce the number of displayed object (to reduce the 2D density) and then to point it.

In this case, the 3D VE is partially displayed on the screen then the user does not have a whole view of the scene. Therefore, this type of techniques is based on a preceding navigation phase to find the target in the huge environment and to display it on the Smartphones screen. The pointing task requires multiple navigations, zoom

FIGURE 45 – Camera-based interaction

FIGURE 46 – Jump and Refine interaction technique

in and zoom out subtasks in a repetitive way. This step is a confusing step and may cause a loss of orientation for the user.

Another category of interaction techniques consists of displaying the entire 3D VE on the Smartphones screens. In this case, the objects size is reduced. Therefore, each object has a smaller size.

Rekimoto [Rekimoto, 1996] proposes a technique based on the tilt of the device itself as input. To interact on the environment displayed on the devices screen, the user manipulates both tilt and buttons. This interaction technique provides : selection from a cylindrical or a pie menu, navigation of a large 2D space, viewing a 3D object from different directions as shown in Figure 48.

Decle and Hachet [Decle and Hachet, 2009] study the rotation of a 3D object
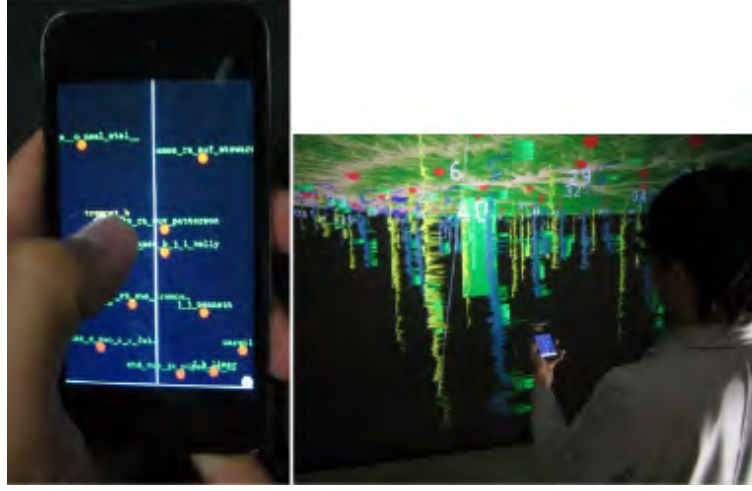
FIGURE 47 – Touch Selection Interface interaction technique

FIGURE 48 – Tilting Operations for Small Screen Interfaces

displayed on a mobile screen via manipulation based on mobile phones touch. To overcome the occlusion provided by the thumb, this interaction uses horizontal and vertical sliding movements : horizontal movements produce rotation around the "up vector" of the camera and vertical movements produce rotation around the "Right vector" of the camera. Therefore, the user is able to move the camera around the 3D model by drawing successive strokes as shown in Figure 49.

Daiber [Daiber et al., 2012] proposes sensor-based interactions for mobile 3D gaming. The navigation is realized through rotation and flipping gestures. The object manipulation is realized through the combination of a direct touch with tilting and rotation of the mobile as shown in Figure 50. The object selection is realized by a
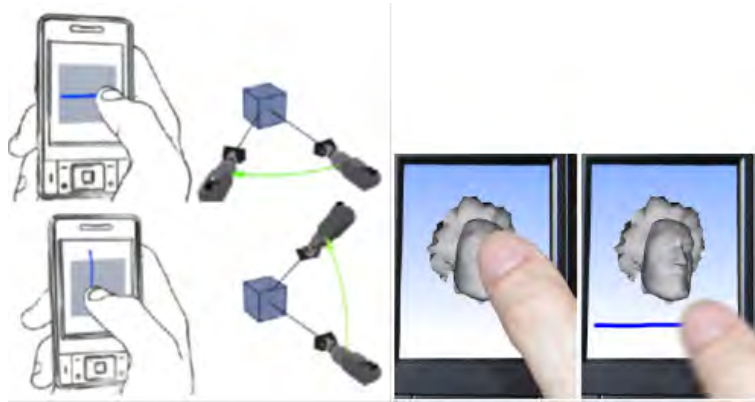
FIGURE 49 – the rotation of a 3D object displayed on a mobile screen via manipulation based on mobile phones touch

simple click and the uppermost object will be pointed. However, these techniques are not useful in case of a dense VE to point partially or fully occluded objects.

FIGURE 50 – sensor-based interactions for mobile 3D gaming

Developing 3D interaction techniques for handheld devices with multi-touch displays is a realistic and natural way of interacting. However, these techniques are facing limitations and problems due to the small size of the device, and the occlusion causing by the user finger. [Telkenaroglu and Capin, 2013] proposes two selection techniques : (1) Dual-Finger Midpoint Ray-Casting and (2) Dual-Finger Offset Ray-Casting from a set of Dual-Finger 3D Interaction Techniques for mobile devices. The Dual-Finger Midpoint Ray-Casting technique consists of using two fingers for interaction. The two fingers positions are used to draw a crosshair marking their midpoint location and a ray is generated from the crosshair toward the scene. Intersected ob-

ject by the ray is highlighted. A zoom technique is performed by moving the user finger while the user has two contact points on the touchscreen. When there is a highlighted object, any third touch action validates the selection process as shown in Figure 51 -up. The Dual-Finger Offset Ray-Casting technique consists of using only one finger as a pointer and an offset. A crosshair is displayed on the scene. Its position is the amount of offset added to the finger position. The crosshair will always follow the finger. A ray is emitted to the scene from the crosshair and intersects the first object as shown in Figure 51 -down. The proposed two selection techniques are designed for handheld AR using two fingers to select small and partly occluded objects in sparse as well as dense AR. They are able to produce fast and accurate results. But, they present some constraints for the totally occluded objects, for the huge number of intersected objects on the very dense environments and for the objects near the screen corners and edges. On the other hand, the low resolution of these screens makes the use of a pointer more difficult since it would be too small and requires too much precision to manipulate.

DrillSample interaction technique [Mossel et al., 2013b] is designed for a 3D selection in a dense AR environment displayed on one-handed mobile device. DrillSample is a two-step selection technique used for pointing visible, partly or fully occluded objects with high similarity in visual appearance. In the first step, the user clicks the target in the scene using single touch as shown in Figure 52 - left. From the 2D touch, a ray is emitted in the 3D scene and the objects casted by the ray are selected. In the second step the AR scene view is turned off and all selected objects are represented as 3D virtual copies for selection refinement. The selected objected are displayed aligned on a horizontal line conserving their depth information as shown in Figure 52 -right. By using : firstly, the built-in Inertial Measurement Unit the user can rotate around the target, secondly, the swipe motion, he can browse the objects or make a zoom in or a zoom out. Finally, by using one-finger touch input, the user selects the target. DrillSample is useful for selecting dense and similar targets. However, DrillSample is not supported for selecting : objects of varying size since small objects may appear too small and big objects might be trimmed at the near image plane, or small objects at a larger distance due to the combination of the use of the ray
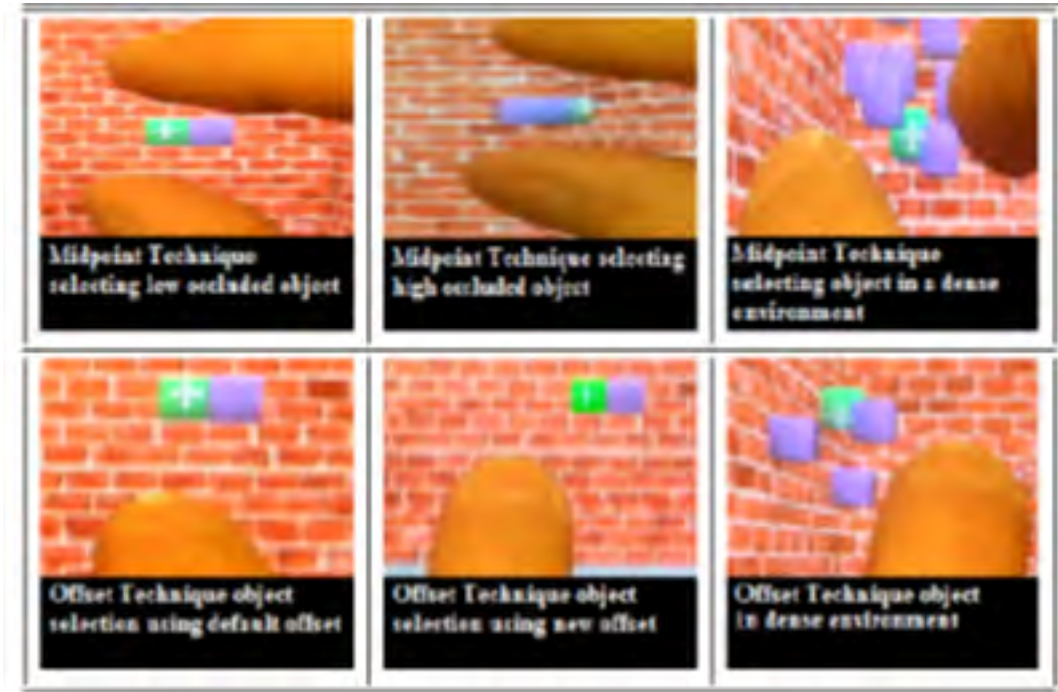
FIGURE 51 – up) Dual-Finger Midpoint Ray-Casting technique and down) Dual-Finger Offset Ray-Casting technique

casting technique and the imprecise finger touch, or objects separated with variable and large distances since the horizontal ray cannot be compressed, or finally, a huge number of objects since the clones overview will look too small. The performance of DrillSample interaction technique can be ameliorated by using a 3D selection volume instead of a ray, a selection algorithm to reduce the number of selected objects or an optimal adjustment of the horizontal line length. DrillSample takes an advantage over the Expand interaction technique by preserving the full original spatial context of the selected objects. But this technique displays the objects on a horizontal view and not on a depth view.

LayerStroke [Wu et al., 2015] is a 3D interaction technique using a 2D touch-based selection technique, designed to select small targets located in a dense 3D VE on a smartphone. It is a dual step technique. The first step consists of executing an algorithm to divide the targets into different layers. Users can select a layer by
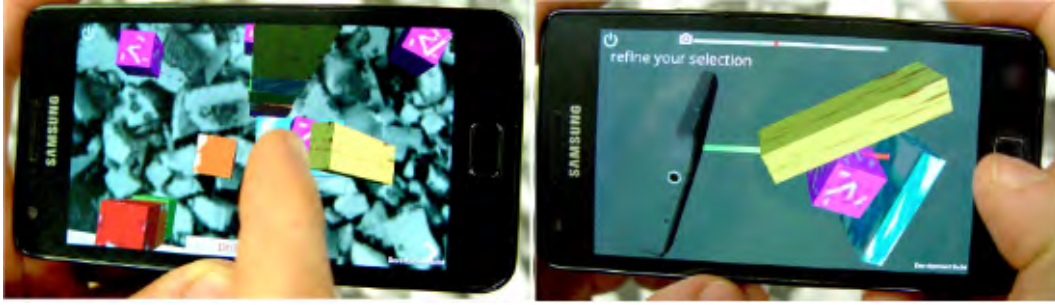
FIGURE 52 – DrillSample interaction technique : left) the user clicks the target in the scene using single touch, right) selected objected are displayed aligned on a horizontal line conserving their depth information

drawing a stroke on the screen. After that, to select a target inside the selected layer, the user taps the corresponding tiles as shown in Figure 53. The advantage of this technique resides on selecting accurately small targets without controlling very precisely the finger movement due to the large tessellation tiles and without magnifying the initial view. LayerStroke reduces the occlusion problem residing by pointing with a fat finger but cannot solve the total occlusion when targets overlay each other.

In conclusion, the majority of interaction techniques proposed in literature are designed to solve the occlusion problem of the fat finger while pointing on a 2D environment displayed on mobiles devices. There is a lake of interaction techniques used to point a target on a 3D VE displayed on Smartphones taking into consideration the depth, the occlusion and the density problems.

4 Smartphone as 3D input device

Some interaction techniques consist of using Smartphones as pointing devices for an Augmented Reality environment or for 3D VE displayed on large screens. In this case, the virtual environment is not displayed on the device screen and the mobile device is used as a 3D input device.

Based on the powerful Smartphones sensors and capabilities (e.g. camera, accele-
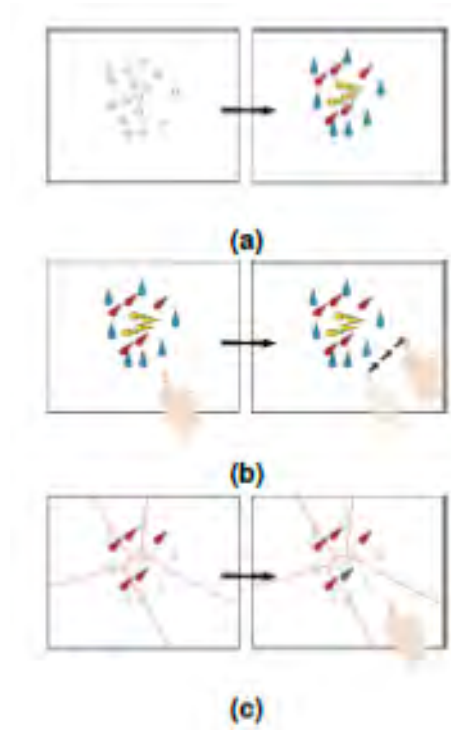
FIGURE 53 – LayerStroke interaction technique : a) divide targets into several layers, b) draw a stroke to select a layer, c) tap the finger in the tile of a target to select it

rometer, gyrometer, GPS, etc.), a set of mobile 3D interaction techniques is proposed. These techniques rely on sensor based input allow the navigation and the manipulation of objects in the scene by tracking the way the user is holding, moving and touching the device. Boring et al. [Boring et al., 2009] introduced three interaction concepts : scroll, tilt and move-gestures to remotely control a pointer on a large public display by using the mobile phone as an input device. Figure 54 illustrates the pointer displacement in function of the mobile device movement (the arrow in Figure 54 shows the pointer displacement). On the other hand, [Benzina et al., 2011] present travel-techniques for VRs and use the mobile phone as a 3D spatial input device.

Smartphones are used as pointing devices for an Augmented Reality environment or for 3D VE displayed on large screens. They are used as multi-Degree-Of-Freedom

FIGURE 54 – scroll, tilt and move techniques

(multi-DOF) controllers, as interaction devices, and gesture recognition for mobile games.

Bauer at al. evaluated the use of a smartphone as an interaction device for large displays [Bauer et al., 2011] they compared its use to the traditional input devices and showed its performance over the keyboard and mouse.

Graf et al. studied the position and motion sensors provided by these smartphones for 3D navigation. Consequently, they proposed the use of a smartphone as an interaction input device with 3D visualizations and presentations [Graf and Jung, 2012]. HandyMap technique operates on an iPod touch interface allowing a ray-casting pointing followed by a refined selection of the secondary 2D overview of the scene using the touchscreen [Prachyabrued et al., 2011] as shown in Figure 55.

Finger Walking in Place (FWIP) interaction technique presents a navigation technique in a CAVE [Kim et al., 2008]. This technique allows a user to travel in a virtual world by considering the finger motion on a multi-touch device as the human leg walking motion in a virtual world ; Figure 56 shows the user moving his hand finger on

87

FIGURE 55 – The HandyMap interaction Technique. Left : the user points the target using a ray in a 3D VE, right : the subset of pointed objects is displayed on the iPod screen ; the refinement phase is applied via the touchscreen based interactions

the touchscreen.

Katzakis et al. explore the use of smartphone sensors for a 3D rotation task [Katzakis and Hori, 2010]. The results show that mobile device is significantly faster than both the mouse and the touch pen.

Rodriguez A. and Leon A. have proposed the use of smartphones as a generic remote 3D interaction device to remotely manipulate 3D content of digital heritage applications. By connecting the device to a remote application via a WiFi connection, the current device orientation is tracked through sensor-fusion, and combined with a single-handed touchscreen input to perform uniform scaling operations (by using the swipe event on the top touchscreen area), 3D translating operations (by uniting the central touch area input to the current device orientation) and 3D rotation operations (by performing touch events on the bottom touch area) as shown in Figure 57 [Rodríguez and León, 2015].

P2Roll and P2Slide [Delamare et al., 2013] present two interaction techniques using a cycling mechanism. P2Roll (Physical Pointing Roll – P2Roll) uses a wrist rolling gesture and P2Slide (Physical Pointing Slide – P2Slide) uses a finger sliding gesture to point, touch, roll left or right or slide left or right a physical target as described in Figure 58.

Researchers have also designed new interaction techniques that transform the Smartphone to be a 3D controller for pointing and manipulating avatars of 3D vir-

FIGURE 56 – Finger Walking in Place (FWIP) interaction technique

tual reality environment displayed on large screens. Tilt & Touch [Du et al., 2011] is an interaction technique that converts a mobile phone into a six degrees-of-freedom controller (including movement (forward/backward, up/down, left/right) and rotation (pitch, yaw, and roll)) via the capacitive touchscreen and built-in MEMS motion sensors as shown in Figure 59. Tilt & Touch provides four basic 3D manipulations : first, navigation and viewpoint transformation by tilting the phone, second, 3D space displacement by tilting and touching, third, zooming by touching the screen and fourth, Object rotation by tilting the phone.

Other interaction techniques were proposed to interact with the physical world while carrying a handheld device within smart environments. These techniques provide solutions to the problem of mobile distal selection of physical objects in augmented environments. This type of techniques provides 3D interactions : scaling, pointing, rotation, zooming, displacement, navigation, translation, scrolling, tilting and moving. Therefore, this type of interaction technique is out of scope of the current study since the virtual environment is not displayed on the Smartphone screen.

Figure 57 – Using the smartphone as a generic remote 3D interaction device
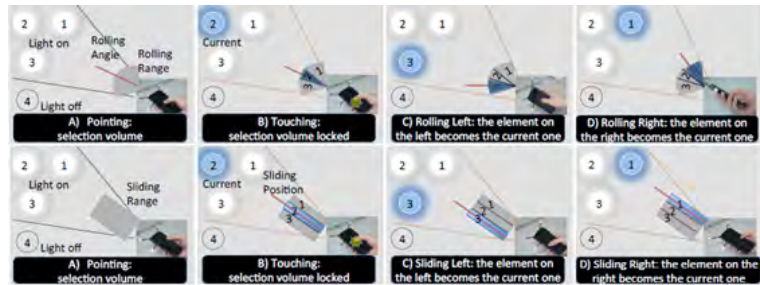
Figure 58 – Up : P2Roll interaction technique. Down : P2Slide interaction technique
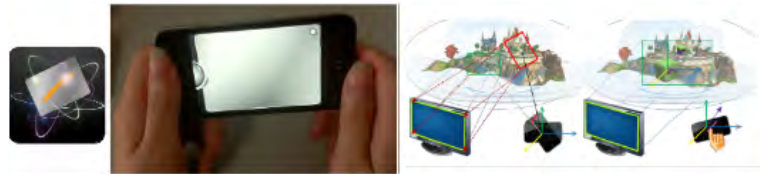
Figure 59 – Tilt & Touch interaction technique

Chapitre II

Classification of interaction techniques in the 3D virtual environment on mobile devices

Once the selection interaction techniques have been described, in this chapter, we will begin by presenting the usage constraints of an interaction technique in 3D virtual environments and the resultant problems affecting the selection performance.

In light of the selection problems described below, we will propose a new classification space [Balaa et al., 2014]. This classification is intended to assist in comparison and design of interaction techniques in 3D on mobile devices. The originality of our approach is to particularly emphasize usage constraints and problems (density, depth, and occlusion) in the context of the selection task. It is no longer a technologically centered approach but a real user centered approach.

1 Usage constraints

3D virtual environments are wide and dense with large number of targets of various sizes. Conversely, the smartphone has a small and flat 2D screen. In general, the user uses his finger to select a target using the touchscreen. But, fat finger

presents a limitation while targeting a small object. The use of 3D applications on mobile devices presents specific constraints due to : first, the insufficient degrees of freedom while manipulating in 3D environment with 2D interaction techniques, second, the finger occlusion when the user's finger covers the screen area [Wobbrock et al., 2008], third, the difficulty of executing an accurate, pixel-level selection in dense environment with varying object sizes displayed on a small screen [Telkenaroglu and Capin, 2013], and fourth, the limitation and the small size of the display and interaction area [Telkenaroglu and Capin, 2013].

Following the performance criteria as described before (in chapter 1 paragraph I "Performance metrics"), we have identified three types of constraints : Decrease accuracy, Increase pointing time and Increase reaction time. Decrease accuracy occurs when pointing is difficult and confusing. Since, it is difficult to avoid the environment object to make the right selection of the target. Increase pointing time happens when the user performs multiple actions to accomplish the pointing task. Accordingly, the total time required to point correctly the target will increase. Increase reaction time arises when the user executes a rearrangement phase to select the right target [Hick, 1952, Hyman, 1953]. Since, while performing a rearrangement phase, user needs an additional time after each action to find the target new position. Consequently, the reaction time increases and leads of the augmentation of the total pointing time.

These constraints create some problems or factors affecting selection performance : The environment density [Vanacken et al., 2007], the depth of targets [Hwang and Jeon, 2009] and the occlusion [Vanacken et al., 2009].

1.1 The environment density

Assume that the virtual environment is wide with a large number of targets of different sizes. To display a large number of targets with various sizes on the 2D screen of mobile devices, the scene must be shrunk, or the virtual environment and its components must be minimized. Then, the scale is reduced ; the objects size and space are minimized. Therefore the distance between objects is reduced. Consequently, to select an object, there may be a problem due to the presence of several objects around

the desired one, and the different objects are very close to each other [Grossman and Balakrishnan, 2006]. In this case, how to quickly and accurately select the desired object ?

The use of direct interaction to point 3D objects presents a lack of precision caused by the fingers on the screen. Indeed, when the user will touch the screen, the contact surface of the finger on the screen is large and can cover multiple objects (cf. Figure 60). Thus, it is difficult to determine which point will be taken into account. Depending on this point, one object could be selected rather than the other. As a result, the denser the 3D EV is, the less precise the selection will be.
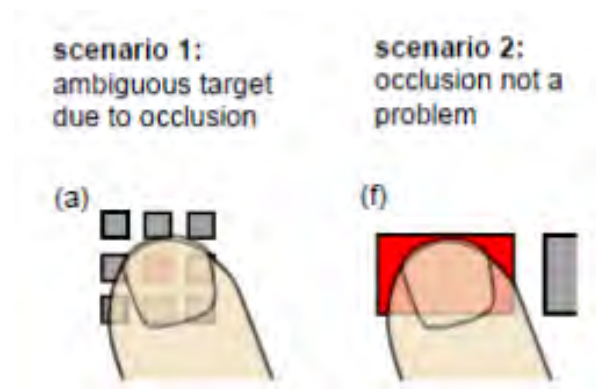
FIGURE 60 – Occlusion problem due to the finger's contact area [Vogel and Baudisch, 2007]

On the other hand, the low resolution of these screens also makes the use of a pointer more difficult to manipulate because it would be too small and require too much precision. The less of precision in the displacement and the selection in a crowded and dense environment create the selection of N selected items at the same time instead of one item. As a result, the target will have to be selected from a set of objects [Bowman, 2002]. Consequently the selection task problem consists of choosing the desired object accurately from the multiple selected objects.

1.1.1 Navigation subtask

Navigation subtask solves the environment density problem by changing scales and views of the environment due to zoom, pan or rotate actions of the exploratory movement.

1.1.2 Pointing subtask

In case of dense and crowded environment, moving a 3D point cursor on the 3D VE does not solve the problem. Using a 3D Point Cursor technique presents some limitations in terms of the 3D environment density problem. The density problem occurs due to the fact that the target is surrounded by multiple objects that make the selection of this target a complicated task. And in case of small target, we have to make many displacement of the position of the input device to catch the target. This fact augments the response time.

The density problem isn't solved by using a 3D volume cursor. Due to the fact that the target is surrounded by multiple objects, the number of targets inside the volume cursor will be more than one. Consequently, the selection of this target is a complex task because we have to choose the target between the N selected objects or we have to use a dynamic selection area. The Silk cursor technique does not solve the density problem. Indeed, only one object must be in the volume so that the user can select it. If several objects are inside, the selection is not possible and the technique is not accurate. The density problem is not solved either by the ray cursor technique, the Go-Go and HOMER techniques, which require a higher selection precision. In fact, to point the target precisely, the user has to make fine displacement of the input device, and it requires more accurate movements and multiple adjustments of the ray direction in depth.

To overcome the problem of selecting a group of N aligned objects in depth, researchers define multiple techniques based on a variable length ray emission or a ray with a depth marker to identify and point the correct target. The difficulty resides on making multiple adjustments to make the ray crosses accurately the desired object. The density problem is not solved while using a Ray with variable length

technique either the depth marker technique. This type of techniques is efficient if the ray crosses exactly and quickly the target. However, using a lock by The Lock ray technique provides a separation between the ray emission and the cursor adjustment. This technique uses a continuous displacement of the input device without making modifications of the ray position. Consequently, the lock eliminates the hand shake problem and partially solves the density problem.

To overcome the density problem, some techniques use a transparent dynamic selection volume instead of a ray or a point cursor to augment the selection area. Considering a volume selection area augments the chances to find the target inside from the first attempt and does not require a higher selection precision to point the target. A selection volume is defined by a ray and the volume around it. This volume can have different forms : a cylinder in the case of the transparent cylinder technique [Dang et al., 2003] or a cone in the case of the Spotlight technique [Liang and Green, 1993]. Or, the selection volume is defined by a point cursor and the volume around while using the transparent sphere technique [Dang et al., 2003]. These techniques select the target and a set of objects around it. The increase of the selection area minimizes the pointing problem in a dense environment but requests a second refinement selection step to point the target from the preselected set of objects. These techniques are accurate ; don't require accurate movements or multiple adjustments in depth of the input device position to catch the target. They don't present a hand shake problem and don't depend on the target size. However, the limitation of these techniques is in the huge number of objects existing in the subset. Consequently, they request a large response time because we have to : choose the target from a long list while using the menu technique, wait for the target to be highlighted while using the circulation technique, wait for the generation of the selection criteria while using the Virtual pointer metaphor [Steinicke et al., 2006], wait for the calculation of the refinement algorithm while using the Spotlight interaction technique [Liang and Green, 1993] or add a navigation phase while using the shadow cone technique [Steed and Parker, 2004].

Using the transparent sphere technique, the density problem is partially solved because it is possible to select a target from a set of items. The transparent sphere

technique suggests the use of one of the two proposed techniques : the menu technique or the circulation technique. The menu technique solves the density problem but lose the information about the depth position. The circulation technique solves the density problem but requests a large response time. The transparent cylinder interaction technique presents a solution to the 3D environment density problem since the target is selected using a circulation technique or Floating menu technique. This technique minimizes the density problem but will produce an immense response time resulting from the selection of a large number of objects inside the cylinder. Considering a dynamic selection volume presents a solution to the 3D environment density problem because it augments the selection area, this technique minimizes errors but will lead to a delay and an immense response time. The response time can be reduced by reducing the number of objects selected by adjusting the dynamic selection area. Once found, the object should be centered in the middle of the dynamic selection volume and the area around the object is gradually shrunk to minimize the total number of the selected objects. Using the Spotlight technique, the huge number of objects inside the cone is reduced while using a dynamic selection volume. The dynamic selection area is used to select a target very close to the ray emitted but not hit by the ray. Consequently, Spotlight technique does not require a higher selection precision to point the target ; it is accurate and presents solutions to the problems resulting in case of small target, hand shake and the big number of selected objects in a dense environment. Solving the density problem cannot be achieved by a ray or a point cursor based techniques in a non-transparent environment. Using a dynamic area selection in front of a ray or a point cursor solves the 3D environment density problem and is better from the menu because its response time is less. In case of using the menu technique, the user have to search for the active object in a large list because the list contains all the object in the environment, the response time is immense and the performance decreases. Even though using the dynamic area selection case is better for the response time problem, the use of a menu is an error prevention method, accurate with no ambiguity and natural.

Using a curved ray has the advantage of selecting a deep target without passing through the environment objects. This technique does not depend on the target

size and presents solutions for the hand shake problem. However, these techniques face the same problem as the other rays in the case of a dense 3D VE with a high density of objects since ; it becomes difficult to get around all objects to point a particular target. The curve with fixed length solves partially the 3D environment density problem. The user has to move many time the wand position in order to select a target. This interaction technique is restricted by the problems resulting in case of small target and the response time because we have to make many displacement of the position of the input device to catch the target. On the other hand the Flexible pointer technique allows an adjustment of the curve length and shape to point the target efficiently. Solving the density problem is achieved by multiple adjustments of the form and the length of a flexible curve in a transparent environment.

Other technique solves the environment density by modifying the amplitude of the transfer function which transforms the coordinates of the finger position into a crosshair coordinates in the scene as described in the Dual-Finger Midpoint Ray-Casting and the Dual-Finger Offset Ray-Casting from a set of Dual-Finger 3D Interaction Techniques for mobile devices proposed by researchers [Telkenaroglu and Capin, 2013]. The proposed two selection techniques present some constraints for the huge number of intersected objects on the very dense environments and for the objects near the screen corners and edges. On the other hand, the low resolution of these screens makes the use of a pointer more difficult since it would be too small and requires too much precision to manipulate. Consequently, it seems more appropriate to use a volume selection area to select multiple objects and then point the target among these pre-selected items. However, the proposed selection techniques are not suitable for use on smartphone for example using a menu to select from several objects could hide the entire 3D VE. By using a small screen, the best solution appears to maintain a visualization of the different pre-selected objects, then to refine the selection from it.

The environment density is also solved by further proposed techniques which consist of designing a virtual modification of the distance between targets or the target size : the Semantic 3D pointing [Elmqvist and Fekete, 2008] consists of shrinking empty space and expanding target sizes on the screen by dynamically adjusting the

ratio between the cursor movement on the screen and the movement of the input device. In a dense area, the distance measured from the cursor to the closest target is small thus the cursor movement is slowed down to allow for higher accuracy. However, in an empty space, the distance is big thus the cursor movement is speeded up.

1.2   The depth of targets

The environment elements are displayed in 3D. Accordingly, to select a target, the user must also moves in depth. As a result, the user faces two problems caused by depth : the first concerns the perception of depth : the user must understand where the pointer is relative to the objects in the 3D VE ; the second concerns the pointing of the target : How to quickly and accurately point to the desired object ? [Grossman and Balakrishnan, 2006].

Depth adds additional complexity to the selection task. The pointing task encounters problems due to : firstly, in the case of indirect interaction, how to determine the position of the cursor in depth to reach an object that may be farther or in front of the pointer ; secondly, how to move in depth to reach the intended target taking into account the obstacles (other objects placed in front) in the scene, and thirdly, how to select the desired object accurately and quickly.

On the other hand, common interaction techniques on mobile devices are designed for use in 2 dimensions and do not offer a degree of freedom sufficient to interact properly with the third dimension (depth) of the 3D VE. To facilitate interaction with the 3D VE, specific input devices are designed to provide more degrees of freedom [Froehlich et al., 2006, Perelman et al., 2015]. However, these devices are designed for PC use or volumetric display and cannot be transposed directly for use on a smartphone. Similarly, immersive applications use devices that are not suitable for smartphones. The interactions are limited to the capabilities offered by the devices of the smartphone such as the touchscreen, the physical button and the camera and the sensors such as the accelerometer, the gyroscope.

1.2.1 Navigation subtask

Navigation subtask partially solves the depth problem by helping the user to realize the perception of depth and to determine the position of the target in the 3D VE. The general movement which is an exploratory movement gives information and knowledge about the environment and the target position accessible by its principal possible actions : the rotation, the pan, and the zoom [Phillips and Badler, 1988].

Navigation subtask offered solutions to another constraint by offering the possibility to move in depth. Moving in depth using navigation can be achieved by : a Targeted movement for example (Go-To/Fly-To) which consists of jumping the camera to the goal position that the user had specified [Mackinlay et al., 1990], or a Specified trajectory movement for example StyleCam experience which consists of creating constrained path for navigation [Burtnyk et al., 2002] or a Specified coordinate movement for example CAD and 3D editing software which consists of precising the exact position and orientation of the target. It's a guided navigation that limits the user's freedom while travelling through the 3D VE.

The target is now visible ; Navigation subtask partially solves the depth problem since it does not present a way to point the target precisely.

1.2.2 Pointing subtask

Pointing subtask proposes different techniques to move in depth then solves the depth problem. Pointing subtask helps the user to realize the perception of depth and to determine the position of the target in the 3D VE, to move in depth and then, to point the target.

Firstly, a cursor controlled by a pointing device is moved in the 3D VE. The cursor could be the cross-hair of the 3D Point cursor technique [Grossman and Balakrishnan, 2006] circulating in the environment. The point cursor metaphor used a cursor to circulate in depth in the 3D environment by changing and setting the value of the z coordinate. This characteristic is important to solve the depth target problem. However, this task is complicated since it requires more accurate movements and multiple adjustments in depth. The solution proposed to resolve this problem

consists in using a volume selection area rather than a small cursor. In the Silk cursor technique [Zhai et al., 1994], the point cursor is replaced by a rectangular volume. Accordingly, this gives information on the position in depth of the object in relation to the volume cursor. Consequently, the depth position of the cursor is not well defined and it is replaced by a range of values therefore, we cannot precisely define the target position. The semi transparency of the cursor facilitates the perception of it in space, but that does not solve the problem of pointing in depth : Consequently, the depth problem is partially solved by using the Silk cursor technique. The transparent sphere technique [Dang et al., 2003] proposes to replace the crosshair by a transparent sphere. Consequently, the number of the objects inside the sphere is N and the depth position of the cursor is not well defined and it is replaced by a sphere volume then we cannot precisely define the target position. This interaction technique uses list menu or circulation techniques to point the target between the N objects inside the sphere volume. Consequently the depth problem is partially solved by using the transparent sphere technique. To determine the cursor position in the 3D VE, this interaction technique uses a linear transformation between the cursor of the 3D VE and the input device. This property presents a depth problem if the target depth position is greater than the hand length. The solution should be changing the characteristics of the transfer function. In summary, pointing techniques, based on the point cursor, do not support the navigation phase. They require an additional accuracy effort from the user to point a target in depth. On other hand, pointing techniques, based on 3D volume cursor partially solves the depth problems, since ; it solves the problem of perception but not the problem of pointing.

Secondly, by emitting a ray in the 3D VE toward the desired object to point it [Liang and Green, 1994], it's the Ray cursor technique [Bowman and Hodges, 1999, Grossman and Balakrishnan, 2006]. The techniques based on a ray emission point the target more easily in depth by projecting a ray to infinity in the 3D VE. The Ray cursor technique solves the depth problem thanks to the virtual ray projected to infinity. It is thus faster to point an object that is far away in the depth. Based on ray emission, different techniques like Go-Go and HOMER solve the depth problem. These techniques are used in a sparse environment to point a visible target. But Go-

Go is limited by the real hand length and the difficulty of grabbing. Then the depth problem is partially solved by the Go-Go interaction technique. The techniques based on a ray emission face the problem of a target placed after N aligned objects for the reason that the ray should pass across the N objects. Ray with variable length or depth cursor techniques are proposed as a possible solution. The Ray with variable length technique [Dang et al., 2003] identifies and points the correct target from the group of the aligned objects. While, The Depth ray [Grossman and Balakrishnan, 2006] interaction technique combines a ray cursor technique and a depth cursor, the depth marker travels between the objects following the input device displacements to reach the target. Although, the Lock Ray technique combines a depth ray and a lock. Using an infinite ray has the advantage of selecting a deep target. Thus, the user can detect the position of the target and make the adjustment to reach it. This task solves the depth target problem, but it is complicated since it requires more accurate movements and multiple adjustments in depth.

Thirdly, using a curved ray has the advantage of selecting a deep target without passing through the environment objects. Thus, the user can detect the position of the target and make the adjustment to reach it. This characteristic solves the depth target problem. But this task is complicated since it requires more accurate movements and multiple adjustments in depth. Curve with fixed length and Curve with variable length techniques partially solves the depth problem because the curve help avoid targets which is in front of the desired target. To avoid the problem of crossing a large number of objects by a ray, Curved ray techniques present a possible solution that allows of bending the ray launched by the user and thus bypass the objects that would face the desired one. The flexible pointer technique [Feiner, 2003] solves the depth problem because this technique enables multiples adjustments and curvatures. Thus, the final curve does not have necessary a well-known shape. Using an adjustable and non-fixed curve shape helps avoid targets which are in front of the desired target and to circulate in depth.

In summary, Pointing subtask solves the depth constraints by helping the user to realize the perception of depth and to determine the position of the target in the 3D VE when the user controls the circulation of a 3D pointer, the emission of a ray

or the multiples adjustments and curvatures of a curve or by helping the user to move and circulate in the depth via the 3D cursor, the ray or the curve or by helping the user to point the target once the cursor hit the target or the ray and the curve intersected it.

1.3   The occlusion

The occlusion problem occurs when one or more elements of the scene can mask the target partially or totally and thus reduce its visibility and ability to be selected [Feiner, 2003, Rosa and Nagel, 2010]. In addition, in the case of a direct interaction, the user uses his finger to point a target directly on the screen. This interaction can create an additional occlusion. The problem of occlusion is how to reach an object if it is hidden by another or by the interaction process ?

In case of total occlusion, the target is totally invisible and hidden by 3D VE elements ; therefore the user cannot find the object to point. The problem of occlusion is to look for the object before you can point it. In the case of a small screen, the higher density of the environment increases the risk of occlusion because the objects are closer to each other.

In the case of partial occlusion, the target is partially hidden by the 3D VE elements, so precise selection is difficult because of the superposition of the objects as shown in Figure 61 (The red target is partially hidden by the 3D VE elements). Object selection is more difficult because the visible surface of the object to be achieved is reduced. Accordingly, several objects close to the target may be difficult to avoid when trying to select the target. Consequently, the target will have to be selected with a set of objects [Bowman, 2002]. As a result, the selection task will be divided into two parts : selecting a set of objects containing the target, and then choosing precisely the desired object from the multiple selected objects.

1.3.1   Navigation subtask

The exploratory movement of the Navigation subtask solves the occlusion constraints when the target is hidden by others. The solution is offered by three possible tech-
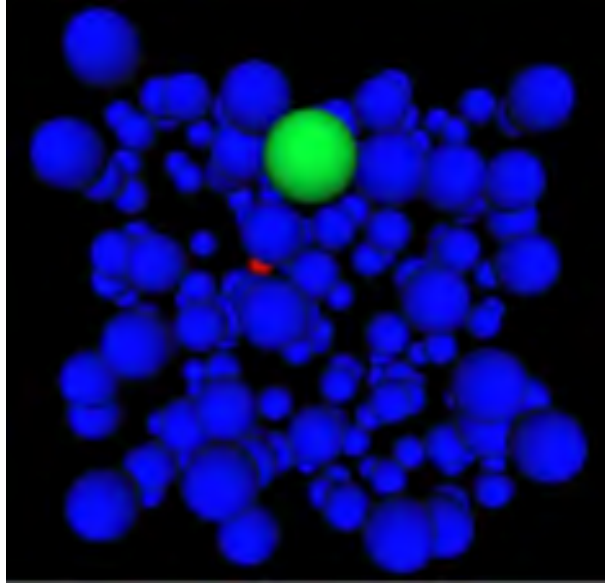
FIGURE 61 – The red target is partially hidden by the 3D VE elements [Rosa and Nagel, 2010]

niques : the rotation, the pan, and the zoom [Phillips and Badler, 1988]. Rotate is defined by the orbiting of the camera around a central point in any direction ; Pan is defined by the translation of the camera along x and y axes ; and Zoom is defined by the translation of the camera along its line of sight.

1.3.2   Pointing subtask

In case of dense and crowded environment, the target can be occluded. Using a 3D point cursor technique faces the occlusion problem : due to the fact that the target is surrounded by multiple objects that make the selection of this target a complicated task. Otherwise, the ray cursor technique is confronted with the problem of occlusion : if the target to be reached is hidden (even partially), the ray can intersect the object in front of the target. In this case, the ray selects the first target encountered and not the target to be reached. Therefore, this technique does not solve the occlusion problem. Moving the virtual hand on the ray proposes a possible solution to the occlusion problem. Indeed, with Go-Go, the virtual ray does not

103

automatically select the first object crossed. Consequently, the user can select, with the virtual hand, another object which is crossed by the ray. Consequently, the Go-Go and HOMER techniques solve partially the occlusion problem. Changing the ray length or circulating a depth cursor profit to identify and point the correct target from the group of the N aligned objects. Since circulating the depth marker profit to identify and point the correct target from the aligned objects. Consequently, The Ray with variable length technique and the Depth ray technique solve partially the occlusion problem. Accordingly, Lock Ray technique solves the occlusion problem. This technique is the best choice in the case of visible, partially occluded targets or totally occluded with transparency [Sellen et al., 1992].

Using a transparent environment solves the occlusion problem and simplifies the pointing subtask. The execution of the point cursor, the ray cursor and the curve techniques became more accurate. Then, the curves techniques allow user to avoid targets located in front of the desired one even if the curves must perform several movements to point precisely the target in case of the Flexible pointer technique. The curve techniques are generally used in case of a visible target or expected target position.

Pointing subtask proposes the adding of a second selection step to the infinite ray propagating in depth or to the volume cursor circulating in depth :

A first example consists of adding a menu to a ray cursor while using the Floating menu technique [Dang et al., 2003] or a menu to the sphere cursor while using the transparent sphere technique [Dang et al., 2003]. This type of techniques solves the occlusion problems by displaying all the visible, partially occluded and totally occluded selected items in an accessible menu. In the menu technique, a floating menu is displayed, accordingly this list menu contains all the selected objects, and consequently the user can choose the target from the menu by clicking the input device button. This technique is functional in case of visible, partially occluded or totally occluded targets with transparency. However, the limitation of these combined techniques is in the huge number of objects displayed in the menu.

The menu is used to identify and point the correct target from the group of the N aligned objects. The menu can have different forms : a 2D flower menu in the

Flower ray technique [Grossman and Balakrishnan, 2006], a virtual sphere menu in the Daisy menu technique [Liang and Green, 1993], a circular menu in the Ring menu technique [Liang and Green, 1993]. Consequently, Flower ray technique [Kurtenbach and Buxton, 1993] offers a solution to the occlusion problem since it transforms the 3D problem into a 2D menu problem. Searching for the target in the 2D menu minimize errors. Daisy menu technique transforms the 3D problem into a 3D menu problem. The technique displays the objects in a simple and clear form. Consequently, searching for the target in the 3D menu reduces the occlusion problem. The Ring menu technique solves the occlusion problems since it transforms the 3D problem into a 3D menu problem. Ring menu presents a solution to the delay by moving and rotating the ring. The Floating menu interaction technique transforms the 3D problem into a 1D menu problem, searching for the target in the 1D menu consist of searching in a long list. The menu is used to identify and point the correct target from a set of preselected objects. For example, the transparent cylinder technique uses a ray cursor and adds a menu and transparency to solve the occlusion problem. The number of objects selected by the ray can be reduced by adjusting the cylinder volume and the number of objects inside.

Using menus solves the occlusion problem. The limitation of these combined techniques is in the huge number of objects displayed in the menu. Searching for the target in a menu will lead to a delay and an immense response time. The response time can be reduced by reducing the number of objects selected by the ray. Another constraint is due to the loss of information about the objects depth locations and positions. Finally, displaying the menu may occlude the 3D VE.

A Second example consists of adding a numeric selection algorithm while using the Spotlight technique or the Virtual pointer technique. This type of techniques solves the occlusion problems by testing via an algorithm all the visible, partially occluded and totally occluded selected items. The number of objects tested can be reduced by reducing the number of objects selected by the ray. Adjusting the dynamic activation area reduces the number of objects inside and the response time. Consequently, the techniques performance augments.

A Third example consists of adding a Navigation phase while using two rays by

the Two ray selection technique [Dang et al., 2003] or two cones by the Shadow cone interaction technique. The navigation phase changes the viewpoint to make the target visible and clear and consequently to reduce the occlusion. In case of multiple targets selected, the Two ray selection technique proposes to emit a second ray from a different view-point to choose the target between the intersected objects. However, an algorithm is needed to memorize the objects crossed by the first ray. The Shadow cone interaction technique suggests to navigate the environment, and then to send a second cone. The targets weight will be adjusted and help to choose between the intersected targets.

A Fourth example consists of adding volumetric meshes. The Starfish technique [Wonner et al., 2012] is a precise interaction technique to select targets in high-density 3D virtual reality environments. At any given moment and according to the position of the head, a starfish-shaped volume is dynamically rebuilt to capture targets near the head. Accordingly, Starfish provides a solution for occluded targets.

2   Summary

We propose a new classification of the existing interaction techniques, according to the three discussed problems (density, depth and occlusion) affecting the selection performance in a dense 3D VE. Our classification space is described as follow (cf. Table 1). Each selection technique is represented in a line. Columns are regrouped into 3 categories : Technique, Selection Task and Constraint.

The column "Technique" describes each interaction technique and specifies their input and output devices. This column is divided into three sub columns : "Selection technique", "Device used" and "Display". The column "Selection technique" indicates the name of the selection technique and its reference. The column "Device used" states the interaction device used by the user while performing the selection task. We grouped the input devices into different categories : the first category consists of the traditional input devices for example : the mouse, the wireless mouse and the joystick, the second category regroups the sensors based devices like the motion tracker, the

optical tracker, the position and orientation tracker, the head magnetic tracker, the tracked wand, the tracked gloves and the tracked joystick, the third category is the camera based tracking devices category and finally, the category of the Smartphone accessories for example the directional input buttons, the multi-touch screen, and the touch-screen. The "Display" column refers to the output display where the 3D VE is exposed. Display can be : 3D volumetric display, 3D stereoscopic environment, spatially immersive display (SID), 3D VR headset, Head Mounted Display (HMD), screens, Smartphones screens and multi-touch screens.

The column "Selection Task" describes the selection technique category, the selection technique sub-task and its way of eliminating disambiguation. This column is divided into three sub columns : "Navigation", "Pointing" and "Disambiguation". The column "Navigation" indicates by "Yes" or "No", if the selection technique integrates the navigation sub-task with the pointing sub-task to perform the overall selection action. The column "Pointing" specifies the selection technique category. According to a well-established classification [Dang et al., 2003] selection technique used in 3D Virtual Environment are classified into 3 categories : point cursor, ray casting and curve. Consequently, for each cell of the "Pointing" column is written : point cursor, ray cursor or curve. In addition to the classification proposed by Dang [Dang et al., 2003] if the selection technique uses volume cursors and selects a volume instead of an object, the selection technique category is designated by volume point cursor instead of point cursor and volume ray cursor instead of ray cursor. Considering the environment constraints and the pointing technique category (while using a ray or a volume cursor) ; the number of targets selected by the pointing technique will takes different values and can be one target or a set of N objects. If the interaction technique selects multiple objects, therefore the selection must present a disambiguation. Then, the column "Disambiguation" describes the technique used to eliminate the ambiguity to accomplish the user selection. Therefore, the cells might have different values : arm extension technique, variable length technique, curved ray technique, moving further pointer, using a circulation technique, adding navigation phase, using a dynamic selection volume, using a menu, executing a selection algorithm or object rearrangement on the screen.

The column "Constraint" describes the ability of each interaction technique to solve the pointing problems. As described before (in paragraph A "Usage constraints"), the user must point a target that is close to several others (density problem); therefore the technique must allow the selection with accuracy of the target among the N surrounding. Moreover, the user must be able to point with the pointing technique the target that would be distant in depth (depth problem); therefore the technique must enable to quickly and accurately travel in depth and select this target. Finally, the target position is known but the target is partially or fully occluded (occlusion problem); therefore the technique must allow the selection of the desired target despite obstacles. Consequently, "Constraint" column is divided in sub columns with each encountered problem : "Density", "Depth" and "Occlusion". While using the interaction technique, if the constraints yield to problems consequently it is noticeable in the corresponding cell of the three columns. Following the performance criteria as described before (in chapter 1 paragraph I "Performance metrics"), we have identified three types of constraints. Accordingly, the cells values are "Decrease accuracy", "Increase pointing time" or "Increase reaction time".

Our classification consists in representing each technique in the table. Thus, this table allows us to quickly see the advantages and weaknesses of a pointing technique, and also allows comparing two techniques together. This representation also allows us to study more precisely the resolution of a problem by analyzing the various solutions proposed for this particular column. We will use this classification space to compare existing interaction techniques. We highlight the differences of the various techniques to demonstrate that existing classifications do not sufficiently highlight the weaknesses and forces of each technique. The following table (cf. Table 1) summarizes the classification of the techniques we have presented in the previous chapter.

TABLE II.1 –

Selection technique	Device used	Display	Nav.	Pointing	Disamb.	Density	Depth	Occlusion
3D Point cursor [Grossman and Balakrishnan, 2006]	Motion tracker (6DOF tracker)	3D volumetric display	No	Point cursor	-	Decrease accuracy	Increase pointing time	Decrease accuracy
Ray cursor [Bowman and Hodges, 1999]	Motion tracker (6DOF tracker)	3D volumetric display	No	Ray cursor	-	Decrease accuracy	-	Decrease accuracy
Curve with fixed length [Dang et al., 2003]	Tracked wand	3D stereoscopic environment	No	Curve (Curved Ray cursor)	-	Decrease accuracy	-	Decrease accuracy
Silk cursor [Zhai et al., 1994]	Tracked gloves	Screen (equipped with 3D stereoscopic glasses)	No	Volume Point cursor (semi-transparent rectangular volume cursor)	-	Not applicable	Increase pointing time	Not applicable
Transparent sphere [Dang et al., 2003]	Tracked Wand	3D stereoscopic environment	No	Volume Point cursor (transparent sphere)	Circulation or List menu	-	Increase pointing time + Increase reaction time	-

Selection technique	Device used	Display	Nav.	Pointing	Disamb.	Density	Depth	Occlusion
Starfish [Wonner et al., 2012]	Tracked device	Screen	No	Volume Point cursor (3D pointer delimited by a starfish-shaped surface)	Moving further pointer	Decrease accuracy	Increase pointing time	-
Transparent cylinder [Dang et al., 2003]	Tracked Wand	3D stereoscopic environment	No	Volume Ray cursor (Infinite transparent cylinder)	Circulation or List menu	Increase reaction time	-	-
Spotlight [Liang and Green, 1993]	6DOF bat + 2 Motion trackers	Screen	No	Volume Ray cursor (Transparent cone)	Dynamic selection volume + selection algorithm (distance calculation)	Increase pointing time	-	-

Selection technique	Device used	Display	Nav.	Pointing	Disamb.	Density	Depth	Occlusion
Go-Go [Poupyrev et al., 1996]	two 6DOF Motion trackers placed on the user hands	3D VR head-set	No	Ray cursor (virtual hand)	Arm extension technique (non-linear transformation)	Decrease accuracy	Increase pointing time	Decrease accuracy
HOMER [Bowman and Hodges, 1997]	Joystick using Motion tracker	3D VR head-set	No	Ray cursor (virtual hand)	Arm extension technique (linear transformation)	Decrease accuracy	-	Decrease accuracy
WIM [Stoakley et al., 1995]	Buttonball + clipboard (each carried a position and orientation tracker + Head magnetic tracker)	Head Mounted Display (HMD)	Yes	Point cursor	-	Decrease accuracy	Increase pointing time	-

Selection technique	Device used	Display	Nav.	Pointing	Disamb.	Density	Depth	Occlusion
Ray with variable length [Dang et al., 2003]	Tracked wand	3D stereoscopic environment	No	Ray cursor	Variable length technique (variable length ray)	Increase pointing time	-	Decrease accuracy
Curve with variable length [Dang et al., 2003]	Tracked wand	3D stereoscopic environment	No	Curve (Curved Ray cursor)	Variable length technique (variable length curve)	Increase pointing time	-	Decrease accuracy
Flexible pointer [Feiner, 2003]	two 6DOF Motion trackers placed on user hands	Screen	No	Curve (Curved Ray cursor)	Curved ray	Increase pointing time	-	Decrease accuracy
Depth ray [Grossman and Balakrishnan, 2006]	Wireless mouse	3D volumetric display	No	Ray cursor	Manipulating further pointer (Depth marker)	Increase pointing time	-	Decrease accuracy

Selection technique	Device used	Display	Nav.	Pointing	Disamb.	Density	Depth	Occlusion
Circulation technique [Dang et al., 2003]	Tracked wand	3D stereoscopic environment	No	Volume Point cursor (sphere) or Volume ray cursor (cylinder) or Ray cursor	Circulation		-	Increase pointing time
Two ray selection technique [Dang et al., 2003]	Tracked wand	3D stereoscopic environment	Yes	Ray cursor	Navigation phase (two rays)	Increase pointing time	-	-
Shadow cone [Steed and Parker, 2004]	Motion tracker + joystick	Spatially immersive display (SID)	Yes	Volume Ray cursor	Navigation phase (two virtual cones)	Increase pointing time	-	-
Smart ray [Grossman and Balakrishnan, 2006]	Wireless mouse	3D volumetric display	Yes	Ray cursor	Navigation phase (several rays) + selection algorithm (weight)	Increase reaction time	-	-

Selection technique	Device used	Display	Nav.	Pointing	Disamb.	Density	Depth	Occlusion
Ray cursor + Transparent sphere [Dang et al., 2003]	Tracked wand with button	3D stereoscopic environment	No	Ray cursor + dynamic volume Point cursor (sphere)	Dynamic selection volume	-	-	Increase pointing time
Ray cursor + Bubble cursor [Vanacken et al., 2009]	Tracker with button	Screen	No	Ray cursor + dynamic volume Point cursor (dynamic sphere shape transformation)	Dynamic selection volume	-	-	Decrease accuracy
Flower ray [Grossman and Balakrishman, 2006]	Wireless mouse	3D volumetric display	No	Ray cursor	2D flower menu	Decrease accuracy + Increase reaction time	-	-
Daisy menu [Liang and Green, 1993]	6DOF bat + 2 Motion trackers	Screen	No	Ray cursor	Virtual sphere menu	Decrease accuracy + Increase reaction time	Increase pointing time	-

Selection technique	Device used	Display	Nav.	Pointing	Disamb.	Density	Depth	Occlusion
Ring menu [Liang and Green, 1993]	6DOF bat + 2 Motion trackers	Screen	No	Ray cursor	Circular menu	Decrease accuracy + Increase reaction time	Increase pointing time	-
Floating menu [Dang et al., 2003]	Tracked wand	3D stereoscopic environment	No	Ray cursor	List menu	Decrease accuracy + Increase reaction time	Decrease accuracy	-
Virtual pointer metaphor [Steinicke et al., 2006]	6DOF Optical tracker (or space mouse)	3D VE	No	Ray cursor	Selection algorithm (minimum distance to the ray) or Dynamic selection volume + selection algorithm	Increase pointing time	-	-

Selection technique	Device used	Display	Nav.	Pointing	Disamb.	Density	Depth	Occlusion
SQUAD [Kopper et al., 2011]	Motion tracking Wand	Screen	No	ray cursor + volume Point cursor (transparent sphere)	Object rearrangement on the screen : (four-item menu)	Increase reaction time	Decrease accuracy	-
Expand [Cashion et al., 2012]	Motion tracker	Screen	No	Ray cursor	Object rearrangement on the screen : (2D virtual grid + Zoom)	Increase reaction time	Decrease accuracy	-

Selection technique	Device used	Display	Nav.	Pointing	Disamb.	Density	Depth	Occlusion
Semantic 3D pointing [Elmqvist and Fekete, 2008]	Mouse	Screen	No	Point cursor	Object rearrangement on the screen : (quadtree representation) + Selection algorithm (spatial query algorithm : non-linear scale function controls the ratio)	Increase reaction time	-	Decrease accuracy

Selection technique	Device used	Display	Nav.	Pointing	Disamb.	Density	Depth	Occlusion
Jump and Refine [Hachet et al., 2007]	Directional input buttons of the mobile phones	Smartphone screen	Yes	Point cursor	Navigation-phase (Zoom : cursor is moved by successive jumps from cells to cells)		Increase pointing time	Decrease accuracy
Touch Selection Interface [Vemavarapu and Borst, 2013]	Camera based tracker	Smartphone screen	Yes	Ray cursor	Object rearrangement on the screen : (Displaying a 2D overview of a 3D sub region of the VE)	Decrease accuracy + Increase reaction time	Increase pointing time	-

Selection technique	Device used	Display	Nav.	Pointing	Disamb.	Density	Depth	Occlusion
Dual-Finger Midpoint Ray-Casting [Telkenaroglu and Capin, 2013]	Smartphone multi-touch screen + two fingers	Multi-touch screen	No	Ray cursor + Point cursor (midpoint location of two fingers positions)	Navigation phase (Zoom)	Decrease accuracy	-	Decrease accuracy
Dual-Finger Offset Ray-Casting [Telkenaroglu and Capin, 2013]	Smartphone multi-touch screen + one finger	Multi-touch screen	No	Ray cursor + Point cursor	moving further pointer (Adding an offset to the finger position)	Decrease accuracy	-	Decrease accuracy
DrillSample [Mossel et al., 2013a]	Smartphone touch-screen	Smartphone screen	No	Ray cursor	Object rearrangement on the screen (Selected objects aligned on a horizontal line)	Increase reaction time	Increase pointing time	-

Selection technique	Device used	Display	Nav.	Pointing	Disamb.	Density	Depth	Occlusion
LayerStroke [Wu et al., 2015]	Smartphone touch-screen	Smartphone screen	No	Point cursor	Object rearrangement on the screen (Algorithm to Rearrange objects into Layers and tiles)	Increase reaction time	Increase reaction time	Increase reaction time

Chapitre III

DichotoZoom

1 Interaction technique

We studied in the previous chapters the selection interaction techniques in a 3D VE and we discussed and classified these techniques presented in literature. These techniques are used to partially solve the three selection problems and to select an object. This is due to lack of interaction techniques to interact with a 3D dense virtual environment composed of many objects and displayed on a Smartphone.

As we have seen, to ease the selection of a target in a 3D VE, many techniques are based on a ray emission or a cursor with a volume selection area. Due to the environment density of the 3D VE, when pointing in the scene with such techniques, several elements of the scene can be selected at once. Therefore, in order to finalize the selection task, an additional selection phase of these pre-selected items is necessary.

In this chapter, we will define the DichotoZoom technique [Balaa et al., 2018] then we will compare and evaluate it to the Circulation technique, suggested by the literature. Using this evaluation, we focus our work on the selection of an object in a 3D VE displayed on a Smartphone in which a large number of objects are present. We have chosen to evaluate more specifically on the selection of an item from a set of pre-selected objects. This study shows the effectiveness of DichotoZoom when the set contain significantly numerous elements.

In the following sections, we will describe the two interaction techniques : the DichotoZoom and the Circulation techniques using physical buttons then we will detail the experimental protocol and the results of the evaluation used to compare these two interaction techniques for selecting an object in a 3D environment.

1.1   Principle

DichotoZoom is a selection technique that minimizes the number of actions used to achieve the desired target. It is an object selection technique in a 3D virtual environment displayed on a Smartphone. DichotoZoom is used to select a target from N already pre-selected objects ; it is a second step of the selection-pointing task. Therefore, this technique can be associated with any pointing task solution that lets you select several targets simultaneously while using a Point cursor techniques or a Ray casting techniques.

DichotoZoom interaction technique is based on the principle of the dichotomy. After each selection action performed by the user, the number of preselected targets is divided by two. At each step, the set of preselected targets is divided into two sets of an equal number of targets. The sets are created according to the depth of each target from the user point of view (cf. Figure 62). The N / 2 targets closest to the user's point of view constitute the first set (purple objects of the Figure 62), while the more distant N / 2 constitute the second set (yellow objects of the Figure 62). The user then chooses the set in which the target is located. Then, the 3D VE only displays the selected subset of targets and readjusts the user's point of view if the farthest subset was chosen. The operation is then repeated on this new subset until the desired object is located in the foreground of the 3D VE, and the user can select it without constraints.

1.2   Implementation

To implement DichotoZoom in 3D VE displayed on Smartphone, we combined it to the ray casting technique and a cylinder, in order to be able to pre-select a
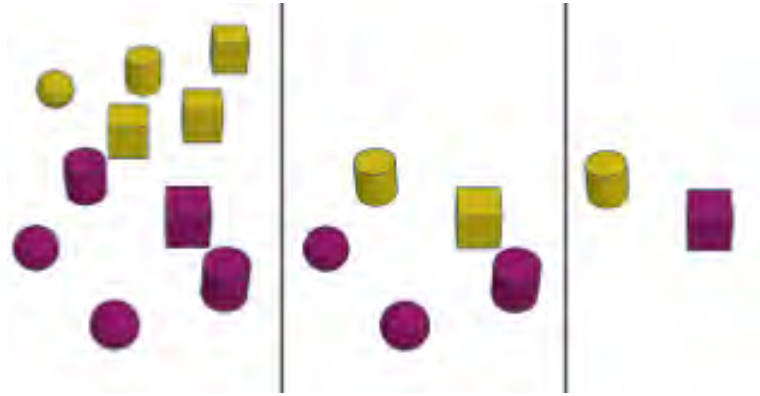
FigURE 62 – Example of selection in a set of preselected elements using Dichoto-Zoom : the target to reach is the represented cube in purple on the left image

set of targets in a complete 3D VE. We chose to replace the ray by a cylinder in order to select a larger number of objects, and facilitate the selection of the right target. To emit the ray, the user presses the screen with a finger. As long as he has not raised the finger off the screen, he can move the ray in the scene. The diameter of the cylinder can be adjusted using the two-finger interactions generally used to "zoom in" and "zoom out" on the screen [Lao et al., 2009]. By raising the fingers off the screen, the targets in the cylinder are then preselected.

Once the selection has been made with the cylinder, the preselected elements in the front zone and those in the farthest zone are respectively colored in purple and yellow as shown in Figure 63. The selection of the front zone or the rear zone is then done by pressing the volume level buttons : the "UP" button to choose the nearest set of elements and the "DOWN" button to select the farthest set. The user can also return to the previous set by pressing the "backspace" key.

Once the desired object is in the foreground of the 3D VE without being obscured by other elements, the user can finish his selection by tapping on the screen at the object position.

In order to facilitate the perception of the two subsets defined by DichotoZoom, the preselected elements in the front zone and those in the farthest zone are respectively colored in purple and yellow as shown in Figure 63. We made this choice in our
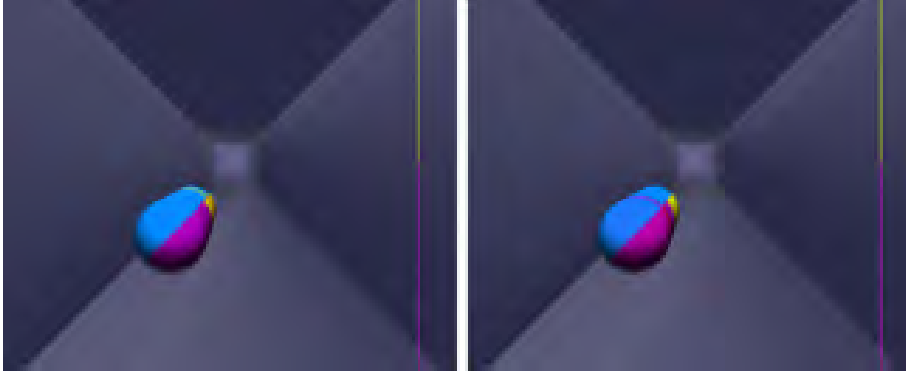
FIGURE 63 – Set of preselected elements : the target to reach is in the farthest zone and colored in yellow (on the left) and in the nearest zone and colored in purple (on the right)

study to easily distinguish the two subsets. In real 3D VEs, where the color of each object is important, other solutions may be considered to distinguish between the two subsets. For example, it would be possible to color only the outline of objects, or to put a 2D plane between the two sub-sets to better distinguish the boundary between the two subsets.

On the other hand, in this first implementation of DichotoZoom, we choose to not display the environment objects that are not selected. Thus, after each iteration, the number of objects displayed is reduced to the half. However, in a real case, we could consider other solutions where all the objects remain visible to keep a global view of the 3D VE. For example, elements that are no longer part of the subset could be displayed in transparency accordingly ; the user can still perceive them and place them in relation to other objects that remain selectable.

Finally, user needs only three separate actions to make DichotoZoom work : two actions to select the near or the far subsets, then an action to undo the last performed selection. For this first study, we chose using the Smartphone buttons to perform these three actions. We made this choice to avoid : learning gestures, detection or recognition problems that we could have with gestures on the screen or with the Smartphone. Therefore, user selects the closest or furthest zone from his point of view respectively with the "UP" and "DOWN" volume buttons. He can cancel the

last action by pressing the "Back" button. Finally, the validation of a selection is done by pressing the touchscreen. Consequently, the target having the least distant in depth is selected.

1.3 Theoretical contribution

The circulation technique requires circulating the environment objects to reach the desired target. Therefore, to reach a target at the Nth position of a set of objects, the user performs N operations before reaching the object in question. Then, the number of actions to achieve and select a desired target is proportional to : first, the density and the number of objects at the environment and second, the depth position from the point of view of the user. If the target is at a far position from the user and in a crowded environment, the user must perform a large number of actions to reach this object.

For a set of N objects, the number of actions to be performed to reach the desired object will be between 1 and N depending on its position at the environment. In case of DichotoZoom, the number of targets present in the set is divided by two after each user action. Therefore, the user will have to perform a maximum of $log_2(N)$ actions to achieve the desired target within a preselected set of N elements. Then, for a higher number of objects in the environment, the DichotoZoom will show the best performance in reaching the desired target from the Circulation technique.

In both cases, the act of passing from one object to another by pressing a button of the device can limit the accuracy problems.

2 A comparative study : comparing the Dichoto-Zoom technique to the Circulation technique

2.1 User evaluation

2.1.1 Objectives

The main objective of this study is to evaluate the relevance of the use of a dichotomy algorithm for the selection of an object. Before even studying the best possible interaction with this algorithm, our goal is to verify that a selection using a dichotomy algorithm does not disturb the user compared to a classic selection called "step by step".

We notice two main risks of disturbances for the user. On one hand, the selection using the dichotomy algorithm could produce more errors compared to the classic selection. On the other hand, the search for the target in the new subset takes too long for a visual search after iteration ; thus, this result might be inconvenient to the user. Therefore, these are various factors related to the perception of the user that we wish to analyze before studying the different interaction modalities that we could use to interact with this algorithm.

The objective of this study is to evaluate and compare the performance and the usability of the proposed technique : DichotoZoom interaction technique. Indeed, we look for measuring the effectiveness of this interaction technique in the 3D pointing tasks : DichotoZoom allows the selection of objects inside a set of already pre-selected ones, with a reduced number of actions. In order to evaluate the effectiveness of the DichotoZoom technique, we performed an evaluation by comparing the DichotoZoom technique to the reference technique : the Circulation technique suggested by the literature.

This experimental study includes two different sessions using respectively the DichotoZoom (DZ) interaction technique using physical buttons and the Circulation interaction technique (CC) using physical buttons to select a target placed randomly in a dense 3D VE. The 3D VE contains significantly numerous objects aligned in

depth. This study shows the effectiveness of DichotoZoom when the set contain significantly numerous elements.

2.1.2 Hypotheses

In case of the DichotoZoom interaction technique, due to the dichotomy algorithm, the number of objects present in the environment is divided by two after each user's action. Consequently, the DichotoZoom technique allows to significantly reducing the number of actions that users have to realize to reach a specific target.

Pointing with DichotoZoom is faster than pointing with Circulation and this speed is discerned for deeper targets according to the algorithm of dichotomy. The latter reduces the time of the search and the number of actions.

However, the DichotoZoom reconfigure the space after each action. The target to be reached can change zone after each action of the user. Consequently, the user is required to wait after each action to see in which zone the desired target is located. Therefore, the user cannot anticipate his actions as he can do with the circulation technique where it can provide roughly the number of actions to perform before beginning the selection task. Hence, we assume that the time required to perform an action using DichotoZoom will be greater than the time consumed by the circulation, because of the time used for the discernment of the next zone (H1).

Finally, despite dropping time in searching for the target and programming the next action, we assume that the task completion time will be reduced by using DichotoZoom because the user will gain in task's time through the decrease in the number of actions to perform (H2).

2.1.3 Participants

We recruited twelve participants : ten males and two females, aged 20 years on average (SD=1). Eleven of them were right-handed. All of them are students at the university and are daily users of a Smartphone.

2.1.4 Apparatus

The material used in the experiment is a Smartphone Samsung Galaxy SII having a 1.2 GHz dual-core processor and running Android 2.3.4. The touchscreen has a size of 4.3 inches diagonally, and a resolution of 480 x 800 pixels. The 3D virtual environment was implemented using Open GL. During the experiment, the participants were seated at a desk. They could hold the smartphone as they wish.

2.1.5 The choice of Task and Base technique

As we have seen previously, the selection of target in 3D VE with high density of objects generally breaks down into 3 subtasks : navigation of the environment, selection of a subset of objects via a ray or a volume, then selecting the desired object from that subset. The dichotomy algorithm that we propose only intervenes in the third subtask for the selection of an object from a subset of many objects.

Techniques such as menu [Dang et al., 2003] or SQUAD [Kopper et al., 2011] operate on these three subtasks. However, the preselected objects of these two techniques are reorganized respectively on a linear menu or on a 2D plane to be more easily selectable. In our case study, we want to keep the original disposition and place of the objects in the environment so that the user will not be disturbed when performing the third subtask. Accordingly, we finally chose the Circulation technique [Dang et al., 2003] as the basic technique to be compared with DichotoZoom. Since, the first two subtasks (navigation and selection of the subset of objects) can be performed in the same way with both techniques ; we focused our study exclusively on the last subtask. Participants did not have to perform an initial selection with a ray or volume to determine the subset. The task, we asked the participants for, was to select a target from a set of objects. This set would correspond, in the complete task, to the subset of selected objects. To manipulate the Circulation technique, interactions are also done using the physical buttons. The user circulates the objects to reach the target by using the volume level buttons : the "DOWN" button is used to move towards the target in the depth direction and the "UP" button is used to move in the opposite direction towards the user. So the user will circulate the spheres to

achieve the target. The user also has the option to return to the previous object by pressing the "Back" button and perform an Undo operation. Finally, in the same way as for DichotoZoom, the user presses on the screen to validate his selection.

2.1.6 Procedures

For both sessions, users were advised to select a spherical red and white target randomly from a set of 32 blue and white spheres displayed in a 3D VE. The instruction was to use the proposed selection technique to display the desired target in the foreground of the 3D VE. To simulate depth and partial occlusion, the 32 spheres are aligned in depth so they are all placed at different Z coordinates in depths (the distance between them is greater than the diameter of the spheres), but with quasi identical coordinates in the (X, Y) plane (on the 2D plane the distance between the centers of two successive targets does not exceed the value of the radius of the targets). Thus, we will simulate the occlusion and the density of the spheres present in the set. The spheres were displayed in semi transparency to be distinguished by the participants. The level of transparency has been defined empirically prior to experimentation. Since the evaluation concerns only the selection technique from a set of pre-selected items, the participants did not have to make the selection phase with the cylinder. The set of 32 spheres displayed on the screen is considered as the result of the selection when using the cylinder. The user's task consists of pointing the red and white sphere as soon as possible and accurately.

Using the Circulation technique, the user can circulate from one target to another by pressing the physical "UP" volume button, and can return to the previous sphere by pressing the "Backspace" button. Using the DichotoZoom technique, the user selects the closest area or furthest area from the user's point of view with the "UP" and "DOWN" volume buttons respectively. He can go back by pressing the "Backspace" button.

Finally, for both techniques, once the desired target is in the foreground in the 3D environment, the user validates the selection by pressing the screen at the target position. Consequently, the task is finished and the target is selected. Otherwise, the

user finished the selection task by pressing the screen at the position of any sphere of the VE different from the target and consequently, the test failed.

The validation instructions given to the participants were to point the target with the finger when it is placed at the first position. However, the evaluation did not focus on the accuracy of the score at the time of validation, we considered the task completed when the user pressed on the screen, even if the target was not correctly pointed and we consider as validated target the one at the first position on the 3D VE. Consequently, we count an error if the pointed target in the foreground is not the requested target. However, we do not count an error if the pointing is not carried out correctly on the nearest object but that this one is the target.

2.1.6.a   DichotoZoom interaction technique : using physical buttons   The experimental procedure of the DichotoZoom interaction technique using physical buttons :

After displaying the 3D environment, the spheres are shown as follow : the target is displayed in red and white and the environment spheres are displayed in blue and white. The user is invited to click the volume button (the UP button or the DOWN button) to trigger the DichotoZoom technique. Consequently, the pre-selected objects are displayed. 32 colored spheres are displayed with zoom. The spheres are displayed regarding their position in depth and the 3D VE maintains its initial representation. The spheres are divided into 2 sets according to their number and their locations in depth relatively to the user point of view and classified in Zone1 and Zone2. Zone1 contains the first half of spheres formed by the nearest targets from the user point of view. Zone2 contains the second half of the spheres consisting of the farthest targets. Spheres change their colors. The spheres of Zone1 are colored of blue and purple and the spheres of Zone2 are colored of blue and yellow. The target also changes its color : it has red and purple colors if it is located at Zone1 or red and yellow colors if it is located at Zone2. In addition, user instructions are displayed : a text explaining the selection procedure and a colored vertical line showing the yellow zone and the purple zone.

During the pointing task, participants were asked to reach the target by selecting

the zone where it is located : by clicking the UP volume button to select Zone2 (the yellow zone) or by clicking the DOWN volume button to choose Zone1 (the purple zone). In case of selection error, the user also has the option to return to the previous set by pressing the "Back" button and perform an Undo operation.

2.1.6.b Circulation interaction technique : using physical buttons The experimental procedure of the Circulation interaction technique using physical buttons :

After displaying the 3D environment, the spheres are shown as follow : the target is displayed in red and white and the environment spheres are displayed in blue and white. The user is invited to click the volume button (the UP button or the DOWN button) to trigger the Circulation technique. Consequently, the pre-selected objects are displayed. 32 colored spheres are displayed with zoom. The spheres are displayed regarding their position in depth and the 3D VE maintains its initial representation. The object on the first 3D VE plane is selected.

During the pointing task, the user circulates the objects to reach the target by using the volume level buttons : the "DOWN" button is used to move towards the target in the depth direction and the "UP" button is used to move in the opposite direction towards the user. So the user circulates in depth from one sphere to another to achieve the target.

Each time the user clicks the "DOWN" button, the sphere disappears and the pointer moves to the next object and each time the user clicks the "UP" button, the sphere appears and the pointer moves to the previous object.

In case of selection error, the user also has the option to return to the previous object by pressing the "Back" button and perform an Undo operation.

Once the desired object is on the first 3D VE plane without being occluded by other elements, the user can finish his selection and validate his task by clicking on the screen at the position of the target.

Finally, for both techniques, the user validates his selection by pressing the object once the desired target is located on the front plane in the 3D environment. To validate the selection task, the instruction given to participants was to point with

their fingers on the target, when it is in the front position.

However, the assessment objective does not focus on the pointing accuracy at the validation phase, we consider that the task was completed as soon as the user pointed the screen, even if the target was not properly pointed object and we consider the front object of the 3D VE as the validated target.

2.1.7 Design

Each participant performed two sessions, one for each of the techniques (DichotoZoom and Circulation). The Interaction Technique ordering was counterbalanced across the 12 users. Within each session, three blocks were run for each technique ; each block of trials required 32 positions of the target. Within each repetition, for each trial, the position of the target to achieve was different. The order of the tasks was random. For each task of the same exercise, the position of the target to be validated was different. All positions of the 3D VE were therefore used once at each exercise. Thus, participants realized a pointing task with the target present in each of the 32 possible positions. Each participant performed 2 techniques x 3 blocks x 32 positions = 192 trials.

Thus, we obtained 2304 selection tasks (12 participants x 2 techniques x 3 blocks x 32 positions).

Before using a technique for the first time, each participant was given a training session to familiarize himself with the selection technique. Participants were able to familiarize themselves with the technique they were going to use for 5 minutes before each session.

2.2 Results

2.2.1 Objective Results

To analyze our results, we rely on estimation techniques based on amplitudes effect and 95% confidence intervals, according to the recent APA recommendations.

The used scripts are those used by [Willett et al., 2015] and available online [1].

2.2.1.a   Task completion time   In our analysis, the task completion time is the time calculated from the moment when the 32 spheres of the 3D VE are displayed on the screen, until the user clicks the screen to validate his selection (the selection of the desired sphere). Average task completion times were 8111 ms (CI [7735 ; 8637]) for the Circulation technique against 7569 ms (CI [7219 ; 7943]) for the Dichoto-Zoom technique. The DichotoZoom technique allows to participants to reduce their selection time by 5.93% (CI [-0.72 ; 11.98]) compared to the Circulation technique. However, despite this reduction in time, we can see in Figure 64 that the confidence intervals overlap, this reduction of about 6%. The calculation of the ratio between DichotoZoom and Circulation gives us a ratio of 0.94 (CI [0.88 ; 1.01]). Therefore, the DichotoZoom technique allows reducing the user's selection time about 6%.
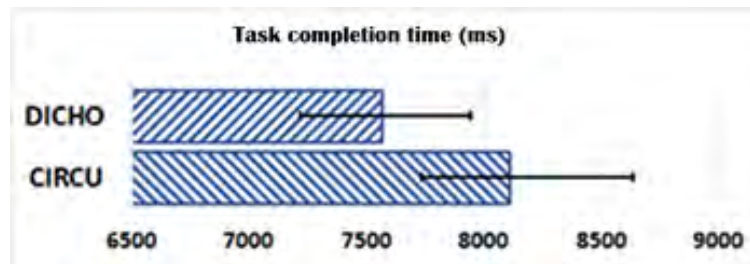
FIGURE 64 – Task completion time in ms for Circulation and DichotoZoom techniques and 95% confidence interval

2.2.1.b   Task completion time according to the series   Participants were asked to repeat the same exercise with each technique 3 times consecutively. Figure 65 shows the task completion time according to the performed exercise. Thus, we can realize that there is an important learning effect for the DichotoZoom technique where the participants performed the tasks in 8170 ms (CI [7668 ; 8728]) for the first exercise, while they performed the same tasks in 7389 ms (CI [7067 ; 7754]) for the second exercise. The ratio between the first two exercises shows that participants

---

1. http ://www.aviz.fr/reliefshearing, section "study material"

were 9.24% (CI [5.50, 13.35]) faster in the second exercise compared to the first. Moreover, we can distinguish that this learning effect fades quickly : since comparing the second and third exercises, participant performances progress much more slowly with 7189 ms (CI [6795 ; 7567]) in the third exercise, specifically, an average gain of 1.58% (CI [-1.25, 3.63]) between the second and third exercise. On the other hand, this learning effect is much less noticeable for the Circulation technique where the participants performed the tasks in 8390 ms (CI [7930 ; 9039]), 8099 ms (CI [7697 ; 8873]), and 7846 ms (CI [7443 ; 8193]) for the 3 consecutive exercises, specifically, gains of 3.33% (CI [0.57 ; 7.26]) and 2.63% (CI [0 ; 7.72]) respectively between the first and the second exercises, and for the second and the third exercises.
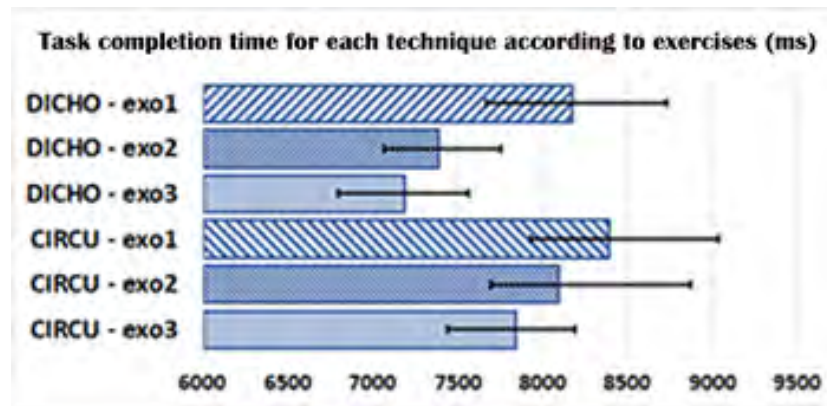
FIGURE 65 – Task completion time in ms for each technique (Circulation vs Dicho-toZoom) according to the exercise and 95% confidence interval

2.2.1.c   Task completion time according to the target position in the set
Figure 66 shows the evolution of the selection time according to the target position to be reached in the set of proposed elements. In case of Circulation technique (blue markers in Figure 66), we can realize that the selection time dependents on the target position in the set : In fact, Computing the linear regression line of the average time of selection obtained depending on the position in the set of elements gives a very high correlation coefficient (R = 0.99). Therefore, the resulting equation (y = 4031.2 + 262.8 x) predicts that for any additional position on the set, the task completion

134

time will be incremented of approximately 263 ms.

FIGURE 66 – Task completion time (ms) evolution according to the target position to be reached in the set

However, when using the DichotoZoom technique, the selection time of a target is not directly related to its position in the set as seen by the red markers in Figure 66. In fact, the calculation of the linear regression line on the average time of selection obtained with the DichotoZoom technique gives a very low coefficient of correlation (R <0.05). This can be explained by the fact that the target position is not directly related to the number of actions to take to reach it : in fact, a very distant target (such as the one located at the 25th position) can be reached with a reduced number of actions (theoretically 2 actions when the set contains 32 elements) than another target much nearer in the set (for example, the target in 7th position requires 5 actions when the set contains the same size).

2.2.1.d   Task completion time according to the number of actions to be performed   Depending on the number of preselected targets and the position of the target to be reached, the number of actions to be performed with DichotoZoom to achieve the desired target may vary. Table 2 summarizes, in our case study, the actions to be performed according to the position of the target. The characters 'U'

135

and 'D' indicate a respectively click on the "UP" button (to select the nearest zone) or on the "DOWN" button (to select the furthest zone). We deduce that in case of a preselected set of 32 targets, the user will perform a maximum of 5 navigation actions between zones to reach the desired target. Once the desired target is located on the foreground plan, the user will have to only point it on the Smartphone screen to select it. This last pointing action is not taken into consideration while counting actions in this section.

TABLE III.1 – Selection actions to perform before pointing to the desired target.

1		9	UD	17	D	25	DD
2	UUUUD	10	UDUUD	18	DUUUD	26	DDUUD
3	UUUD	11	UDUD	19	DUUD	27	DDUD
4	UUUDD	12	UDUDD	20	DUUDD	28	DDUDD
5	UUD	13	UDD	21	DUD	29	DDD
6	UUDUD	14	UDDUD	22	DUDUD	30	DDDUD
7	UUUD	15	UDDD	23	DUDD	31	DDDD
8	UUDDD	16	UDDDD	24	DUDDD	32	DDDDD

Figure 67 shows the task completion time (s) evolution according to the number of actions to be performed. In case of DichotoZoom technique (shown in red in Figure 67), the linear regression line computed by the means of the selection times shows a very high correlation coefficient (R > 0.975). This coefficient very close to 1 confirms the correlation between the selection time and the number of selection actions to be performed for the DichotoZoom technique.

While using the DichotoZoom technique, the equation of the regression line (y = 4887.8 + 675.35 x) predicts that each additional action will increase the task completion time by 675 ms. Otherwise, using the Circulation technique, despite that it is necessary to make an action to navigate one position, the equation of the regression line obtained for the number of action is similar to the one obtained for the first 5 positions (cf. Figure 67). Each additional action will increase the task completion time by 265 ms.
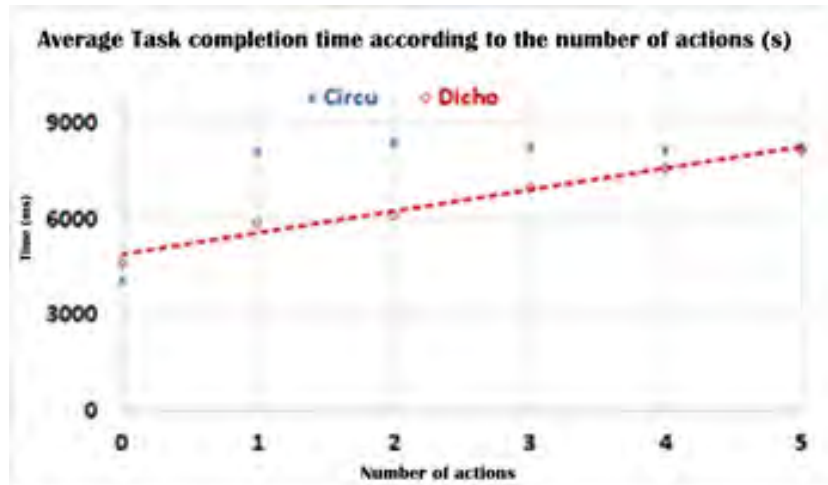
FIGURE 67 – Task completion time (s) evolution according to the number of actions to be performed

This time difference between the two techniques can be explained by the fact that with the DichotoZoom technique, after each action performed, the remaining objects are divided into two sets. The user can therefore hardly anticipate whether the object will be near or far from the point of view of the user. With Circulation, as the object selection progress in ascending order in the depth, it is easier for the user to anticipate knowing that the desired object is still distant and thus quickly perform several successive actions.

Therefore, this comparison showed that DichotoZoom technique was significantly faster than the Circulation technique for a large number of preselected objects. The large number of objects produces a great difference of the number of actions to be performed and therefore this gives advantages for the DichotoZoom technique.

Therefore, our hypothesis H1 is confirmed. Indeed, after each performed action, the remaining items are re-divided into two sets. Thus, users can hardly anticipate on the next target position and zone : if the target is near or far from the user viewpoint. However using the Circulation technique, and knowing that to point a target, the user has to circulate in depth and to pass from object to object in the same direction to reach the target thus, it is easier for the user to anticipate. If the desired object

137

is still distant from the pointer, the user can perform several successively and faster actions.

This comparison showed that DichotoZoom technique was significantly faster than the Circulation technique when the number of objects present in the pre-selected set is large enough and the environment is dense. Since the number of actions to take depends on the number of objects presented in the set, DichotoZoom technique has a higher performance than Circulation technique.

2.2.1.e   Error rate and Correction   To evaluate the techniques accuracy, we measured two factors : firstly, the use of the back or undo event during the task, and secondly, the number of wrong selections.

Figure 68 shows, for each technique, the rate of cancellation's action. Participants had the opportunity to cancel the performed selection action and to return to the previous step by pressing the "Back" button. We measured the rate of cancellation's action as the number of times participants used the back button (to cancel the selection previously made) regarding the total number of actions performed during the task. Using the Circulation technique, participants canceled one selection action in 1.24% (CI [0.76, 1.95]) of cases while they performed the same action in 2.72% (CI [0.99 ; 6, 16]) of cases with DichotoZoom technique. On the other hand, we can see in Figure 68 that the confidence interval of the DichotoZoom technique is much greater than that of the Circulation technique. This informs us that there is more variability among participants in using the cancellation of an action with DichotoZoom than with Circulation.
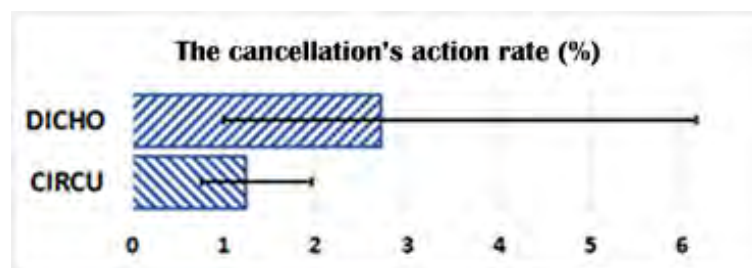
FIGURE 68 – the cancellation's action rate (%) and 95% confidence interval

We also measured the task's validation error. The task's validation error occurs once the participants validated the selection for an object different from the desired target. The task error rate is the number of times participants performed a wrong or inaccurate selection regarding the total number of task to achieve. Participants made fewer errors with the Circulation technique than with the DichotoZoom technique. The circulation technique does not have selection errors. However, the DichotoZoom technique produced in 4.46% (CI [1,06 ; 11,27]) of cases a selection error. The high error rate of DichotoZoom is explained by a very high error rate for two participants : in fact, participants 3 and 5 performed respectively 18.48% and 24.47% error. However, we can also note that 7 of the 12 participants didn't make any mistakes with DichotoZoom.

3   Discussion

We have chosen to evaluate more specifically the selection of an item from a set of pre-selected objects. Accordingly, our study compares two interaction techniques : the DichotoZoom technique and the Circulation technique. The DichotoZoom technique allows to participants to reduce their task completion time compared to the Circulation technique. Furthermore, we can realize that there is an important learning effect for the DichotoZoom technique where the participants performed the tasks for the first exercise, while they performed the same tasks for the second exercise. On the other hand, this learning effect is much less noticeable for the Circulation technique.

Our study shows that the response time required for executing one additional action when using the Circulation technique is smaller than the required response time when using the DichotoZoom technique. Since, user can anticipate his action while using the circulation technique, however he will lose more time for the discernment of the next zone while using the DichotoZoom technique. Accordingly, this result confirms our first hypothesis.

In case of Circulation technique, we can realize that the selection time dependents on the target position in the set. The task completion time will be incremented for

any additional position on the set. However, when using the DichotoZoom technique, the selection time of a target is not directly related to its position in the set, since it depends on the dichotomy algorithm. Therefore, this comparison showed that DichotoZoom technique was significantly faster than the Circulation technique for a large number of preselected objects. The large number of objects produces a great difference of the number of actions to be performed and therefore this gives advantages for the DichotoZoom technique. The task completion time analysis showed that DichotoZoom technique was significantly faster than the Circulation technique to achieve the selection task for a large number of preselected objects. Consequently, our second hypothesis is confirmed. Comparing the rate of cancellation's action, the difference between the two techniques was not significant ; consequently, the two techniques are accurate. Participants made fewer errors with the Circulation technique than with the DichotoZoom technique. In both cases, the act of passing from one object to another by pressing a button of the device reduces the accuracy problems.

4 Summary

Selecting an object in 3D crowded VE holding a high density of objects is often divided into two phases : an initial volume selection of set of objects, then a refinement selection within this set of pre-selected elements. The proposed DichotoZoom technique allows to quickly selecting an object within this set. DichotoZoom is a selection technique that minimizes the number of actions to take to achieve the desired target. This study shows the effectiveness of DichotoZoom when the set contains significantly numerous elements. Then, the higher the number of objects in the environment, the DichotoZoom will show the best performance in reaching the desired target from the Circulation technique.

Chapitre IV

Experiment Evaluation of the DichotoZoom interaction technique modalities

1   Introduction

The variety of input devices that may have mobile devices allows imagining many 3D interaction techniques. Multiple studies described in literature show different ways of interaction using the Smartphone input output devices.

The presence of a keyboard and buttons is a common factor in the vast majority of mobile terminals. Therefore, it is important to study how to interact with a 3D environment using the device keys.

Smartphones provide the user with the possibility to interact with a 3D scene using a touchscreen or a multi-touchscreen. The touchscreen then provides the ability to manipulate a cursor continuously.

A large majority of today's Smartphones is equipped with a video camera having a high resolution and a potential to be compared to professional digital cameras. Then, the possibilities of using a camera to interact with a mobile device in a 3D scene are large.

Some mobile devices are equipped with special sensors such as accelerometers, positioning system and magnetometer, which can be very useful for interacting with a 3D environment. These devices allow you to add degrees of freedom to the Smartphones or to provide a smoother and more intuitive interaction.

Consequently, the most suitable interaction techniques to the Smartphone are restricted to the capabilities offered by interfaces (touchscreen, physical button, and camera) and sensors (accelerometer, gyroscope) integrated on the Smartphone.

Having shown on a preliminary study and through a comparative study, selecting using a dichotomy algorithm does not disturb the user and allows him to select a target more quickly after a little learning phase of the technique. This second study aims to find the best mode of interaction allowing the user to interact with this selection technique on a Smartphone.

In this chapter, we will evaluate the use of DichotoZoom technique with various interaction modalities available on Smartphone : using physical or graphical buttons, using gestural interactions via touchscreen or moving the device itself.

2 DichotoZoom multimodal interactions

The DichotoZoom technique is based on a discrete selection and established on two main actions : the selection of the near area or the far area from the user point of view. To refine the interaction technique, we studied the use of different interaction modalities that seemed most suitable to make discrete selection and therefore to quickly differentiate between the two areas. Hence, we chose to study the selection by using discrete actions (pressing a button) or by using continuous action (a distinctive gesture). And, these selections in two forms : in physical form by using the smartphone itself or in a digital form by interacting directly on the display surface (the touchscreen). Thus, the combination of these two features led us to study four interaction modalities (see Table IV.1) : firstly, the button interaction which is based on the use of physical buttons ; secondly, the touch interaction, which is based on the use of numeric buttons ; thirdly, the Tilt interaction, which is based on the use

of physical gestures ; and finally, the swipe interaction which is based on the use of digital gestures.

TABLE IV.1 – Evaluated DichotoZoom multimodal interactions for target selection

		Interaction	
		Discrete	Continuous
Device	Physical *Smartphone*	Button	Tilt
	Digital *Touchscreen*	Touch	Swipe

Their respective tasks with the DichotoZoom technique are described in the following sections :

2.1 DichotoZoom interaction technique : using physical buttons

Button is the interaction modality based on the Smartphone physical buttons. The volume buttons are used to choose between zones. The user chooses the near area by clicking the "DOWN" volume button and the far area by clicking the "UP" volume button from the user point of view as shown in Figure 69. The "back" button is used to cancel and undo the last selection.

2.2 DichotoZoom interaction technique : using gestural interactions via touchscreen

Touch is the interaction technique based on the use of digital buttons displayed on the screen. The buttons are displayed on the edge of the screen to not disturb the use of the 3D VE. The buttons for selecting the near or far area are respectively at the top and bottom corners of the thumb side, so that they can be easily accessed by the thumb of the hand holding the smartphone as shown in Figure 70. The cancel button for the last action is at the bottom of the screen in the opposite corner.
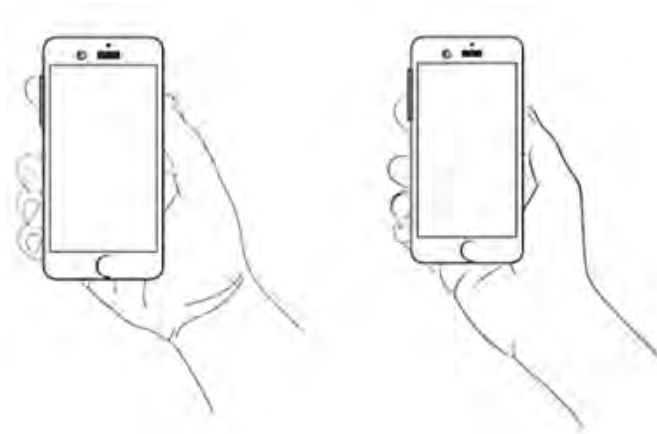
FIGURE 69 – DZ using physical buttons : Left : clicking the "UP" volume button, Right : clicking the "DOWN" volume button

2.3 DichotoZoom interaction technique : using gestural interactions via Accelerometer

Tilt is the gestural modality based on the movements (the inclinations) made directly on the Smartphone. The inclinations are calculated according to a neutral position of the Smartphone. The neutral position is defined as the smartphone horizontal position, with the touchscreen facing the sky. A tilt action consists of moving the Smartphone from its neutral position and tilts it to the desired direction, then returning it to its neutral position. The user tilts the Smartphone screen towards him to choose the near area from the user point of view in the 3D VE ; and tilts the Smartphone screen forward to choose the farthest area as shown in Figure 71. To cancel and undo the last selection, the user must tilt the Smartphone to the left.

2.4 DichotoZoom interaction technique : using gestural interactions via swipe

Swipe is a gestural interaction made directly on the touchscreen. The user makes gestures from the top towards the screen's bottom to select the near area from the user point of view and from the bottom towards the screen's top to select the far
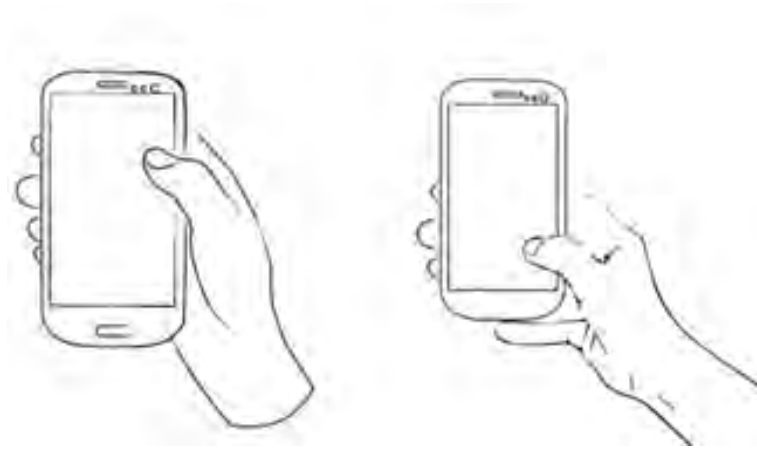
FIGURE 70 – DZ using gestural interactions via touchscreen. Left : clicking at the top right screen's area, Right : clicking at the bottom right screen's area
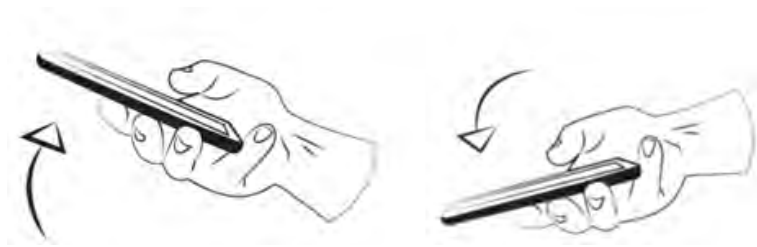
FIGURE 71 – DZ using gestural interactions via Accelerometer. Left : tilting the Smartphone towards the user's body, Right : tilting the Smartphone faraway the user's body

area as shown in Figure 72. A gesture from right to left is used to cancel and undo the last selection.

3   User evaluation

The objective of this study is to evaluate and compare performance, usability and user preference of the four proposed interaction modalities of the DichotoZoom interaction technique. Indeed, we are seeking for measuring and comparing the effectiveness of the four DichotoZoom interaction modalities while pointing a target in
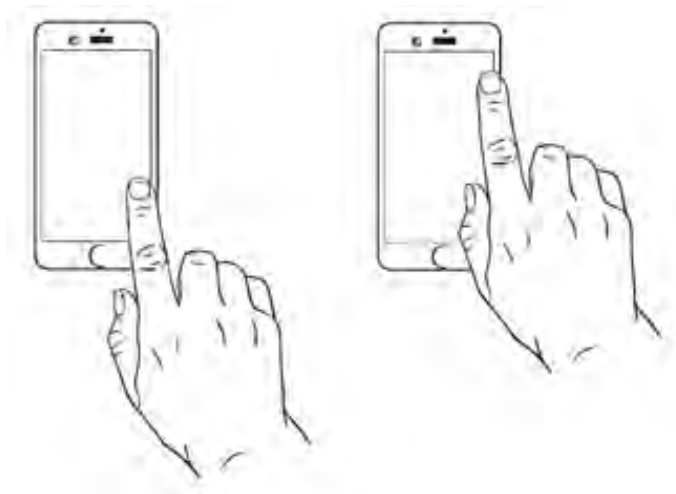
FIGURE 72 – DZ using gestural interactions via swipe. Left : making gestures from the top towards the screen's bottom, Right : making gestures from the bottom towards the screen's top

3D VE.

The four interaction modalities : the Button, the Touch the Tilt and the Swipe are introduced in the previous section.

In order to evaluate the effectiveness of the four DichotoZoom interaction modalities, we performed an evaluation by comparing these four interaction techniques while pointing a target in a 3D VE. This experimental study includes four different sessions using respectively the four interaction modalities to select a target placed randomly in a dense 3D VE. The 3D VE contains significantly numerous objects aligned in depth. This study compares the effectiveness of the four interaction modalities DichotoZoom when the set contain significantly numerous elements. The study will concern both the user satisfaction and preference, and the pointing performance criteria.

3.1 Hypotheses

We can divide the four DichotoZoom interaction modalities into two groups based on the interaction types : the button interaction and the gestural interaction. The

button interaction is composed of the physical and virtual button interactions, which are respectively the Button and the Touch modalities. The gestural interaction is composed of the swipe and the Smartphone movement, representing the Swipe and the Tilt modalities. The button interaction is characterized by the simple pointing action without changing on the 3D VE display and view. The gestural interaction is characterized by the simple pointing action with movement which can occlude the 3D VE for a while.

However, the Tilt modality is the most interesting modality since it is stylish and the easiest to be understood by the user : the user will adjust the position of the Smartphone depending on the position of the target. Using gestural interaction, user will lose time and concentration on the 3D VE while pointing a target. Hence, we assume that the interaction modality had an influence on the task completion time. The faster modality is the Touch and the slowest modality is the Tilt (H1). The time required to perform an action using Tilt modality will be greater than the time consumed by the Button modality, because of the time used for the readjustment of the neutral Smartphone position and the concentration on the new 3D VE configuration.

Since pointing using DichotoZoom is a simple choice between two areas, the four interaction modalities are accurate. We assume that the Touch and the Button modalities are more accurate than the Tilt since the user does not need to move the Smartphone (H2).

3.2   Participants

We recruited twelve participants : six males and six females, aged 31.4 years on average (SD=4). Nine of them were right-handed and three of them were left-handed. All of them are daily users of a Smartphone.

3.3   Apparatus

The material used in the experiment is a Smartphone Samsung Galaxy SII having a 1.2 GHz dual-core processor and running Android 2.3.4. The touchscreen has a size of 4.3 inches diagonally, and a resolution of 480 x 800 pixels. The 3D virtual environment was implemented using Open GL.

3.4   Procedures

For all the four sessions, users are advised to select a spherical red and white target randomly from a set of 32 blue and white spheres displayed in a 3D VE. To simulate depth and partial occlusion, the 32 spheres are aligned in depth. They are all placed at different Z coordinates in depths, but with quasi-identical coordinates in the (X, Y) plane. The spheres were displayed in semi transparency to be distinguished by the participants. Since the evaluation concerns only the selection technique from a set of pre-selected items, the participants did not have to make the selection phase with the cylinder. The set of 32 spheres displayed on the screen is considered as the result of the selection when using the cylinder. The user's task consists of pointing the red and white sphere as soon as possible and accurately.

Afterward, the user validates the selection by pressing the screen at the target position consequently, the task is finished and the target is selected. Otherwise, the user finished the selection task by pressing the screen at the position of any sphere of the VE different from the target and consequently, the test failed.

After displaying the 3D environment, the spheres are shown as follow : the target is displayed in red and white and the environment spheres are displayed in blue and white. The user is invited to click the volume button (the UP button or the DOWN button) to trigger the DichotoZoom technique. Later, the user chooses the set where the target is located. The user repeats the selection step until the desired object is located on the 3D VE front plane. Finally, he confirms the selection by pressing the screen at the position of the red and white sphere.

Once the selection is made using the cylinder, the user is invited to click the volume button (UP or DOWN) consequently, the pre-selected objects are displayed.

148

32 colored spheres are displayed with zoom. The spheres are displayed regarding their position in depth and the 3D VE maintains its initial representation. The spheres are divided into 2 sets according to their number and their locations in depth relatively to the user point of view and classified in Zone1 and Zone2. Zone1 contains the first half of spheres formed by the nearest targets from the user point of view. Zone2 contains the second half of the spheres consisting of the farthest targets. Spheres change their colors. The spheres of Zone1 are colored of blue and purple and the spheres of Zone2 are colored of blue and yellow. The target also changes its color : it has red and purple colors if it is located at Zone1 or red and yellow colors if it is located at Zone2. In addition, user instructions are displayed : a text explaining the selection procedure and a colored vertical line showing the yellow zone and the purple zone.

During the pointing task while using the Button interaction modality, participants were asked to reach the target by selecting the zone where it is located : by clicking the UP volume button to select Zone2 (the yellow zone) or by clicking the DOWN volume button to choose Zone1 (the purple zone). In case of selection error, the user also has the option to return to the previous set by pressing the "Back" button and perform an Undo operation.

During the pointing task while using the Touch interaction modality, participants were asked to reach the target by selecting the zone where it is located : by clicking on the top right of the screen to select Zone2 (the yellow zone) or by clicking on the bottom right of the screen to choose Zone1 (the purple zone). In case of selection error, the user also has the option to return to the previous set by clicking on the top left of the screen and perform an Undo operation.

During the pointing task while using the Tilt interaction modality, participants were asked to reach the target by selecting the zone where it is located : by turning the Smartphone backward to select Zone2 (the yellow zone) or by turning the Smartphone forward (towards the user direction) to choose Zone1 (the purple zone). In case of selection error, the user also has the option to return to the previous set by tilting the Smartphone to the left and perform an Undo operation. During the pointing task while using the Swipe interaction modality, participants were asked to

149

reach the target by selecting the zone where it is located : by sliding and moving the user finger from the screens bottom to the screens top to select Zone2 (the yellow zone) or by sliding and moving the user finger from the screens top to the screens bottom to choose Zone1 (the purple zone). In case of selection error, the user also has the option to return to the previous set by sliding and moving the user finger on the screen from right to left and perform an Undo operation.

Finally, for all techniques, the user validates his selection by pressing the object once the desired target is located on the front plane in the 3D environment. To validate the selection task, the instruction given to participants was to point with their fingers on the target, when it is in the front position.

However, the assessment objective does not focus on the pointing accuracy at the validation phase, we consider that the task was completed as soon as the user pointed the screen, even if the target was not properly pointed object and we consider the front object of the 3D VE as the validated target.

3.5   Design

Every one of the twelve participants performed a session for each of the four proposed interaction techniques (the Button, the Touch the Tilt and the Swipe). The Interaction Technique ordering was counterbalanced with a Latin square across the 12 users so that each technique is tested the same number of times at each position. While comparing the DichotoZoom technique to the Circulation technique, we notice that the participant performances evolved mainly between the first and the second exercise and remained almost stable during the third one. Consequently, within each session, two blocks were run for each technique ; each block of trials required 32 positions of the target. Within each repetition, for each trial, the position of the target to achieve was different. Thus, participants realized a pointing task with the target present in each of the 32 possible positions. Each participant performed 4 techniques x 2 blocks x 32 positions = 256 trials. Thus, we obtained 3072 selection tasks (12 participants x 4 techniques x 2 blocks x 32 positions).

Before using a technique for the first time, each participant was given a training

session to familiarize himself with the selection technique.

3.6   Collected Data

We measured the usability of each technique via the SUS questionnaire System [1] [Brooke et al., 1996] and the attractiveness of each technique via the AttrakDiff questionnaire system [2]. After each session, participants were asked to complete the SUS and the AttrakDiff surveys.

Finally, at the end of the four sessions, the participants are asked to rank the four techniques according to their preference and give comments on each technique.

4   Results

4.1   Objective Results

To analyze our results, we rely on estimation techniques based on amplitudes effect and 95% confidence intervals, according to the recent APA recommendations. The used scripts are those used by [Willett et al., 2015] and available online [3].

4.1.1   Task completion time

Computing the average task completion time for each modality shows that participants were more efficient with the Touch modality (6968 ms, CI [6626 ; 7395]) than with the Button modality (7566 ms, CI [7274 ; 8039]). Otherwise, the average task completion time for the Swipe (9022 ms, CI [8566 ; 9488]) and the Tilt (9144 ms, CI [8743 ; 10176]) modalities are much higher compared to the other two techniques. Figure 73 shows the ratios of the three new modalities compared to the one used in the previous chapter (Button). Results reveals that only the Touch modality has a ratio lower than 1 (average = 0.921 and CI [0.896 ; 0.950]), which shows that this modality

---

1. SUS : the System Usability Scale Questionnaire
2. http ://attrakdiff.de/index-en.html
3. http ://www.aviz.fr/reliefshearing, section "study material"

is more effective than the Button modality. Otherwise, the other two modalities have ratios much higher than 1 (Swipe / Button = 1.195 with a CI [1.148, 1.243] and Tilt / Button = 1.211 with CI [1.156, 1.3]), which shows that the participants have always performed the tasks more slowly while using these two modalities compared to the Button modality initially used.
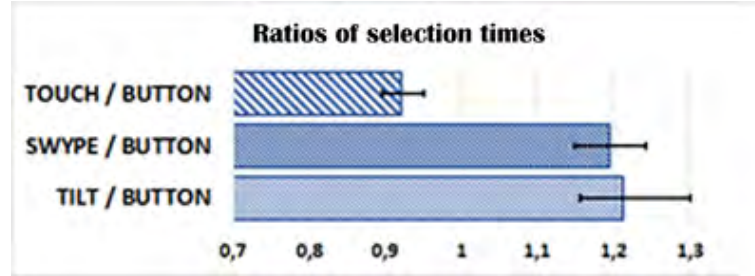
FIGURE 73 – Ratios of selection times and 95% confidence intervals. The ratios represent the time taken with different modalities against the time set with the Button modality.

This difference in selection time between different modalities is supported by the analysis of the selection time of a zone according to the number of performed actions (cf. Figure 74). Results reveals that the linear regression lines of Touch and Button techniques have slopes (respectively y = 681.81x + 4241.5 and y = 700.65x + 4799.5) lower than those of Tilt and Swipe techniques (respectively y = 1052.7x + 4872.6 and y = 1181.1x + 4313.3). A selection action with one of the "continuous" techniques is therefore more expensive in time than with a "discrete" technique.

4.1.2 Error rate and Correction

We evaluated the techniques accuracy by measuring two factors : the selection error as the number of wrong selections and the cancellation rate as the use of the back or undo event during the task. Regarding the selection errors rate shown in Figure 75, the Touch and the Button modalities prove to be the most accurate interaction modalities and produced respectively only in 0.26% (CI [0.0 ; 0.65]) and 0.13% (CI [0.0 ; 0.39]) of cases a selection error. The Swipe and Tilt interaction
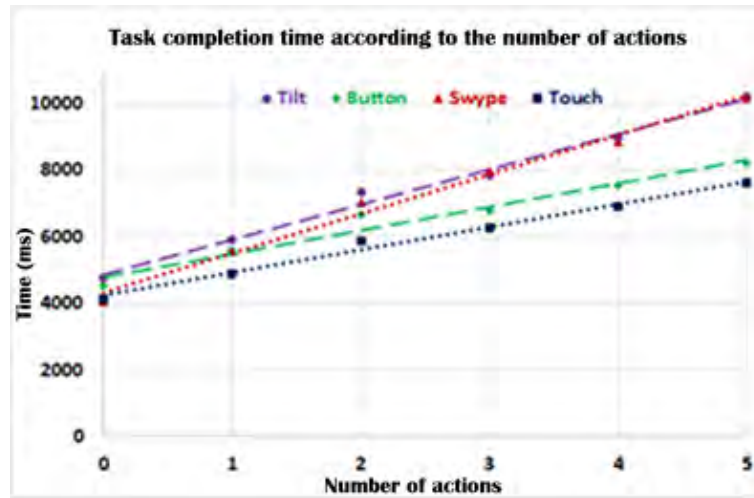
FIGURE 74 – task completion time of a zone selection by modality according to the number of realized actions

modalities have higher error rates : Swipe modalities with 0.66% (CI [0.13, 1.31]) and Tilt with 0.57% (CI [0.13, 1.44]) of cases produced a selection error. However, the error rates for each modality remain quite low (less than 1%). Otherwise, the confidence intervals for each modality overlap. Therefore, the interaction modality does not have a significant difference between the four modalities on selection error rate.
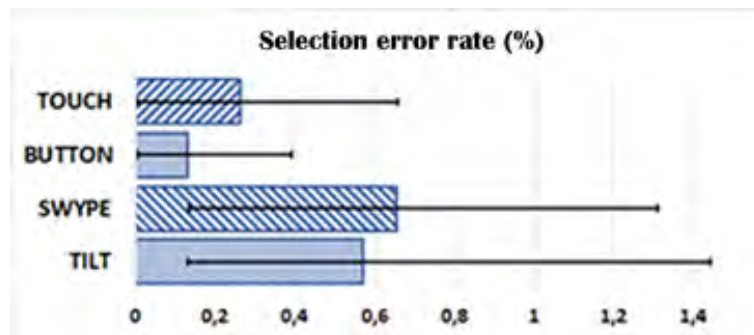
FIGURE 75 – Selection error rate and 95% confidence intervals

The accuracy of the interaction modalities is especially measured according to the cancellation rate. Regarding the Touch and the Button modalities, the back event

has relatively low usage (respectively 0.54% CI [0.21, 1.16] and 0.29% CI [0.16, 0.44]). However, it is much used in the case of Swipe modality (2.64% CI [1.12, 4.78]) and Tilt modality (2.78% CI [1.64, 4.77]). The confidence intervals for each modality (cf. Figure 76) are almost separated between Touch and Button modalities on the one hand, and Swipe and Tilt modalities on the other hand. Therefore, the use of the back or undo event differences between the first two mentioned and the last two modalities.
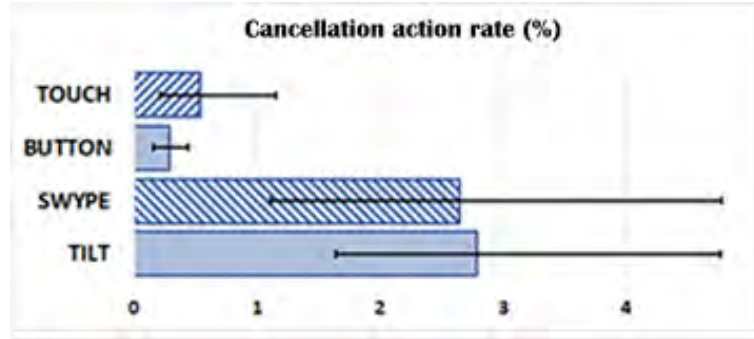
FIGURE 76 – Cancellation action rate and 95% confidence intervals

4.2 Subjective Results

Three aspects are considered for the subjective evaluation : usability, attractiveness and user preferences.

4.2.1 Usability

A SUS score [Brooke et al., 1996] was computed for each technique. The scores show better usability for the two so-called "discrete" techniques. In fact, the Button and the Touch modalities have scores of 92 and can be rated "Excellent" in terms of the qualifiers used for the results [Brooke et al., 1996]. The Swipe and Tilt modalities have lower scores (respectively 78 and 74), but these scores remain relatively high scores and rated "Good".

4.2.2 Attractiveness evaluation using AttrakDiff

To evaluate and compare the attractiveness of the four DichotoZoom interaction technique modalities, the AttrakDiff evaluation method[4] is used.

The analysis of the diagram of average values (cf. Figure 77) in terms of the technique's usability (pragmatic quality PQ) shows that Button modality (PQ = 2.35) is considered very useful and usable, then the Touch modality (PQ = 2.13), and then the Swipe modality (PQ = 1.50) and finally, the Tilt modality (PQ = 1.25). Consequently, the user is more convenient by the use of Button than the hand motion.

The analysis of the diagram of average values (cf. Figure 77) allows on the hedonic quality – identity (HQ-I) dimension to rank the four techniques based on the identification with the user from the most till the least in the following order : first the Tilt modality (HQ-I = 0.58), then the Touch modality (HQ-I = 0.50), then the Swipe modality (HQ-I = 0.08) and finally, the Button modality (HQ-I = -0.33). Consequently the use of the hand motion is more convenience to the user to transmit his professionalism, style and innovation through the application than the Button.

The hedonic quality – stimulation (HQ-S) dimension of the diagram of average values (cf. Figure 77) describes the user's stimulation and motivation. He was motivated the most by the Touch modality (HQ-S = 2.38), then by the Tilt modality (HQ-S = 1.96), then by the Swipe modality (HQ-S = 1.92) and finally, by the Button modality (HQ-S = 0.67). Consequently, the use of interesting and novel features like : the Touch, the hand motion and the Swipe is more motivating for the user than the use of the standard feature like the Button.

Considering the overall user's impression (Attractiveness ATT), the diagram of average values (cf. Figure 77) shows that Touch modality (ATT = 2.50) is considered very attractive, then the Button modality (ATT = 1.96), then the Tilt modality (ATT = 1.83) and finally, the Swipe modality (ATT = 1.50). Consequently, the user gets the good impression from the Touch because it's the most popular technique used on a flat screen. This classification is homogenous with the user's preference.

---

4. http ://attrakdiff.de/index-en.html
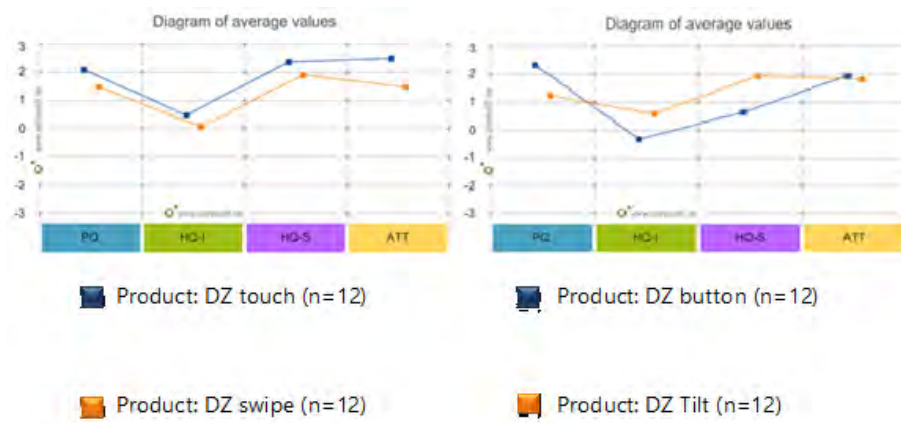
Figure 77 – diagram of average of Touch vs. Swipe vs. Button vs. Tilt

Considering the portfolio evaluation of the Button modality, Figure 78 shows that the Button modality was rated as "task oriented". The pragmatic quality (PQ = 2,35 ; Confidence = 0,18) indicates that this technique is very pragmatic and supports its users. The hedonic quality (HQ = 0,17 ; Confidence = 1,08) indicates that the user is stimulated by the Button modality but improvements for identification are required because the confidence interval is too large.

The portfolio evaluation of the Tilt modality shows in Figure 78 that the Tilt modality was rated as "rather desired". The pragmatic quality (PQ = 1,25 ; Confidence = 1,00) shows that this technique is pragmatic and supports its users. The hedonic quality (HQ = 1,27 ; Confidence = 0,64) shows that the user is stimulated by the Tilt modality. The HQ confidence interval and the PQ confidence interval are large. Improvements are required in term of stimulation and usability.

Discussing the portfolio evaluation of the Swipe modality, Figure 78 shows that the Swipe modality was rated as "rather desired". The pragmatic quality (PQ = 1,50 ; Confidence = 0,79) indicates that this technique is pragmatic and supports its users. The hedonic quality (HQ = 1,00 ; Confidence = 0, 53) indicates that the user is stimulated by the Swipe modality but improvements for identification are required. The HQ confidence interval and the PQ confidence interval are large. Improvements are required in term of stimulation and usability. The Swipe modality confidence

rectangle is smaller than the Tilt modality confidence rectangle. The Tilt modality stimulates users more than the Swipe modality but Swipe is more usable.

The portfolio evaluation of the Touch modality shows in Figure 78 that the Touch modality was rated as "desired". The pragmatic quality (PQ = 2,13; Confidence = 0,41) indicates that this technique is very pragmatic and supports its users. The hedonic quality (HQ = 1,44; Confidence = 0,38) indicates that the user is stimulated the most by the Touch modality. The Touch modality has the smaller confidence rectangle. It is the best DichotoZoom modality.
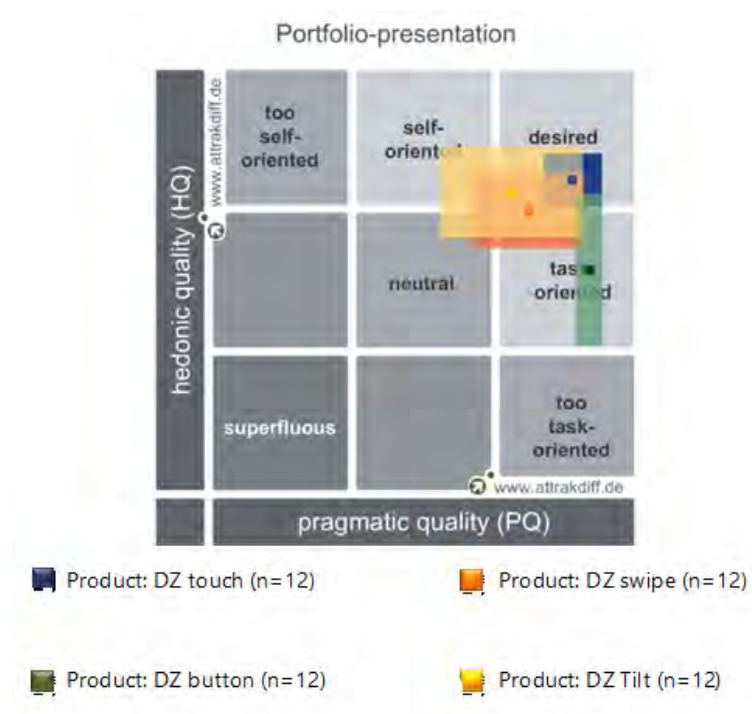
FIGURE 78 – Portfolio with average values of the dimensions PQ and HQ and the confidence rectangle of Touch vs. Swipe vs. Button vs. Tilt

4.2.3   User Preferences

At the end of the experiment, participants were asked to rank the four interaction modalities from the best to the worst according to their preference. Figure 79

presents these rankings, by showing for each technique the number of times that the participants ranked them respectively at 1st, 2nd, 3rd or 4th position.
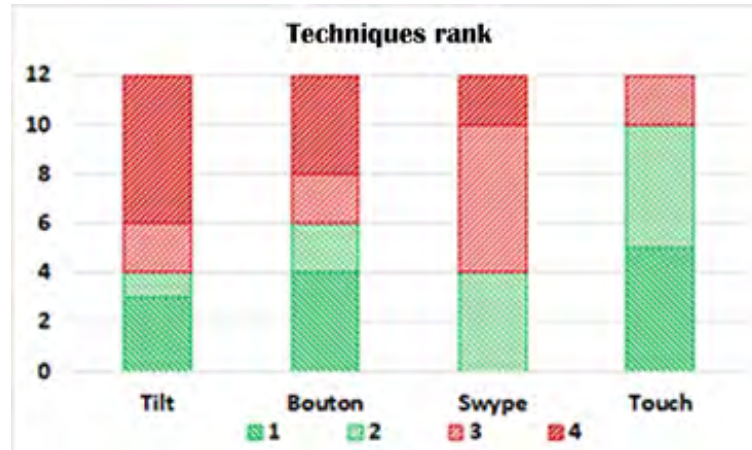
FIGURE 79 – Techniques rank : number of times each technique was positioned at 1st (dark green), 2nd (light green), 3rd (light red) and 4th (dark red) position

The participant preference for the two button-based versions is reliable with the obtained SUS scores : The Touch modality is the most favorite and appreciated after by being ranked 5 times at the first position and 5 times at the second position. The Button modality is ranked 4 times as the most appreciated technique by the participants. In contrast, the Tilt modality was ranked 6 times at the last position, while the Swipe modality was never be nominated as a preferred technique for participants.

Finally, at the end of experiments, participants were asked to comment on the different techniques. They declared their preference for the Touch and Button techniques for their simplicity of use (said respectively 10 and 6 times), and that these techniques did not require to make efforts during their use (said 2 times for each). However, the Button technique was judged "old" (3 times) and "already seen" (2 times). Although the Swipe and Tilt techniques were judged "Fun" "Modern" (4 and 7 times respectively), these techniques appeared more difficult and complicated to use (4 and 3 times) and could require a concentration effort (1 time each) and were less accurate (once each).

158

5 Discussion

Regarding the task completion time, the faster modality is the Touch (6968 ms, CI [6626 ; 7395]) then the Button (7566 ms, CI [7274 ; 8039]) then the Swipe (9022 ms, CI [8566 ; 9488]) and the slowest is Tilt (9144 ms, CI [8743 ; 10176]). Accordingly, the Touch modality is faster than other modalities and the Tilt modality is the slowest modality. Considering that, firstly, the task completion time is directly related to the response time for each action, secondly, the Tilt modality requires more time to adjust the neutral Smartphone position and the concentrate on the new 3D VE configuration, thirdly, the Touch modality is the faster interaction because it requires a simple touch at any position on a relatively large screen area therefore it does not need precision and clicking a screen is more familiar to the user, and finally the selection time difference is directly related to the time required to perform an action. Consequently, the interaction modality has a significant effect on the response time for each technique. Accordingly, the interaction modality had an influence on the task completion time and has a significant effect on completion time for each technique. The result of our experimentation confirms our first hypothesis.

Our results also validate the second hypothesis ; in fact, we evaluated the techniques accuracy by measuring two factors : the number of wrong selections and the use of the back or undo event during the task. Regarding the selection errors, the interaction modality does not have a significant effect on selection error rate. The Touch and the Button modalities prove to be the most accurate but all the four modalities are precise with low error rate (less than 1%). Since, the four modalities simply require two distinct discrete actions to perform the selection. Accordingly, the four interaction modalities are accurate and the Touch and the Button modalities are more accurate than the Swipe and Tilt modalities, because the continuous interactions (physical and digital) require a hand movement not only a simple click on physical or digital devices.

In the case of Touch and Button modalities, the back event has relatively low usage (respectively 0.54% CI [0.21, 1.16] and 0.29% CI [0.16, 0.44]) ; however, it is much used in the case of Swipe modality and Tilt modality (respectively 2.64%

CI [1.12, 4.78] and 2.78% CI [1.64, 4.77]) since the Smartphone keeps its initial position and display, therefore, the user does not need to move the Smartphone and consequently keeps concentration on the 3D VE. Consequently, the use of the back or undo event differences between the Touch and the Button modalities and the Swipe and Tilt modalities. Regarding to the rate of use of the back or undo event during the selection task : the interaction modality has a significant effect on the use of the back or undo event during the selection task for each technique.

Based on user evaluation in terms of usability, the Button and Touch techniques are rated "excellent" against the Tilt technique and the Swipe technique. Although the scores of the Tilt modality and the Swipe modality are lower, they remain relatively high scores, which mean that these two modalities have still been well received by the participants.

In addition, in terms of attractiveness, the analysis of the diagram of average values allows on the pragmatic quality (PQ) dimension and based on usefulness and usability to show that the user is more convenient by the use of Button than the hand motion. The hedonic quality – identity (HQ-I) dimension, based on the identification with the user, shows that the use of the hand motion is more convenience to the user to transmit his professionalism, style and innovation through the application than the Button. The analysis of the hedonic quality – stimulation (HQ-S) dimension, based on stimulation and motivation, reveals that the use of interesting and novel features likes : the Touch, the hand motion and the Swipe is more motivating than the use of the standard feature like the Button. Finally, the attractiveness (ATT) dimension, based on attractiveness, shows that the user gets the good impression from the Touch because it is the most popular technique used on a flat screen. This result is homogenous with the user's preference.

The average values of the AttrakDiff dimensions for Button modality shows that the mean values for the dimensions pragmatic quality (PQ = 2.35), hedonic quality – stimulation (HQ-S = 0.67), and attractiveness (ATT = 1.96) lie above the mean value of the rating scale. However, the mean value for the dimension hedonic quality – identity (HQ-I = -0.33) lies below the mean value of the rating scale. The best mean value has been attained in dimension pragmatic quality (PQ = 2.35).

This value indicates that users rate the Button modality as very usable. The lowest mean value has been achieved in the dimension hedonic quality – identity (HQ-I = -0.33). Button modality doesn't provide the user with identification and doesn't meet ordinary standards. Improvement to Button modality must be done to bind the user more strongly to the application. With regard to hedonic quality – stimulation (HQ-S = 0.67), the product is located in the average region. The mean value is optimal and Button modality meets ordinary standards. The Button's attractiveness (ATT = 1.96) value is located in the above-average region. It can be deduced that users perceive the application as very attractive. According to the architects of the questionnaire AttrakDiff, Button modality is very usable and very attractive. Button modality meets the ordinary standards in motivating and stimulating the users but needs more improvement to be identified with the user.

The average values of the AttrakDiff dimensions for Tilt modality indicates that all mean values for the dimensions lie above the mean value of the rating scale. The best mean value has been attained in dimension hedonic quality – stimulation (HQ-S = 1.96). This value indicates that users rate the Tilt modality as very motivating and stimulating. The lowest mean value has been achieved in the dimension hedonic quality – identity (HQ-I = 0.58). Tilt modality is located in the average region. Identification is provided to the user and thus Tilt meets ordinary standards. The Tilt modality's attractiveness (ATT = 1.83) value is located in the above-average region. It can be deduced that users perceive the application as very attractive. With regard to pragmatic quality (PQ = 1.25), the Tilt modality is located in the above-average region. It can be deduced that users perceive the application as very usable. According to the architects of the questionnaire AttrakDiff, Tilt modality is very motivating, very stimulating, and very attractive. Tilt is very usable and meets the ordinary standards at the identification dimension with the user.

The average values of the AttrakDiff dimensions for Touch modality illustrates that all mean values for the dimensions lie above the mean value of the rating scale. The best mean value has been attained in dimension attractiveness (ATT = 2.50). This value indicates that users rate the Touch modality as very attractive. The lowest mean value has been achieved in the dimension hedonic quality – identity (HQ-I =

161

0.50). Touch modality is located in the average region. Identification is provided to the user and thus Touch modality meets ordinary standards. With regard to pragmatic quality (PQ = 2.13), the Touch modality is located in the highest region. It can be deduced that users perceive the application as very usable. With regard to hedonic quality – stimulation (HQ-S = 2.38), the Touch modality is located in the highest region. It can be deduced that users perceive the application as very motivating and stimulating. According to the architects of the questionnaire AttrakDiff, Touch modality is very attractive, very motivating, very stimulating, and very usable. Touch modality meets the ordinary standards at the identification dimension with the user.

The average values of the AttrakDiff dimensions for Swipe modality reveals that all mean values for the dimensions lie above the mean value of the rating scale. The best mean value has been attained in dimension hedonic quality – stimulation (HQ-S = 1.92). This value indicates that users rate the Swipe modality as very motivating and very stimulating. The lowest mean value has been achieved in the dimension hedonic quality – identity (HQ-I = 0.08). Swipe modality is located in the average region. Identification is provided to the user and thus Swipe modality reached the minimum standards. With regard to pragmatic quality (PQ = 1.50), the Swipe modality is located in the above-average region. It can be deduced that users perceive the application as very usable. With regard to hedonic attractiveness (ATT = 1.50), the Swipe modality is located in the above-average region. It can be deduced that users perceive the application as very attractive. According to the architects of the questionnaire AttrakDiff, Swipe modality is very attractive, very motivating, very stimulating, and very usable. Touch modality meets the minimum standards at the identification dimension with the user.

In terms of user preference, regarding to the study of the result of the rankings and arrangements made by the user, the user preference was confirmed for the two button-based modalities : the Touch modality and the Button modality. The Touch modality is rated as the most preferred technique after being ranked five times at the first position and five times at the second position. The Button modality is also classified four times as the most preferred technique by participants. In contrast, the Tilt modality was classified six times at the last position.

162

Globally, the majority of users expressed that the touch modality is practical, simple and easy to use. The touch modality is considered funny, exciting, very simple, error free and accurate. The only negative comment is not entertaining. 8.33% of participants considered that touch modality doesn't provide amusement in manipulation. The button modality is considered simple, easy, practical and déjà vu. In addition users stated that the button modality is error free, reliable, very practical, classical and not bad. Participants mentioned that button modality doesn't need effort. The negative user's comments are old-fashion, unimaginative, and boring. 16.67% of participants considered button modality as an outdated technique. Users expressed that the Tilt modality is stylish, creative, modern, funny, interactive and novel. It is a simple, practical, good and predictable technique. 8.33% of participants considered the Tilt modality as complicated and impractical technique however the same percentage value of participants considered it as simple and practical one. But they noticed that it requires high physical and mental effort and needs some practice. The Tilt modality is more complicated and prone to error thus it is considered as difficult, more difficult and too difficult to manipulate and needs some practice. The Swipe modality is considered stylish, predictable and relatively difficult technique. In addition users stated that the Swipe modality is funny. Participants mentioned that the Swipe modality requires more concentration and attention. The negative user's comments are not fast, complicated and could damage the 3D VE. 8.33% of participants considered the Swipe modality as an impractical technique. 16.67% of participants considered the Swipe modality as a relatively difficult technique. Participants mentioned that the Swipe modality is a choice between two values consequently user's task doesn't require a movement.

Finally, the most frequently mentioned negative comments relate to the entertainment for Touch interaction ; the novelty for Button interaction ; and the difficulty for both Tilt and Swipe modality. And, the most frequently mentioned positive comments refer to a practical, simple and easy to use technique for Touch modality ; a simple technique for Button modality ; a good and stylish technique for Tilt modality ; and a stylish and predictable technique for the Swipe modality.

6 Summary

We have presented four DichotoZoom interaction modalities that improve target selection in a dense 3D VE. However, DichotoZoom interaction technique simply requires two distinct discrete actions to perform the selection. As a result, simple button modality (physical or Digital) on the Smartphone are enough to make this technique work quickly. The Touch modality is proving to be the fastest and most accurate mode. Moreover, the fact of using physical Buttons presents several advantages over other modalities1) it does not t reserve place on the screen, as it would be the case with graphic buttons; 2) does not obscure the 3D VE displayed on the screen (problem that could arise by pointing to the screen or by gesturing on the 3D VE); 3) and this keeps the smartphone always horizontal. Otherwise, the physical buttons can't be used for any type of interaction on 3D VE, it is interesting to be able to use them for this selection and thus be able to keep the other modalities for more complex tasks such as manipulation of objects in the 3D VE.

Finally, since this technique is easy to use and does not require special resources, it could be easily used on different devices that are even smaller or have few means of interaction (for example, Smartwatches or glasses).

Chapitre V

Conclusions and Future Work

The 3D object selection and manipulation remains the object of an important research field following the 3D application revolution on Smartphones. Designing an effective and easy to use 3D interaction technique allows progressing and evolving the 3D applications in a VE. In this work, we have come with two main results. In the first we have designed a novel interaction technique to select a target on a dense 3D VE displayed on a Smartphone screen. In the second we have developed a new classification framework to represent the 3D interaction techniques previously described on literature. Additionally, we have used this taxonomy as a framework for evaluating the known techniques and subsequently for designing a new technique to solve the Selection problems.

We have discussed : the selection problems (density, depth and occlusion), the selection tasks (Navigation subtasks, Pointing subtasks and Validation subtasks), the selection techniques (the Ray casting, the Curve, and the Point cursor) and the number of targets selected. We have presented a classification of large number of 3D interaction techniques defined in literature. The new classification takes into consideration the three major selection concerns : the selection problems, the selection subtasks and the number of targets selected. The following table (cf. Table 4) resumes the evaluation of our interaction technique DichotoZoom (DZ) according to the classification axis. Later, we have defined an innovative 3D VE interaction

technique : DichotoZoom as a progressive refinement selection technique based on the dichotomy algorithm which allows the selection of object inside a set of already pre-selected ones, with a limited number of actions, quickly and accurately. We have elaborated four modalities of the DichotoZoom interaction technique based on the Smartphone properties, capabilities and devices (using physical or graphical buttons, using gestural interactions via touchscreen or moving the device itself). DichotoZoom technique simply requires two distinct actions to be able to perform the selection. The user takes one of two actions, e.g. move backward of forward, to select the near or the far subsets, and can thus be easily manipulated using the different modalities.

We have assessed DichotoZoom in two experiments. The first experiment has compared the DichotoZoom technique to the existing Circulating technique and has evaluated them both subjectively and objectively. The subjective evaluation of the technique has revealed the users' attitude towards Dichotozoom in comparison to the Circulation one. Complementary, the objective study has shown that Button modality of the DichotoZoom technique provides a feasible solution for increasing the precision and speed of 3D selection tasks on small devices. Moreover, the limitations of the mobile devices could be resolved when using DichotoZoom interaction technique. Our experimental study also reveals that the DichotoZoom technique was significantly faster than the Circulation technique when the number of objects present in the pre-selected set is large enough and the environment is dense. The second experiment has evaluated the four modalities of DichotoZoom. Its subjective results have revealed that Button and Touch interaction modalities have the potential to increase the overall usability of 3D applications. Button modality has shown the best results from the users' perspective rather than gestural interactions which are more commonly preferred solutions in today's touch-based 3D applications. The Button modality has proved to be the fastest and most accurate mode and the most preferred by the users.

Our work has contributed to the general knowledge in the field of 3D interaction techniques used in a dense 3D virtual environment displayed on small screens. It has presented a first look at the task of 3D object selection with the expectation that it would inspire future work in designing object interaction techniques

for 3D virtual environments. Our work has advanced the state of the art of 3D interaction techniques and has provided guidelines to design more powerful selection techniques [Schürg, 2015, Sundén et al., 2014, Wu et al., 2016]. With the new suggested classification and the proposed technique, Selection technique in a crowded virtual environment displayed on a Smartphone can advance into a new domain that was previously impractical considering the known interaction techniques. Our work shows the importance of using simple button modality (physical or Digital) in an interaction technique. Finally, our work proves that DichotoZoom is able to yield fast and accurate results for the 3D object selection difficulties that users are likely to encounter in any virtual environment. As a result, simple buttons (physical or Digital) are proved enough to make this technique work quickly. This technique is easy to use and does not require special resources or extra devices ; it could be easily deployed on different devices even with a smaller display and with few interaction devices such as the Smartwatchs or lenses.

In future work, researches may progress the DichotoZoom study in the following fields. In our experiments we have examined the refinement selection using the dichotomy algorithm of the DichotoZoom technique after selecting a set of N pre-selected objects. Future research should also compare the effectiveness of DichotoZoom on the overall selection process. While we have focused on studying how to select a target in a crowded 3D VE, it might be convenient to study how to adapt this technique to move the target in the virtual environment after selection. We have seen that the mobile hardware devices capability has an important effect on the design of the techniques. Future research on the DichotoZoom technique should also compare the effectiveness of pointing using the built-in camera.

One key factor in this research is the task completion time. Hence the DichotoZoom is considerably faster when the pre-selected set presents numerous objects. Future enhancement will consist of designing a new Dichotozoom modality which interrupts the dichotomy algorithm when the target becomes visible, to click directly on the target and consequently to reduce the completion time.

We have seen that user displeasure results from the transition phase between two user's actions and the different ways this information is visually presented. Since it

is important for the user to better visualize the target selection area, other proposed modification consists of improving the objects visual display on the scene. Consequently, we suggest studying the different visual feedback that could be offered to users in order to minimize the effect of the user's annoyance issues and to concretize the display so that it could be used in real 3D environments.

Bibliographie

[Ahrens et al., 2008] Ahrens, J., Raake, A., Rohs, M., Schäfers, T., and Spors, S. (2008). Designing low-dimensional interaction for mobile navigation in 3d audio spaces. In *Audio Engineering Society Conference : 34th International Conference : New Trends in Audio for Mobile and Handheld Devices.*

[Angus and Sowizral, 1995] Angus, I. G. and Sowizral, H. A. (1995). Embedding the 2d interaction metaphor in a real 3d virtual environment. In *Stereoscopic Displays and Virtual Reality Systems II*, volume 2409, pages 282–293. International Society for Optics and Photonics.

[Balaa et al., 2018] Balaa, E., Raynal, M., Cabric, F., and Dubois, E. (2018). Dichotozoom : Technique de sélection d'objets dans un environnement virtuel 3d affiché sur smartphone.

[Balaa et al., 2014] Balaa, E., Raynal, M., Issa, Y. B., and Dubois, E. (2014). Classification of interaction techniques in the 3d virtual environment on mobile devices. In *International Conference on Virtual, Augmented and Mixed Reality*, pages 3–13. Springer.

[Bauer et al., 2011] Bauer, J., Thelen, S., and Ebert, A. (2011). Using smart phones for large-display interaction. In *2011 International Conference on User Science and Engineering (i-USEr)*, pages 42–47. IEEE.

[Benzina et al., 2011] Benzina, A., Toennis, M., Klinker, G., and Ashry, M. (2011). Phone-based motion control in vr : analysis of degrees of freedom. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 1519–1524. ACM.

[Bergé et al., 2015] Bergé, L.-P., Dubois, E., and Raynal, M. (2015). Design and eva-
luation of an "around the smartphone" technique for 3d manipulations on distant
display. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*,
SUI '15, pages 69–78, New York, NY, USA. ACM.

[Bergé et al., 2014a] Bergé, L.-P., Perelman, G., Hamelin, A., Raynal, M., Sanza, C.,
Houry-Panchetti, M., Cabanac, R., and Dubois, E. (2014a). Smartphone based
3d navigation techniques in an astronomical observatory context : Implementa-
tion and evaluation in a software platform. *International Journal on Advances in
Software*, 7(3 & 4) :551–566.

[Bergé et al., 2014b] Bergé, L.-P., Perelman, G., Raynal, M., Sanza, C., Serrano, M.,
Houry-Panchetti, M., Cabanac, R., and Dubois, E. (2014b). Smartphone-based 3d
navigation technique for use in a museum exhibit. In *The Seventh International
Conference on Advances in Computer-Human Interactions (ACHI 2014)*, pages
252–257. Citeseer.

[Bergé et al., 2014c] Bergé, L.-P., Serrano, M., Perelman, G., and Dubois, E. (2014c).
Exploring smartphone-based interaction with overview+ detail interfaces on 3d
public displays. In *Proceedings of the 16th international conference on Human-
computer interaction with mobile devices & services*, pages 125–134. ACM.

[Billinghurst et al., 2002] Billinghurst, M., Cheok, A., Prince, S., and Kato, H.
(2002). Real world teleconferencing. *IEEE Computer Graphics and Applications*,
22(6) :11–13.

[Boring et al., 2009] Boring, S., Jurmu, M., and Butz, A. (2009). Scroll, tilt or move
it : using mobile phones to continuously control pointers on large public displays.
In *Proceedings of the 21st Annual Conference of the Australian Computer-Human
Interaction Special Interest Group : Design : Open 24/7*, pages 161–168. ACM.

[Bowman et al., 2004] Bowman, D., Kruijff, E., LaViola Jr, J. J., and Poupyrev,
I. P. (2004). *3D User interfaces : theory and practice, CourseSmart eTextbook*.
Addison-Wesley.

[Bowman et al., 2001a] Bowman, D., Wingrave, C., Campbell, J., and Ly, V. (2001a). Using pinch gloves (tm) for both natural and abstract interaction techniques in virtual environments.

[Bowman, 2002] Bowman, D. A. (2002). Principles for the design of performance-oriented interaction techniques. In *Handbook of Virtual Environments*, pages 317–340. CRC Press.

[Bowman and Hodges, 1997] Bowman, D. A. and Hodges, L. F. (1997). An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. *SI3D*, 97 :35–38.

[Bowman and Hodges, 1999] Bowman, D. A. and Hodges, L. F. (1999). Formalizing the design, evaluation, and application of interaction techniques for immersive virtual environments. *Journal of Visual Languages & Computing*, 10(1) :37–53.

[Bowman et al., 2001b] Bowman, D. A., Johnson, D. B., and Hodges, L. F. (2001b). Testbed evaluation of virtual environment interaction techniques. *Presence : Teleoperators & Virtual Environments*, 10(1) :75–95.

[Bowman et al., 1997] Bowman, D. A., Koller, D., and Hodges, L. F. (1997). Travel in immersive virtual environments : An evaluation of viewpoint motion control techniques. In *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*, pages 45–52. IEEE.

[Bowman et al., 2001c] Bowman, D. A., Kruijff, E., LaViola Jr, J. J., and Poupyrev, I. (2001c). An introduction to 3-d user interface design. *Presence : Teleoperators & Virtual Environments*, 10(1) :96–108.

[Brooke et al., 1996] Brooke, J. et al. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194) :4–7.

[Burtnyk et al., 2002] Burtnyk, N., Khan, A., Fitzmaurice, G., Balakrishnan, R., and Kurtenbach, G. (2002). Stylecam : interactive stylized 3d navigation using integrated spatial & temporal controls. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*, pages 101–110. ACM.

171

[Çapın et al., 2006] Çapın, T., Haro, A., Setlur, V., and Wilkinson, S. (2006). Camera-based virtual environment interaction on mobile devices. In *International Symposium on Computer and Information Sciences*, pages 765–773. Springer.

[Cashion et al., 2012] Cashion, J., Wingrave, C., and LaViola Jr, J. J. (2012). Dense and dynamic 3d selection for game-based virtual environments. *IEEE transactions on visualization and computer graphics*, 18(4) :634–642.

[Chittaro and Burigat, 2004] Chittaro, L. and Burigat, S. (2004). 3d location-pointing as a navigation aid in virtual environments. In *Proceedings of the working conference on Advanced visual interfaces*, pages 267–274. ACM.

[Chittaro et al., 2003] Chittaro, L., Ranon, R., and Ieronutti, L. (2003). Guiding visitors of web3d worlds through automatically generated tours. In *Proceedings of the eighth international conference on 3D Web technology*, pages 27–38. ACM.

[Chittaro and Scagnetto, 2001] Chittaro, L. and Scagnetto, I. (2001). Is semitransparency useful for navigating virtual environments? In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 159–166. ACM.

[Daiber et al., 2012] Daiber, F., Li, L., and Krüger, A. (2012). Designing gestures for mobile 3d gaming. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, page 3. ACM.

[Dang et al., 2003] Dang, N.-T., Le, H.-H. L. H.-H., and Tavanti, M. (2003). Visualization and interaction on flight trajectory in a 3d stereoscopic environment. In *Digital Avionics Systems Conference, 2003. DASC'03. The 22nd*, volume 2, pages 9–A. IEEE.

[Darken and Sibert, 1996] Darken, R. P. and Sibert, J. L. (1996). Wayfinding strategies and behaviors in large virtual worlds. In *CHI*, volume 96, pages 142–149.

[De Haan et al., 2005] De Haan, G., Koutek, M., and Post, F. H. (2005). Intenselect : Using dynamic object rating for assisting 3d object selection. In *IPT/EGVE*, pages 201–209. Citeseer.

[de Souza et al., 2010] de Souza, M., Carvalho, D. D. B., Barth, P., Ramos, J. V., Comunello, E., and von Wangenheim, A. (2010). Using acceleration data from

smartphones to interact with 3d medical data. In *2010 23rd SIBGRAPI Conference on Graphics, Patterns and Images*, pages 339–345. IEEE.

[Decle and Hachet, 2009] Decle, F. and Hachet, M. (2009). A study of direct versus planned 3d camera manipulation on touch-based mobile phones. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, page 32. ACM.

[Delamare et al., 2013] Delamare, W., Coutrix, C., and Nigay, L. (2013). Mobile pointing task in the physical world : balancing focus and performance while disambiguating. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*, pages 89–98. ACM.

[Du et al., 2011] Du, Y., Ren, H., Pan, G., and Li, S. (2011). Tilt & touch : Mobile phone for 3d interaction. In *Proceedings of the 13th International Conference on Ubiquitous Computing*, UbiComp '11, pages 485–486, New York, NY, USA. ACM.

[Elmqvist and Fekete, 2008] Elmqvist, N. and Fekete, J.-D. (2008). Semantic pointing for object picking in complex 3d environments. In *Proceedings of graphics interface 2008*, pages 243–250. Canadian Information Processing Society.

[Feiner, 2003] Feiner, A. O. S. (2003). The flexible pointer : An interaction technique for selection in augmented and virtual reality. In *Proc. UIST'03*, pages 81–82.

[Figueroa et al., 2010] Figueroa, P., Kitamura, Y., Kuntz, S., Vanacken, L., Maesen, S., De Weyer, T., Notelaers, S., Octavia, J. R., Beznosyk, A., Coninx, K., et al. (2010). 3dui 2010 contest grand prize winners. *IEEE Computer Graphics and Applications*, 30(6) :86–96.

[Froehlich et al., 2006] Froehlich, B., Hochstrate, J., Skuk, V., and Huckauf, A. (2006). The globefish and the globemouse : two new six degree of freedom input devices for graphics applications. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 191–199. ACM.

[Garner et al., 2005] Garner, M., Zoller, E., Trotter, P., and Anderson, N. (2005). Next-generation devices : the impact of convergence.

[Gordon et al., 2002] Gordon, G., Billinghurst, M., Bell, M., Woodfill, J., Kowalik, B., Erendi, A., and Tilander, J. (2002). The use of dense stereo range data in augmented reality. In *Proceedings of the 1st international Symposium on Mixed and Augmented Reality*, page 14. IEEE Computer Society.

[Graf and Jung, 2012] Graf, H. and Jung, K. (2012). The smartphone as a 3d input device. In *2012 IEEE Second International Conference on Consumer Electronics-Berlin (ICCE-Berlin)*, pages 254–257. IEEE.

[Grossman and Balakrishnan, 2004] Grossman, T. and Balakrishnan, R. (2004). Pointing at trivariate targets in 3d environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 447–454. ACM.

[Grossman and Balakrishnan, 2005] Grossman, T. and Balakrishnan, R. (2005). The bubble cursor : enhancing target acquisition by dynamic resizing of the cursor's activation area. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 281–290. ACM.

[Grossman and Balakrishnan, 2006] Grossman, T. and Balakrishnan, R. (2006). The design and evaluation of selection techniques for 3d volumetric displays. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*, pages 3–12. ACM.

[Guiard et al., 2004] Guiard, Y., Blanch, R., and Beaudouin-Lafon, M. (2004). Object pointing : a complement to bitmap pointing in guis. In *Proceedings of Graphics Interface 2004*, pages 9–16. Canadian Human-Computer Communications Society.

[Hachet, 2010] Hachet, M. (2010). *3D User Interfaces, from Mobile Devices to Immersive Virtual Environments*. PhD thesis.

[Hachet et al., 2007] Hachet, M., Pouderoux, J., Tyndiuk, F., and Guitton, P. (2007). Jump and refine for rapid pointing on mobile phones. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 167–170. ACM.

[Hagedorn and Döllner, 2008] Hagedorn, B. and Döllner, J. (2008). Sketch-based navigation in 3d virtual environments. In *International Symposium on Smart Graphics*, pages 239–246. Springer.

174

[Hamelin and Dubois, 2015] Hamelin, A. and Dubois, E. (2015). Design and evaluation of an interaction technique for volume selection in a 3d point cloud. In *Proceedings of the 27th Conference on l'Interaction Homme-Machine*, page 3. ACM.

[Hand, 1997] Hand, C. (1997). A survey of 3d interaction techniques. In *Computer graphics forum*, volume 16, pages 269–281. Wiley Online Library.

[Hassenzahl, 2004] Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. *Human-computer interaction*, 19(4) :319–349.

[Hick, 1952] Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of experimental psychology*, 4(1) :11–26.

[Hinckley et al., 1994] Hinckley, K., Pausch, R., Goble, J. C., and Kassell, N. F. (1994). A survey of design issues in spatial input. In *Proceedings of the 7th annual ACM symposium on User interface software and technology*, pages 213–222. ACM.

[Horvitz et al., 2012] Horvitz, E. J., Apacible, J., Sarin, R., and Liao, L. (2012). Prediction, expectation, and surprise : Methods, designs, and study of a deployed traffic forecasting service. *arXiv preprint arXiv :1207.1352.*

[Hwang and Jeon, 2009] Hwang, M. S. and Jeon, J. W. (2009). Design of the 3d input method based on touch device for mobile. In *2009 Fifth International Joint Conference on INC, IMS and IDC*, pages 1607–1610. IEEE.

[Hyman, 1953] Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of experimental psychology*, 45(3) :188.

[Hynninen, 2012] Hynninen, T. (2012). First-person shooter controls on touchscreen devices : A heuristic evaluation of three games on the ipod touch. Master's thesis.

[Igarashi et al., 1998] Igarashi, T., Kadobayashi, R., Mase, K., and Tanaka, H. (1998). Path drawing for 3 d walkthrough. In *ACM Symposium on User Interface Software and Technology*, pages 173–174.

[Jankowski and Hachet, 2013] Jankowski, J. and Hachet, M. (2013). A survey of interaction techniques for interactive 3d environments. In *Eurographics 2013 - STAR, May 2013, Girona, Spain*, pages 65–93.

[Katzakis and Hori, 2010] Katzakis, N. and Hori, M. (2010). Mobile devices as multi-dof controllers. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 139–140. IEEE.

[Khan et al., 2008] Khan, A., Mordatch, I., Fitzmaurice, G., Matejka, J., and Kurtenbach, G. (2008). Viewcube : a 3d orientation indicator and controller. In *Proceedings of the 2008 symposium on Interactive 3D graphics and games*, pages 17–25. ACM.

[Kim et al., 2008] Kim, J.-S., Gračanin, D., Matković, K., and Quek, F. (2008). Finger walking in place (fwip) : A traveling technique in virtual environments. In *International Symposium on Smart Graphics*, pages 58–69. Springer.

[Kopper et al., 2011] Kopper, R., Bacim, F., and Bowman, D. A. (2011). Rapid and accurate 3d selection by progressive refinement. In *2011 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 67–74. IEEE.

[Kurtenbach and Buxton, 1993] Kurtenbach, G. and Buxton, W. (1993). The limits of expert performance using hierarchic marking menus. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pages 482–487. ACM.

[Kurtenbach and Buxton, 1994] Kurtenbach, G. and Buxton, W. (1994). User learning and performance with marking menus. In *CHI*, volume 94, pages 258–264. Citeseer.

[Lao et al., 2009] Lao, S., Heng, X., Zhang, G., Ling, Y., and Wang, P. (2009). A gestural interaction design model for multi-touch displays. In *Proceedings of the 23rd British HCI Group annual conference on people and computers : Celebrating people and technology*, pages 440–446. British Computer Society.

[Lehtinen et al., 2012] Lehtinen, V., Nurminen, A., and Oulasvirta, A. (2012). Integrating spatial sensing to an interactive mobile 3d map. In *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 11–14. IEEE.

[Lewis and Sauro, 2009] Lewis, J. R. and Sauro, J. (2009). The factor structure of the system usability scale. In *International conference on human centered design*, pages 94–103. Springer.

[Liang and Green, 1993] Liang, J. and Green, M. (1993). Geometric modeling using six degrees of freedom input devices. In *3rd Int'l Conference on CAD and Computer Graphics*, pages 217–222. Citeseer.

[Liang and Green, 1994] Liang, J. and Green, M. (1994). Jdcad : A highly interactive 3d modeling system. *Computers & graphics*, 18(4) :499–506.

[Lucas, 2005] Lucas, J. F. (2005). *Design and evaluation of 3D multiple object selection techniques.* PhD thesis, Virginia Tech.

[Mackinlay et al., 1990] Mackinlay, J. D., Card, S. K., and Robertson, G. G. (1990). Rapid controlled movement through a virtual 3d workspace. In *ACM SIGGRAPH computer graphics*, volume 24, pages 171–176. ACM.

[Marzo et al., 2014] Marzo, A., Bossavit, B., and Hachet, M. (2014). Combining multi-touch input and device movement for 3d manipulations in mobile augmented reality environments. In *Proceedings of the 2Nd ACM Symposium on Spatial User Interaction*, SUI '14, pages 13–16, New York, NY, USA. ACM.

[Medeiros et al., 2013] Medeiros, D., Teixeira, L., Carvalho, F., Santos, I., and Raposo, A. (2013). A tablet-based 3d interaction tool for virtual engineering environments. In *Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, pages 211–218. ACM.

[Mossel et al., 2013a] Mossel, A., Venditti, B., and Kaufmann, H. (2013a). 3dtouch and homer-s : intuitive manipulation techniques for one-handed handheld augmented reality. In *Proceedings of the Virtual Reality International Conference : Laval Virtual*, page 12. ACM.

[Mossel et al., 2013b] Mossel, A., Venditti, B., and Kaufmann, H. (2013b). Drill-sample : precise selection in dense handheld augmented reality environments. In *Proceedings of the Virtual Reality International Conference : Laval Virtual*, page 10. ACM.

[Naito et al., 2009] Naito, M., Shizuki, B., Tanaka, J., and Hosobe, H. (2009). Interaction techniques using a spherical cursor for 3d targets acquisition and indicating in volumetric displays. In *2009 13th International Conference Information Visualisation*, pages 607–612. IEEE.

177

[Nieuwenhuisen and Overmars,] Nieuwenhuisen, D. and Overmars, M. Motion planning for camera movements in virtual environments 2002. *Utrecht, the Netherlands : Utrecht University Information and Computing Sciences.*

[Olwal and Feiner, 2004] Olwal, A. and Feiner, S. (2004). Unit : modular development of distributed interaction techniques for highly interactive user interfaces. In *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, pages 131–138. ACM.

[Pausch et al., 1995] Pausch, R. F., Burnette, T., Brockway, D., and Weiblen, M. E. (1995). Navigation and locomotion in virtual worlds via flight into hand-held miniatures. In *Siggraph*, volume 95, pages 399–400.

[Perelman et al., 2015] Perelman, G., Serrano, M., Raynal, M., Picard, C., Derras, M., and Dubois, E. (2015). The roly-poly mouse : Designing a rolling input device unifying 2d and 3d interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 327–336. ACM.

[Phillips and Badler, 1988] Phillips, C. B. and Badler, N. I. (1988). Jack : A toolkit for manipulating articulated figures. In *Proceedings of the 1st annual ACM SIGGRAPH symposium on User Interface Software*, pages 221–229. ACM.

[Piazza et al., 2013] Piazza, T., Fjeld, M., Ramos, G., Yantac, A., and Zhao, S. (2013). Holy smartphones and tablets, batman ! : mobile interaction's dynamic duo. In *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction*, pages 63–72. ACM.

[Pierce and Pausch, 2002] Pierce, J. S. and Pausch, R. (2002). Comparing voodoo dolls and homer : exploring the importance of feedback in virtual environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 105–112. ACM.

[Pietroszek et al., 2015] Pietroszek, K., Wallace, J. R., and Lank, E. (2015). Tiltcasting : 3d interaction on large displays using a mobile device. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15, pages 57–62, New York, NY, USA. ACM.

[Potter et al., 1988] Potter, R. L., Weldon, L. J., and Shneiderman, B. (1988). Improving the accuracy of touch screens : an experimental evaluation of three strategies. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 27–32. ACM.

[Poupyrev et al., 1996] Poupyrev, I., Billinghurst, M., Weghorst, S., and Ichikawa, T. (1996). The go-go interaction technique : non-linear mapping for direct manipulation in vr. In *ACM Symposium on User Interface Software and Technology*, pages 79–80. Citeseer.

[Prachyabrued et al., 2011] Prachyabrued, M., Ducrest, D. L., and Borst, C. W. (2011). Handymap : a selection interface for cluttered vr environments using a tracked hand-held touch device. In *International Symposium on Visual Computing*, pages 45–54. Springer.

[Ramos et al., 2006] Ramos, G., Robertson, G., Czerwinski, M., Czerwinski, M., Tan, D., Baudisch, P., Hinckley, K., Hinckley, K., and Agrawala, M. (2006). Tumble ! splat ! helping users access and manipulate occluded content in 2d drawings. In *Proceedings of the working conference on Advanced visual interfaces*, pages 428–435. ACM.

[Raynal et al., 2013] Raynal, M., Dubois, E., and Schmitt, B. (2013). Towards unification for pointing task evaluation in 3d desktop virtual environment. In *International Conference on Human Factors in Computing and Informatics*, pages 562–580. Springer.

[Rekimoto, 1996] Rekimoto, J. (1996). Tilting operations for small screen interfaces (tech note). In *UIST*, volume 96, pages 167–168.

[Riboldi and Maciel, 2010] Riboldi, G. and Maciel, A. (2010). Permeating the architectural past in 4d : an augmented reality interactive system. In *XIIth Symposium on Virtual and Augmented Reality. Brazil : Brazilian Computer Society*, pages 133–140.

[Robbins et al., 2004] Robbins, D. C., Cutrell, E., Sarin, R., and Horvitz, E. (2004). Zonezoom : Map navigation for smartphones with recursive view segmentation. In

Proceedings of the Working Conference on Advanced Visual Interfaces, AVI '04, pages 231–234, New York, NY, USA. ACM.

[Rodríguez and León, 2015] Rodríguez, A. and León, A. (2015). Smartphone-based remote 3d interaction for digital heritage applications. In *2015 Digital Heritage*, volume 1, pages 297–300. IEEE.

[Rohs et al., 2007] Rohs, M., Schöning, J., Raubal, M., Essl, G., and Krüger, A. (2007). Map navigation with mobile devices : virtual versus physical movement with and without visual context. In *Proceedings of the 9th international conference on Multimodal interfaces*, pages 146–153. ACM.

[Rosa and Nagel, 2010] Rosa, D. A. W. and Nagel, H. H. (2010). Selection techniques for dense and occluded virtual 3d environments, supported by depth feedback : Double, bound and depth bubble cursors. In *2010 XXIX International Conference of the Chilean Computer Science Society*, pages 218–225. IEEE.

[Roudaut et al., 2008] Roudaut, A., Huot, S., and Lecolinet, E. (2008). Taptap and magstick : improving one-handed target acquisition on small touch-screens. In *Proceedings of the working conference on Advanced visual interfaces*, pages 146–153. ACM.

[Ruddle et al., 2000] Ruddle, R. A., Howes, A., Payne, S. J., and Jones, D. M. (2000). The effects of hyperlinks on navigation in virtual environments. *International Journal of Human-Computer Studies*, 53(4) :551–581.

[Salomon et al., 2003] Salomon, B., Garber, M., Lin, M. C., and Manocha, D. (2003). Interactive navigation in complex environments using path planning. In *Proceedings of the 2003 symposium on Interactive 3D graphics*, pages 41–50. ACM.

[Schmidt et al., 2006] Schmidt, R., Wyvill, B., Sousa, M. C., and Jorge, J. A. (2006). Shapeshop : Sketch-based solid modeling with blobtrees. In *ACM SIGGRAPH 2006 Courses*, page 14. ACM.

[Schmitt et al., 2011] Schmitt, B., Dubois, E., Raynal, M., Bach, C., and Croenne, D. (2011). Evaluation d'une technique d'interaction ubiquitaire pour le pointage de données complexes et spatialisées. *Mobilité et Ubiquité 2011 UBIMOB'11*.

[Schmitt et al., 2012] Schmitt, B., Raynal, M., and Dubois, E. (2012). Poster : A composite approach to evaluate two interaction techniques for a 3d pointing task. In *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 159–160. IEEE.

[Schürg, 2015] Schürg, T. (2015). Development and evaluation of interaction concepts for mobile augmented and virtual reality applications considering external controllers. *Rheinisch-Westfälische Technische Hochschule Aachen, Aachen, Germany, Tech. Rep. 52062*.

[Scoditti et al., 2011] Scoditti, A., Vincent, T., Coutaz, J., Blanch, R., and Mandran, N. (2011). Touchover : Decoupling positioning from selection on touch-based handheld devices. In *Proceedings of the 23rd Conference on L'Interaction Homme-Machine*, IHM '11, pages 6 :1–6 :4, New York, NY, USA. ACM.

[Sellen et al., 1992] Sellen, A. J., Kurtenbach, G. P., and Buxton, W. A. (1992). The prevention of mode errors through sensory feedback. *Human-computer interaction*, 7(2) :141–164.

[Serino et al., 2016] Serino, M., Cordrey, K., McLaughlin, L., and Milanaik, R. L. (2016). Pokémon go and augmented virtual reality games : a cautionary commentary for parents and pediatricians. *Current opinion in pediatrics*, 28(5) :673–677.

[Song and Norman, 1993] Song, D. and Norman, M. (1993). Nonlinear interactive motion control techniques for virtual space navigation. In *Proceedings of IEEE Virtual Reality Annual International Symposium*, pages 111–117. IEEE.

[Steed, 2006] Steed, A. (2006). Towards a general model for selection in virtual environments. In *3D User Interfaces (3DUI'06)*, pages 103–110. IEEE.

[Steed and Parker, 2004] Steed, A. and Parker, C. (2004). 3d selection strategies for head tracked and non-head tracked operation of spatially immersive displays. In *8th International Immersive Projection Technology Workshop*, pages 13–14.

[Steinicke et al., 2006] Steinicke, F., Ropinski, T., and Hinrichs, K. (2006). Object selection in virtual environments using an improved virtual pointer metaphor. In *Computer Vision and Graphics*, pages 320–326. Springer.

181

[Stoakley et al., 1995] Stoakley, R., Conway, M. J., and Pausch, R. (1995). Virtual reality on a wim : interactive worlds in miniature. In *CHI*, volume 95, pages 265–272.

[Stuerzlinger and Wingrave, 2011] Stuerzlinger, W. and Wingrave, C. A. (2011). The value of constraints for 3d user interfaces. In *Virtual Realities*, pages 203–223. Springer.

[Stuerzlinger et al., 2006] Stuerzlinger, W., Zaman, L., Pavlovych, A., and Oh, J.-Y. (2006). The design and realization of covid : a system for collaborative virtual 3d design. *Virtual Reality*, 10(2) :135–147.

[Sundén et al., 2014] Sundén, E., Bock, A., Jönsson, D., Ynnerman, A., and Ropinski, T. (2014). Interaction techniques as a communication channel when presenting 3d visualizations. In *2014 IEEE VIS International Workshop on 3DVis (3DVis)*, pages 61–65. IEEE.

[Teather and Stuerzlinger, 2008] Teather, R. J. and Stuerzlinger, W. (2008). Assessing the effects of orientation and device on (constrained) 3d movement techniques. In *2008 IEEE symposium on 3D user interfaces*, pages 43–50. IEEE.

[Teather and Stuerzlinger, 2013] Teather, R. J. and Stuerzlinger, W. (2013). Pointing at 3d target projections with one-eyed and stereo cursors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 159–168. ACM.

[Telkenaroglu and Capin, 2013] Telkenaroglu, C. and Capin, T. (2013). Dual-finger 3d interaction techniques for mobile devices. *Personal and ubiquitous computing*, 17(7) :1551–1572.

[van Ballegooij and Eliëns, 2001] van Ballegooij, A. and Eliëns, A. (2001). Navigation by query in virtual worlds. In *Virtual Reality Modeling Language Symposium : Proceedings of the sixth international conference on 3 D Web technology*, volume 2001, pages 77–83. Citeseer.

[Van Dam, 1997] Van Dam, A. (1997). Post-wimp user interfaces. *Communications of the ACM*, 40(2) :63–67.

[Vanacken et al., 2007] Vanacken, L., Grossman, T., and Coninx, K. (2007). Exploring the effects of environment density and target visibility on object selection in 3d virtual environments. In *2007 IEEE symposium on 3D user interfaces*. IEEE.

[Vanacken et al., 2009] Vanacken, L., Grossman, T., and Coninx, K. (2009). Multimodal selection techniques for dense and occluded 3d virtual environments. *International Journal of Human-Computer Studies*, 67(3) :237–255.

[Vemavarapu and Borst, 2013] Vemavarapu, P. V. and Borst, C. W. (2013). Evaluation of a handheld touch device as an alternative to standard ray-based selection in a geosciences visualization environment. In *Workshop on Off-the-Shelf Virtual Reality*.

[Vinson, 1999] Vinson, N. G. (1999). Design guidelines for landmarks to support navigation in virtual environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 278–285. ACM.

[Vogel and Baudisch, 2007] Vogel, D. and Baudisch, P. (2007). Shift : a technique for operating pen-based interfaces using touch. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 657–666. ACM.

[Watts et al., 2010] Watts, C., Sharlin, E., and Woytiuk, P. (2010). Exploring interpersonal touch-based interaction and player socialization in prism squad : Go. In *3rd Annual International Conference on Computer Games, Multimedia and Allied Technology, Singapore*.

[Willett et al., 2015] Willett, W., Jenny, B., Isenberg, T., and Dragicevic, P. (2015). Lightweight relief shearing for enhanced terrain perception on interactive maps. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3563–3572. ACM.

[Wobbrock et al., 2008] Wobbrock, J. O., Myers, B. A., and Aung, H. H. (2008). The performance of hand postures in front-and back-of-device interaction for mobile computing. *International Journal of Human-Computer Studies*, 66(12) :857–875.

[Wonner et al., 2012] Wonner, J., Grosjean, J., Capobianco, A., and Bechmann, D. (2012). Starfish : a selection technique for dense virtual environments. In *Pro-

ceedings of the 18th ACM symposium on Virtual reality software and technology, pages 101–104. ACM.

[Wu et al., 2015] Wu, S., Chellali, A., Otmane, S., and Moreau, G. (2015). Layers-troke : a layer based selector for small target acquisition. In *Proceedings of the 27th Conference on l'Interaction Homme-Machine*, page 38. ACM.

[Wu et al., 2016] Wu, S., Guo, X., Lu, S., Xiao, B., and Zhao, C. (2016). Semantic-based urban modern agricultural park virtual interactive design system. *International Journal of Digital Content Technology and its Applications*, 10(1) :89.

[Wyss et al., 2006] Wyss, H. P., Blach, R., and Bues, M. (2006). isith-intersection-based spatial interaction for two hands. In *3D User Interfaces (3DUI'06)*, pages 59–61. IEEE.

[Yatani et al., 2008] Yatani, K., Partridge, K., Bern, M., and Newman, M. W. (2008). Escape : a target selection technique using visually-cued gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 285–294. ACM.

[Yu et al., 2012] Yu, L., Efstathiou, K., Isenberg, P., and Isenberg, T. (2012). Efficient structure-aware selection techniques for 3d point cloud visualizations with 2dof input. *IEEE Transactions on Visualization and Computer Graphics*, 18(12) :2245–2254.

[Yu et al., 2015] Yu, L., Efstathiou, K., Isenberg, P., and Isenberg, T. (2015). Cast : Effective and efficient user interaction for context-aware selection in 3d particle clouds. *IEEE transactions on visualization and computer graphics*, 22(1) :886–895.

[Zaman et al., 2010] Zaman, L., Natapov, D., and Teather, R. J. (2010). Touchs-creens vs. traditional controllers in handheld gaming. In *Proceedings of the international academic conference on the future of game design and technology*, pages 183–190. ACM.

[Zeleznik et al., 2007] Zeleznik, R. C., Herndon, K. P., and Hughes, J. F. (2007). Sketch : An interface for sketching 3d scenes. In *ACM SIGGRAPH 2007 courses*, page 19. ACM.

[Zhai et al., 1994] Zhai, S., Buxton, W., and Milgram, P. (1994). The" silk cursor" : Investigating transparency for 3d target acquisition. In *CHI*, volume 94, pages 459–464. Citeseer.

[Zhai et al., 1997] Zhai, S., Milgram, P., and Rastogi, A. (1997). Anisotropic human performance in six degree-of-freedom tracking : An evaluation of three-dimensional display and control interfaces. *IEEE Transactions on Systems, Man, and Cybernetics-Part A : Systems and Humans*, 27(4) :518–528.