# 3D on Tabletops

**Ursula Marciniak**
RWTH Aachen University
52056 Aachen, Germany
ursula.marciniak@rwth-aachen.de

**Thomas Kutz**
RWTH Aachen University
52056 Aachen, Germany
thomas.kutz@rwth-aachen.de

## ABSTRACT
Current tabletop systems are increasingly dealing with 3D applications. Users using these tabletops have to interact with 3D models to accomplish their tasks. These tabletop systems have to provide interaction techniques which allow effective and accurate manipulations of 3D data. Usually, tabletops only provide a 2D input space, namely the table surface, which is why 3D interactions may become unnatural. Thus, an extension of the input space into the third dimension may be helpful. However, the aspects to be considered when developing interaction concepts depend on the used visualization of the 3D models on the tabletop. There are several technologies to visualize 3D data. It can be displayed on the 2D surface of the table or the 3D impression can be improved by using stereoscopic or virtual reality systems. There are also tabletops using physical objects, like blocks or clay, for the representation of 3D data. For each technology different aspects concerning the interaction with the 3D models have to be taken into account. This report gives an overview of current 3D visualization technologies for tabletops. The main problems concerning their usability are examined and solutions from the ongoing research in this area are presented.

### Author Keywords
Tabletop, tangible user interface, multi-touch interface, 3D visualization, interaction techniques

### ACM Classification Keywords
H.5.2 Interaction styles: []

## INTRODUCTION
In this report, we consider the concept of 3D on tabletops. It will be shown how 3D informations can be visualized, which technologies are available, and how the usability can be improved. A tabletop itself is a digital table where 2D or 3D information is displayed. Furthermore, the users have the possibility to interact with the displayed information to manipulate the data. Later on we will see some examples of different interaction types, methods and their fields of application.

**Figure 1. 2-dimensional (left) and 3-dimensional (right) game pieces to evaluate the spatial seeing ability of humans, where users are supposed to keep the pieces in mind [11].**

To introduce and motivate the topic of 3D on tabletops we will now name some advantages of 3D over 2D. A user study on the accuracy of perception of graphical elements on tabletops from different view angles shows some interesting results [16]. Basically, there are two types of view distortion. The first one occurs when multiple users are sitting around one table. Then each user sees the content from a different orientation. The second one occurs when large variances in the user's viewing angle appear. Accordingly, the seated user may see 2D information distorted. This depends on the view angles and on the location of the information on the display. The accuracy of perception decreases the more the tabletop is tilted. Some graphical elements are more robust against these distortion problems than others. However, this problem reveals the benefits of some 3D visualization technologies which overcome the problem of view distortion.

A further user study evaluating the spatial seeing ability of humans, points out that this ability empowers the user to store information more effectively in memory. The spatial abilities help to store more information about particular objects including their position in space and enables the user to work more accurately [11]. The user study used 2D and 3D memory game pieces which are visualized in Figure 1. The results show that even though completing the 3D memory game took longer, the user did so in fewer attempts. For this reason this study can be seen as a motivation for extending the tabletop into the third dimension.

However, there are also some problems with dealing with 3D data on tabletops. The main challenges are on the one hand the visualization and on the other hand the usability. Some questions regarding the visualization may be: Which visualization methods provide the user with the best 3D perception? Where do the advantages outweigh the resulting problems? The second issue regarding the interaction methods brings up questions like: How can the user interact with the 3D data? How can it become realistic and intuitive?

The named problems will be discussed in the following chapters. At first we consider the part of the visualization followed by the challenge of the usability of 3D tabletops. The usability chapter contains the question how the physical world can be taken as an example for the digital world. We will see what the differences between using physical objects contra virtual objects are. Thus, we can discuss what the problems of dealing with 3D on tabletops are and we will also see that the usability problems depend on the used technology for the visualization. Afterwards we offer a short summary and discuss the problems and advantages of the presented technologies.

## VISUALIZATION OF 3D DATA

Several technologies for visualization of 3D data have been developed. This chapter structures the existing technologies as done in [5]. Depending on the used technology, different problems concerning the interaction with the visualized data occur.

### 3D on 2D displays

One way to visualize 3D data on a tabletop is to project the 3D objects on a 2D display. Thus, the resulting imagery is only 2D and can be manipulated by common multi-touch displays in a direct way, as proposed by Hancock and Cockburn [7] in their *shallow-depth 3D* concept. One problem of the visualization of 3D objects on 2D displays is the perspective distortion when viewing the table from different positions. Especially when the system is to be used by more than one user simultaneously, perspective distortion is a huge disadvantage. Furthermore, some tasks like picking up an object and putting it on another one, are not natural when using only the 2D table surface. Therefore, an extension of the input space into the third dimension is useful to provide more natural interaction techniques. Solutions for this problem will be presented in the chapter *Usability of 3D tabletops*.

### Stereoscopic technologies

Stereoscopic technologies provide each eye of the user with a slightly different image which results in a depth perception. The visualized 3D data seems to pop out of the display. Usually, the user has to wear glasses for achieving this effect. There are several kinds of this technology, like the shutter glasses or polarization methods. However, the viewing angle and position for which a usable visualization can be achieved, is limited. Moreover, looking around an object is not possible when using only stereoscopy. For such a 360-degree view a motion tracking system is required which detects the position and viewing angle of the user's head and adapts the visualization accordingly. The problem with this visualization is that every user has to be provided with another image in order to avoid high view distortions which can be hard when only using one display.

The system described in [12] is an example for a tabletop using stereoscopic technologies. The system called *Illusion-Hole* combines 2D and 3D data on one display. 3D data is visualized via stereoscopic images and can be perceived by the users wearing polarized glasses with a head-tracking system. Thereby, the users look through a hole and perceive different images depending on their position around the table. With this concept a 360-degree view is realized for up to four users. Figure 2 shows the system in use.

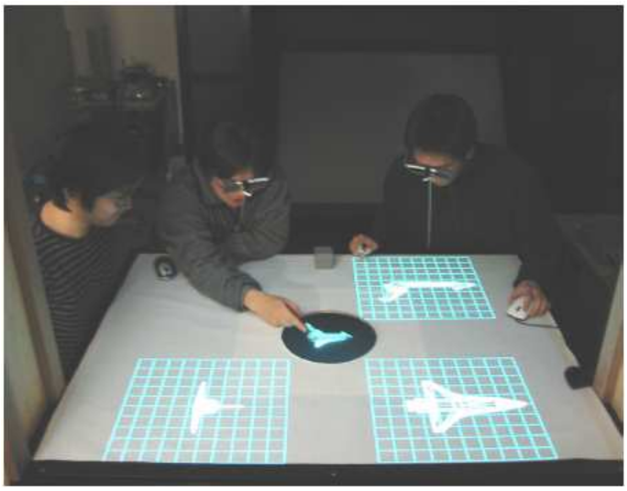

**Figure 2. The IllusionHole system in use. The object in the center is perceived in 3D through the polarized glasses [12].**

### Augmented and virtual reality

Another possibility to let the user perceive 3D data is to use head mounted displays, as it is often the case in virtual reality systems. Wearing such a display, the user feels completely immersed in a virtual environment. Since in tabletop environments the users want to interact with each other, head mounted displays that only augment reality are better suited and thus more often used. Such systems augment reality, for example by displaying additional 3D information. Since every user has its own display, there are no problems concerning the perspective distortion of 3D objects.

An example for a system using such a head mounted augmented reality display is *VITA* [2] which serves for the visualization of an archaeological dig. In VITA the users wear tracked head-mounted displays to augment the 2D data with 3D information. There are two modes in which the system can be used. In the *world-in-miniature* mode the dig environment is visualized on a tabletop next to a touch-sensitive surface where a 2D user interface is placed. On this interface the user selects interesting data which is then visualized in 3D on the tabletop. This mode is shown in Figure 3. The second mode is called *life-size-world* and displays the whole environment in the head-mounted display at its actual size. The user can walk through the environment and examine archaeological finds. There is also the possibility to display the surrounding panorama which results in a completely immersive virtual reality.

**Figure 3. World-in-miniature mode in VITA. An overview of the dig site is visualized and can be manipulated by using the tabletop's touch surface [2].**

**Figure 4. Illuminating Clay. The users are forming a landscape whereas the system projects additional information in real-time onto the clay [13].**

### Spatially augmented reality

In spatially augmented reality systems the third dimension is directly created by using physical proxies, such as blocks or clay. On this objects additional information is projected. Since here the third dimension is not virtual but physical, no glasses or head mounted displays are required. This technology provides new interaction concepts, since the data can be manipulated directly by manipulating the underlying physical proxies. Such interfaces are called *tangible user interfaces*. Tabletop systems using this technology are aware of the form and position of the physical proxies and adapt the projected imagery accordingly. However, the usability of such a system is constrained to the form of the tangibles. Therefore, such systems are mostly constructed for special purposes and cannot be used for different application domains.

An example of such a system is *Illuminating Clay* [13]. Using this tabletop the user can analyze landscape models by deforming clay. A laser in the ceiling measures the topography of the formed landscape which is used as input data for several analysis functions. The results are then projected on the clay giving the user further information, as shown in Figure 4.

## USABILITY OF 3D TABLETOPS

This chapter is focusing on the usability of 3D tabletops. At first the topic of dealing with 3D data on 2D tabletops will be discussed. For that purpose two different approaches, which solve two main problems in that area, are presented. Afterwards a short motivation for the use of the 3rd dimension follows. A study on the comparison of interactions with digital and physical objects will show that the physical world is still the best resource for developing interaction concepts. Based on that study some approaches will be presented for the extension of the input space. Further challenges are on the one hand the connection of 2D and 3D contexts in augmented reality systems and on the other hand the possibility of sensing objects in higher levels when using tangible user interfaces. These two topics will be discussed at the end of this chapter.

### 3D data on a 2D tabletop display

There are two main problems when dealing with 3D data on a 2D display. The first problem is view distortion and the second one is providing intuitive techniques for interacting with 3D objects on a 2D display.

#### *The effects of changing projection*

The problem of view distortion was addressed by a user study, where 3D data models were displayed on a 2D tabletop display [7]. The authors discuss the problem of distortion when the viewing angle changes and they present ideas to leverage this problem. The focus of their study was on how the degree of discrepancy between the center of projection and the observer's point of view affects perception of object orientation. There are different ways of projecting 3D data which can lead to distorted images and which makes interpreting angles and orientations very difficult. Errors increase when displacing the center of the projection from the observer's viewpoint which designers have to take into consideration. For multi-user tables, they found out that a neutral center of projection, which is adjusted for all participating users, combined with parallel projection geometry provides a compromise for multi-user situations.

The experimental task of the twenty-four participants was to determine the orientation of target objects. Figure 5 shows a diagram of the experimental setup, where the participants stood at the ends of a bottom-projected tabletop display. For this purpose the viewpoints of the users were tracked with a Vicon motion tracking system. Additional markers were attached at the end of a string at the corners of the tabletop and an installed wand recorded the answers of the participants about the angle to the target object. In the experimental task two 3D target objects (long thin cylinder inside a shorter thicker cylinder) were displayed at two different distances for each user, as shown in Figure 5.

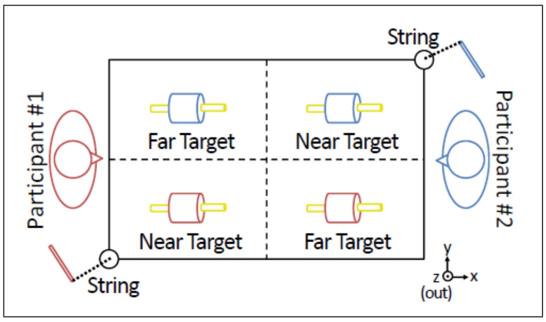


**Figure 5. The setup of the experimental task of determining the orientations of target objects [7].**

The authors provide several practical recommendations that can be emerged from their study. At first they have shown that decisions made about the projection geometry are important. Especially, attaching the center of projection to one person's point of view can introduce errors of up to 60 degree in perception for another person at the table, and using a center of projection above the table together with a perspective projection can introduce errors over 40 degree.

Their study provides some evidence that a parallel projection with a center of projection directly above the table eases some of the problems. Furthermore, a center of projection that is very far above the table may win some of the benefits of a parallel projection, while maintaining some of the perspective depth cues.

### *Shallow Depth 3D*

The next challenge is the question how to interact with 3D data on 2D interaction surfaces. This topic was discussed 2007 by Hancock et al., who presented the concept of *shallow depth 3D* [6]. The authors recognized that the use of the third dimension to pile, sort and store objects does not extend far above the table. So *shallow depth 3D* is an approach for 3D interaction with limited depth. They present a user study that examines the efficiency of different interaction possibilities. Basically there are three possible interactions which will be described in the following:

*One-Touch Input:* The user has to touch the object with the index finger of his dominant hand at one point. The position on the object holds until the finger will be released. In this way it is possible to rotate the object in x, y, and z-axis and an output degree of freedom of 5 (5DOF) can be achieved. An illustration of the motions are shown in Figure 6. Nevertheless, rotating a side of a virtual cube to the surface involves touching that side and dragging, which may require a re-touch.
*Translation:* To translate the object, the user has to touch it at its center and move it in one direction.
*Rotation on plane:* To perform a rotation on plane, the user has to touch the object next to its center and move it.
*Rotation through entire space:* For rotations through the entire space, the user has to touch anywhere else on the object and move it.

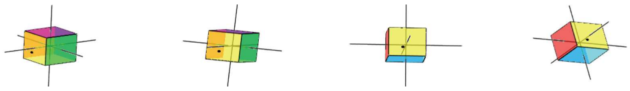

**Figure 6. A motion sequence of using one-touch interactions. The black dot represents the touch point of the finger [6].**

*Two-Touch Input:* With this method, using two points of contact (4DOF), the output degree of freedom can be extended up to six. The user can perform 2D rotations and translations with the index finger of the dominant hand, while performing at the same time pitch and roll rotations with the index finger of the non-dominant hand.
*Translation:* Translations work like in the one-touch method.
*Rotation on plane:* To perform a rotation on plane, the user has to grab anywhere outside the center and move the object in the desired direction.
*Rotation through entire space:* To rotate the object through the entire space, the user has to hold the object with the index finger of the dominant hand and drag it with the index finger of the non-dominant hand.

*Three-Touch Input:* This method uses three points of contact. The first one (index finger of the dominant hand) is used for translation, the second one (thumb of the dominant hand) for yaw about the first point, and the third contact point (index finger of the non-dominant hand) for pitch and roll about the center of the object.
*Translation:* To translate the object, the user has to grab anywhere on the object and move it.
*Rotation on plane:* For rotations on plane, the user has to twist the object with both fingers of the dominant hand.
*Rotation through entire space:* Rotations through the entire space work like in the two-touch method.

Although there is a risk that this freedom can be confusing for users, the study reveals, that the interactions are fastest and most accurate when using the three-touch technique. Beside that, the participants also felt most comfortable when using this interaction method.

The study consists of two tasks, the *Passing-Task* and the *Docking-Task*. The primary measure is the task completion time. After each experiment, the scientists analyzed the interaction techniques, including the spent time for touching, translating and rotating the objects. They also checked the locations of the objects that the users touched.

For the experiments they took a front-projected tabletop display using DiamondTouch [4]. The multi-finger interaction was possible by using distinct DiamondTouch sensors and a DiamondTouch pad.

*Passing Task:* The first task was to pass a cube to one of three *virtual* people with a certain side of the cube facing upward and toward that target person. The start position of the cube was the center of the table. This virtual persons were located to the left, right and opposite side of the table, as shown in Figure 7. The target person was indicated by a red coloring. This experiment was done for every side of the cube, with

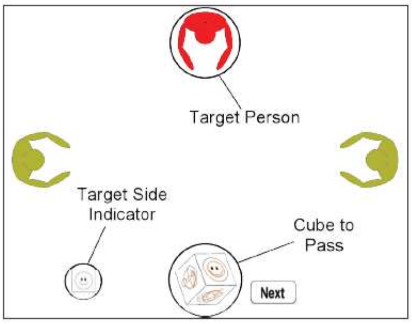


**Figure 7. The Passing-Task, where the user has to pass a cube to another virtual person with a certain upward facing cube side [6].**

every technique and to each possible virtual person (54 trials per participant).

*Docking Task:* The docking-task has the function to explore the performance differences in the three mentioned techniques. In this task the participants had to dock a tetrahedron inside another of equal size. The vertices and edges of the objects were colored and the edges were also haloed to help the user in terms of orientation and depth. For example, when a vertex was moved within the target range, the vertex would change his color. Every participant had 40 seconds for each trial. After that given time limit, the next trial started automatically.

This study showed that the techniques which use a higher number of touches were both better in terms of performance and user preference. The authors suppose that these benefits appear because the higher number of touches provided users with the opportunity to independently control more degrees of freedom. They say that shallow-depth was easily understood and interpreted as a natural environment, the people were enthusiastic about manipulation in 3D and a higher number of touches allows more natural and flexible interactions.

## Use of the 3rd dimension

In 2007 Terrenghi et al. published the results of a comparative study on manipulating physical and digital objects [15]. They figured out that the ways of manipulation are fundamentally different from each other. The study deals with two main tasks: A 25-pieces puzzle task and a photo sorting task with 40 photos. Each task was arranged with two different modes: In a physical 3D space and on a multi-touch, interactive 2D tabletop. Thereby, the authors assume that the best resource for designing specific interactive surfaces is still the physical world.

Completing the puzzle task took longer in the digital mode, whereas in the sorting task no significant time differences occurred. After the experiments, the scientists asked the participants to evaluate their experience. They rated their experience of the digital puzzle as more frustrating than the physical puzzle and they were evenly spread over which method in the sorting task was easier. However, they were convinced that physical photo sorting was more enjoyable [15].

Furthermore, the authors observed and evaluated the methods of interaction. One main aspect was the question to which degree the participants are acting one or two-handed and how the nature of these interactions was.

The coded observations showed that there exists a predominance of one-handed interactions in the digital tasks, whereas in the physical tasks bimanual interactions was much more dominant. There were also differences in the nature of two-handed interactions in the physical and digital mode. The participants which used bimanual interactions in the digital mode, used only symmetric actions. In the physical mode 11 of 12 participants used asymmetric actions.

Although the participants were trained to interact bimanually and additionally had no time limits, the results show that they interact differently as in the physical 3D world. This supports the statements of the participants that acting on a 2D tabletop is less enjoyable. Although the tabletop was able to use two-handed asymmetric gestures, nobody intuitively had the motivation to make use of it. Furthermore, the participants needed much more time for solving the tasks and they did not behave like in a 3D environment. The use of the third dimension offers by, for example picking up a photo or puzzle piece, the possibilities of focusing, selecting and keeping objects separate from others close to the body without moving the body towards objects. This offers a greater range of flexibility for dealing and manipulating the artifacts in the different tasks.

Because the 3rd dimension seems to be such an important component for intuitive interactions, some approaches for the use of the 3rd dimension will be now presented.

### *Mid-air Interactions*

Hilliges et al.(2009) [8] present a novel shadow-based technique for connecting the user's hand in the physcial world with virtual objects in the digital world. The goal of the authors is to use the space above the tabletop display to enable a more intuitive manipulation of digital objects in three dimensions. A simple example is the task of picking up a ball and placing it into a cup, like shown in Figure 8. Using 2D techniques makes it difficult and unnatural because all interactions are bound to the surface. An important condition for the developers was the idea of a technique which closely resembles the ways of manipulation of the real world. The authors implemented their technology on two rear projection-

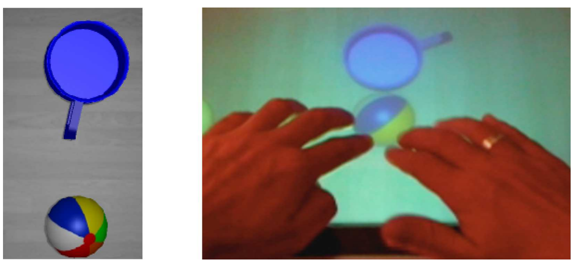

**Figure 8. Limitations of the current 2D techniques: It is difficult to place the ball into the cup [8].**

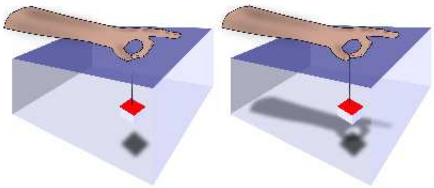

**Figure 9. Shadow feedback for connecting the user's hand with digital objects [8].**

vision tabletops and used existing and new computer vision methods to sense hand gestures and postures above the table.

To avoid the loss of directness they developed a shadow-based technique which provides feedback during interactions in the air. That feedback technique drops shadows from the user's hands above the surface into the virtual 3D scene and makes a mapping between actions in the real (physical) world and interactions within the virtual world possible (see Figure 9). This method allows a closer coupling between the input and output spaces. This approach allows sensing up to a half meter above the tabletop, to estimate the heights of hands and to detect simple pinch gestures. Beside that, they have built a tabletop system which is based on a depth camera and a holoscreen. In the end they have also implemented a tabletop system with high DOF 3D interactions without requiring any user instrumentations. Their future work will consist of building a physics-based user interface to afford a more intuitive and natural user interaction.

### *Z-touch*

A second approach for using the 3rd dimension is the *Z-touch* approach from 2010 [14]. Z-touch is a multi-touch table that can sense postures of fingers or hands within a proximity to the surface. Two key components of the system are the eight multilayered infrared (IR) line laser modules (at each corner and at the center of each edge) and the high-speed camera below the tabletop. Z-touch uses multiple laser layers to detect the distance from the surface and angles of each finger. These laser layers are synchronized with the high-speed camera's shutter signal. Two features are provided within that system: *Hand posture detection* and *hover detection*.

*Hand posture detection:* The system is able to detect the finger angles by matching the appearance of different finger postures in laser plane images from different heights (see Figure 10). The system detects blobs in each laser layer image, as indicated by rectangles in Figure 11. The position and size of these blobs is then used to determine the angle of a user's finger. That allows the control of multiple parameters with a single finger. A simple example would be the control of the direction of a Bezier curve depending on the finger angle.

*Hover detection:* Interaction techniques similar to the mouse-hover mode at desktops and other graphical user interfaces become possible with the Z-Touch approach. Users are e.g. now able to hover over objects to get additional informations without touching them.

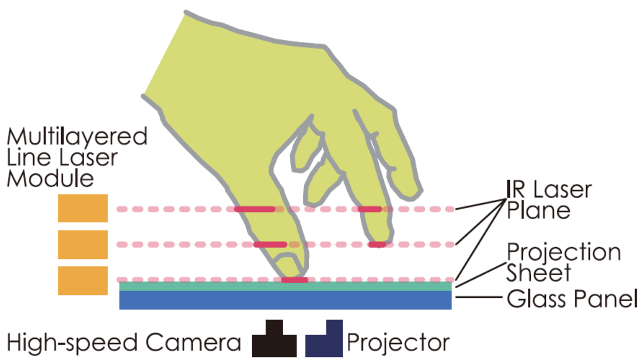


**Figure 10. The Z-Touch configuration. Multiple line laser planes recognize finger angles by matching the detected distances at different heights [14].**

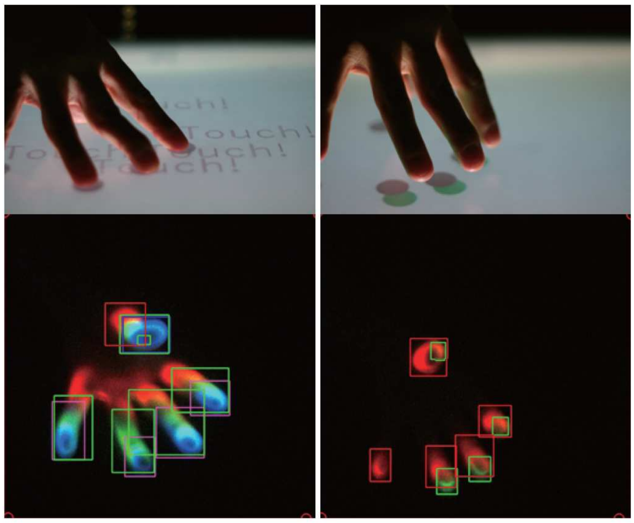

**Figure 11. Touching vs. hovering. The position of a rectangle indicates where a blob has been detected. In this case the right hand is hovering over the surface [14].**

Different colors for each laser plane visualize the depth maps detected by the camera, as shown in the bottom row images in Figure 11. In this case the blue colors mean, that the hand is touching the tabletop. According to this, the right hand is only hovering, while the left hand is touching the surface.

Possible implemented interaction techniques are a drawing application and a map zooming application. In the drawing application the colors correspond to the depth values of the depth map and in the map zooming application the zoom level depends on the lowest height of a finger.

### *Extended Pen Devices*

Withana et al.(2010) [17] and Lee et al.(2010) [10] present two similar input devices, *ImpAct* and *Beyond*, for direct 3D manipulation. Whereas in the aforementioned interaction technique by Hilliges et al. the user's body is separated from the tabletop screen, with the pen-like devices Beyond

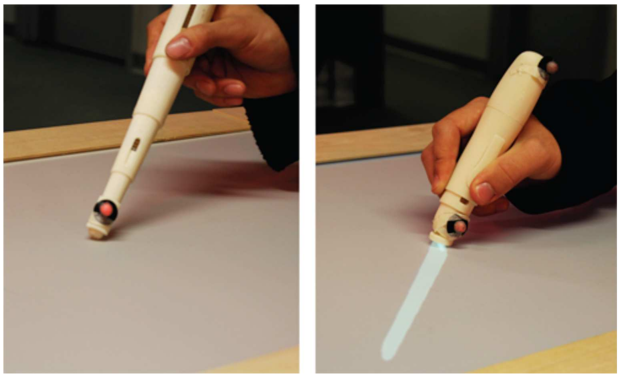

**Figure 12. When pressing against the surface, the Beyond device collapses and is visually extended in the digital space. [10].**

and ImpAct the user can directly manipulate the 3D models. When pressing the devices against the screen, they collapse due to their telescopic construction but simultaneously the devices get visually extended into the digital space, as shown in Figure 12 for the Beyond device. This allows the user to perform manipulation tasks like selection, rotation or drawing directly in the virtual 3D space.

In Beyond [10] additional hand gestures with the free hand are implemented to ease the drawing of lines, curves, squares and ellipses as well as to support the extrusion and rotation of 3D models. Furthermore, the user's head is tracked by a camera and the 3D models are rendered corresponding to the user's head position, which allows to view an object from different angles. However, due to this view-dependent rendering, Beyond is not constructed for multi-user interactions.

**Interaction with stereoscopic and AR displays**

Concerning the interaction with 3D data, tabletops using stereoscopic displays do not differ much from systems using augmented reality. In both technologies the third dimension is realized by using visual effects resulting in a lack of haptic feedback when trying to touch the object. The main difference between the two concepts is the number of displays. Since in stereoscopic tabletops one display is used to visualize the data, it cannot be perceived correctly from all perspectives, as described in the chapter *Visualization of 3D data*. Especially when using motion parallax to allow looking around the object, only the tracked user has a correct perception. In augmented reality systems the 3D data is visualized on each head-mounted display resulting in a correct depth perception for every user.

In order to perform manipulations in the third dimension either hand gestures or further 6DOF input devices are suitable. However, hand gestures provide more natural interaction concepts which can be learned faster than using input devices where for example different functions are mapped to identically looking buttons. Benko et al. [2] present several interaction concepts for their VITA system including 2D and 3D interactions. An overview of the VITA system is given in the chapter *Visualization of 3D data*. The user interacts with the system by using a tabletop for 2D navigation of the dig site. When multiple users participate, they have to wear or sit on conductive pads in order to get distinguished by the tabletop, which is a DiamondTouch [4] table. The user can also perform *cross-dimensional gestures*.

*Cross-dimensional gestures*

Hybrid or cross-dimensional gestures combine interactions in 2D and 3D contexts. Several such cross-dimensional gestures are presented in [3]. These gestures provide a connection between the data visualized on the tabletop display and the data visualized on the head-mounted displays. This kind of connection is required when the user wants to move an object from the 2D tabletop surface to the 3D space visualized on his head-mounted display, for example. The developed gestures can be categorized into the two classes *one-handed* and *two-handed*. In order to perform these hand gestures, the user has to wear an instrumented glove for gesture recognition and a hand-tracker for position recognition.

There are four basic hand gestures which have different meanings on 2D and 3D contexts, as shown in Figure 13. For example, selecting and moving an object in the 2D tabletop environment is implemented by a tap gesture followed by dragging the object, while in the 3D environment it is implemented as a grab gesture followed by moving the object in the 3D space. The cross-dimensional gestures are combinations of these four basic gestures.

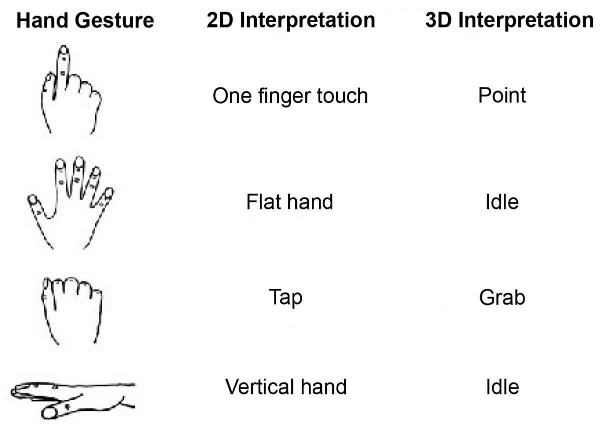


**Figure 13. Basic hand gestures with different interpretations in 2D and 3D environment [3].**

The transition from the tabletop representation to the head-mounted display 3D representation of an object is implemented by the *cross-dimensional pull* gesture. Performing this gesture the user first places his hand flat on the tabletop where the object is displayed. Then the user moves his fingers to the center of the object and finally forms a fist. During this gesture the object on the tabletop shrinks until it disappears and the 3D representation of the object appears on the user's head-mounted display. The sequence is shown in Figure 14 a-c.

The inversive interaction is the *cross-dimensional push* gesture, which moves an object from its 3D representation to a representation on the tabletop. To perform this gesture the

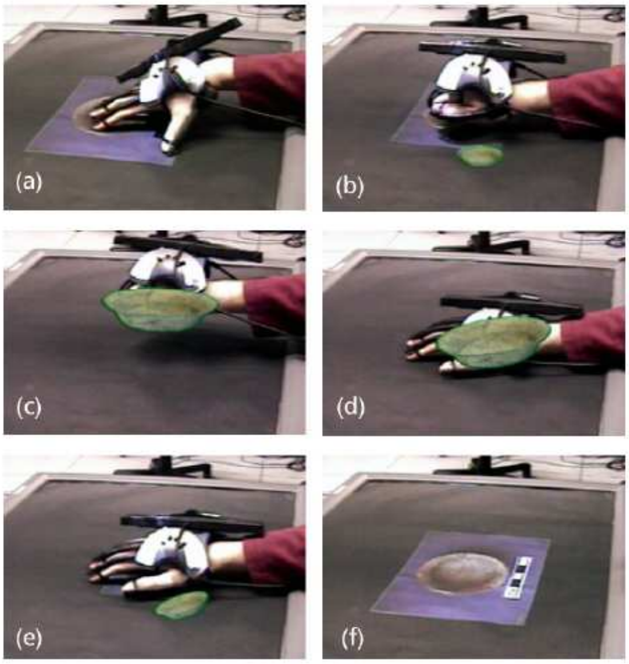


**Figure 14. Cross-dimensional pull gesture (a-c): Selecting the object (a), forming a fist (b) and grabbing the object (c). Cross-dimensional push gesture (d-f): Pushing object towards the table (d), the object shrinks and disappears (e) and appears on the tabletop (f). [3].**

user places a flat hand above the 3D object and pushes the object down until his hand touches the surface of the tabletop. The interaction sequence is shown in Figure 14 d-f. Both cross-dimensional gestures are intuitive and preserve the interaction context.

Another gesture is the *cross-dimensional connect*. By grabbing the object in the 3D space, moving it towards the table and then tapping somewhere on the table, the user can connect the 3D representation of the object with its 2D representation on the tabletop. A leader line visualizes the connection between the two objects, as shown in Figure 15. Tapping again on the table breaks the connection and the 2D representation disappears from the tabletop.

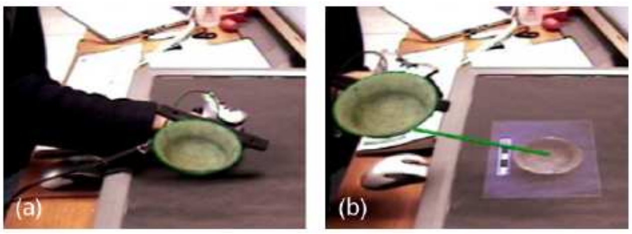


**Figure 15. Cross-dimensional connect gesture. The leader line (green) visualizes the connection between the 3D and 2D objects [3].**

Users normally not only want to select and move objects but also manipulate them by scaling and rotating them. Such interactions are not easy to implement in an intuitive manner when there is no haptic feedback. Therefore, Benko et al. provide two-handed gestures for the 3D manipulation of objects. The hand wearing the instrumented glove is used to hold and rotate the object in the 3D space whereas the other hand uses the tabletop to perform the manipulation tasks like scaling.

*User study*

Benko et. al evaluated their VITA system in a user study with three archeologists and three archeology students. Whereas the feedback concerning the interaction with the system was positive, the users mentioned that wearing all the required devices is uncomfortable when using the system for a longer time. This is a problem of many augmented reality systems, since there head-mounted displays, instrumented gloves (or other input devices) and tracking devices have be worn in order to use the system.

### Spatially augmented reality

Spatially augmented reality tabletop systems provide other interaction concepts since here the 3D objects are not virtual but physical. Such interfaces are called tangible user interfaces (TUI) because they make digital information directly manipulatable by the user's hands. There is a high variety of objects used in such tangible user interfaces such as sand [9], clay [13] or blocks of different shapes [1]. TUIs have several advantages over graphical and multi-touch interfaces. The main advantage is the existence of tactile and even sound feedback. Having this feedback the user can manipulate the objects without using his eyes. Another advantage is the familiarity of the interaction functions because everyone knows how a physical block can be translated or rotated or how sand can be formed to create the desired topography. In graphical or touch interfaces these functions have often to be learned before the user can effectively interact with the system. Furthermore, the user does not need any additional glasses or head-mounted displays to perceive the data without perspective distortion.

Tabletop systems using TUIs for interaction need to detect where the physical objects are placed on the table in order to compute the corresponding output which is then projected on the table or on the objects themselves. Several technologies can be used to implement this detection, like magnetic trackers, electronic connectors or laser range detection systems. Recent approaches use tables with diffuse illumination which can track 2D markers on the bottom side of the

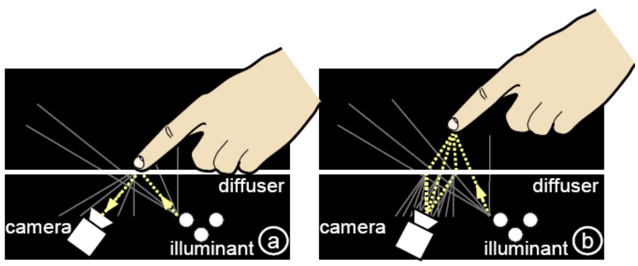


**Figure 16. (a) Light is reflected with less scattering when the finger touches the surface. (b) Light is first scattered before it reaches the camera when the finger is above the surface [1].**

objects. In tabletops using diffuse illumination, the screen is illuminated from under the surface. The markers at the bottom of the objects reflect the light which is sensed by a camera under the screen. The advantage of this approach is that the interaction objects are unpowered and thus need no maintaining batteries which makes a larger number of these objects manageable. However, the disadvantage of such systems is that the camera cannot detect objects above the table surface. This is illustrated in Figure 16. The farther an object is above the surface, the more the light is scattered and the recognized image by the camera gets blurry making it impossible to interpret. The consequence is that 3D structures like putting several blocks on each other are not recognized by the system. The work of Baudisch et al. [1] directly addresses this problem and presents tangible blocks called Luminos which contain glass fiber bundles additionally to their markers.

The main idea of Luminos is to transfer the markers of the higher level blocks down to the table surface such that the camera can detect them. This is done by using glass fiber bundles which guide the light such that the markers of the vertically arranged objects are mapped to a horizontal space, namely the bottom face of the undermost block. This concept is shown in Figure 17a. Each block has a marker on one side of its bottom face, here illustrated as a black stripe. The glass fiber bundle shifts the markers of higher level blocks on their way down. The higher the block, the higher the offset will be. This horizontal arrangement is achieved by using skewed glass fiber bundles, as shown in Figure 17b. One problem of this construction is that the blocks have to be precisely aligned so that the markers do not occlude each other. To leverage this problem, the markers do not only get shifted but also scaled down which can be realized by using glass fiber bundles that have the same effect as if they were stretched in one dimension, as illustrated in Figure 17c. However, this down-scaling at each level exponentially decreases the marker size actually seen by camera. Besides using quadratic blocks, Baudisch et al. also present round blocks, as shown in Figure 18. The *twister* (Figure 18a) arranges the markers in a circle while the *taper* (Figure 18b) maps outer ring markers to inner rings.

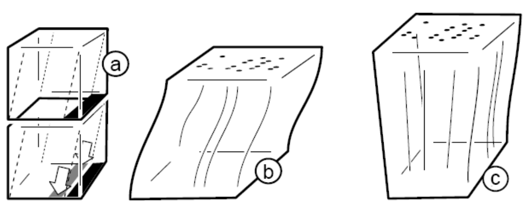


**Figure 17. (a) The marker of the higher block is shifted. (b) Exemplified glass fiber bundle form for shifting. (c) Exemplified glass fiber bundle form for scaling down the markers [1].**

There are several applications of the Lumino blocks possible. One example is the construction kit which is shown in Figure 19. There are three types of blocks: 1x1, 1x2 and 1x4. The 1x2 and 1x4 blocks have different markers for each single unit. This makes it possible for the system to define how

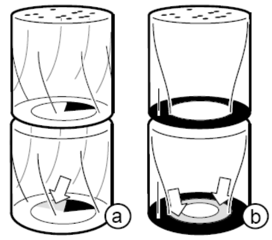


**Figure 18. Round Lumino blocks: Twister (a) and taper (b) [1].**

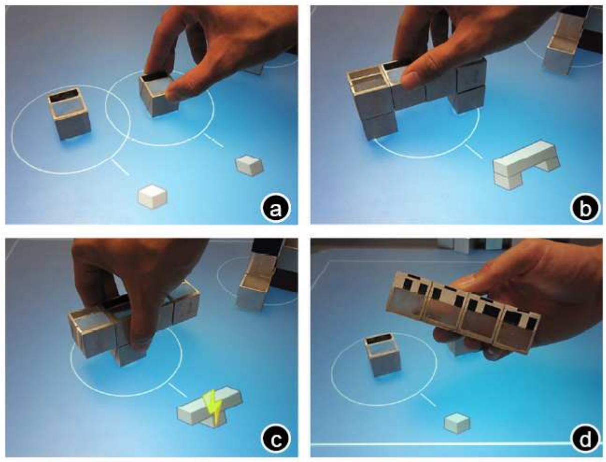


**Figure 19. Lumino construction kit application. Labeling of single (a) and composed (b) blocks. (c) Warning about bad construction. (d) Different markers for each unit of an 1x4 block [1].**

the blocks are arranged. Furthermore, the system assists the user by pointing out bad constructions such as the overhang warning in Figure 19c. Other additional information like 3D models of the construction or the current construction cost are imaginable.

## CONCLUSION

In our report we gave an overview of current visualization technologies for 3D models on tabletops. Therein, we discussed different interaction problems which depend on the used technology. By displaying 3D models on a 2D tabletop display problems of view distortion and interacting occur. The problem is that the interaction is constrained to a 2D surface and tasks which need a third dimension like putting an object upon another one are not realizable in an intuitive way. Therefore, several approaches exist which extend the input space into the third dimension by using the space above the table or by using the digital space inside the table. When using augmented reality systems, there are two contexts in which the data can be represented, namely the tabletop display and the 3D space generated by the head-mounted display. Therefore, interaction techniques which combine these two contexts are required, as proposed by the cross-dimensional hand gestures in VITA. In spatially augmented tabletops physical proxies, like blocks or clay, are used to represent the 3D models. Such systems have to detect the position and identity of the tangible objects on the table surface. However, constructing 3D structures requires the system to sense objects in higher levels which can be difficult when using diffuse illumination. This report discussed the named problems and presented solutions from the ongoing research in this area.

Our report points out that no *perfect* solution for 3D visualizations on tabletops exists. Every mentioned technology has its own drawbacks and advantages concerning its usability. When creating a tabletop system which uses 3D data, first the preferred visualization technology has to be determined. Depending on the used technology, the mentioned interaction problems have to be taken into account when designing interaction techniques for the tabletop system.

## REFERENCES


1. P. Baudisch, T. Becker, and F. Rudeck. Lumino: tangible blocks for tabletop computers based on glass fiber bundles. In *CHI '10: Proceedings of the 28th international conference on Human factors in computing systems*, pages 1165–1174, New York, NY, USA, 2010. ACM.
2. H. Benko, E. W. Ishak, and S. Feiner. Collaborative mixed reality visualization of an archaeological excavation. In *ISMAR '04: Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 132–140, Washington, DC, USA, 2004. IEEE Computer Society.
3. H. Benko, E. W. Ishak, and S. Feiner. Cross-dimensional gestural interaction techniques for hybrid immersive environments. In *Proceedings of the IEEE Virtual Reality. Conference*, pages 209–216, 2005.
4. P. Deitz and D. Leigh. Diamondtouch: A multi-user touch technology. In *Proc. UIST 2001*, pages 219–226. ACM Press, 2001.
5. T. Grossman and D. Wigdor. Going deeper: a taxonomy of 3d on the tabletop. In *2007 IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, pages 137–144. IEEE, October 2007.
6. M. Hancock, S. Carpendale, and A. Cockburn. Shallow-depth 3d interaction: design and evaluation of one-, two- and three-touch techniques. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1147–1156, New York, NY, USA, 2007. ACM.
7. M. Hancock, M. Nacenta, C. Gutwin, and S. Carpendale. The effects of changing projection geometry on the interpretation of 3d orientation on tabletops. In *ITS '09: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, pages 157–164, New York, NY, USA, 2009. ACM.
8. O. Hilliges, S. Izadi, A. D. Wilson, S. Hodges, A. G. Mendoza, and A. Butz. Interactions in the air: adding further depth to interactive tabletops. In *UIST '09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pages 139–148, New York, NY, USA, 2009. ACM.
9. H. Ishii, C. Ratti, B. Piper, Y. Wang, A. Biderman, and E. Ben-Joseph. Bringing clay and sand into digital design — continuous tangible user interfaces. *BT Technology Journal*, 22(4):287–299, 2004.
10. J. Lee, S. Teerapittayanon, and H. Ishii. Beyond: collapsible input device for direct 3d manipulation beyond the screen. In *Adjunct proceedings of the 23nd annual ACM symposium on User interface software and technology*, UIST '10, pages 393–394, New York, NY, USA, 2010. ACM.
11. M. Mathews, M. Challa, C.-T. Chu, G. Jian, H. Seichter, and R. Grasset. Evaluation of spatial abilities through tabletop ar. In *CHINZ '07: Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction*, pages 17–24, New York, NY, USA, 2007. ACM.
12. K. Nakashima, T. Machida, K. Kiyokawa, and H. Takemura. A 2d–3d integrated tabletop environment for multi-user collaboration: Research articles. *Comput. Animat. Virtual Worlds*, 18(1):39–56, 2007.
13. B. Piper, C. Ratti, and H. Ishii. Illuminating clay: a 3-d tangible interface for landscape analysis. In *Proceedings of the CHI 2002*, pages 355–362. ACM Press, 2002.
14. Y. Takeoka, T. Miyaki, and J. Rekimoto. Z-touch: a multi-touch system for detecting spatial gestures near the tabletop. In *ACM SIGGRAPH 2010 Talks*, SIGGRAPH '10, pages 57:1–57:1. ACM, 2010.
15. L. Terrenghi, D. Kirk, A. Sellen, and S. Izadi. Affordances for manipulation of physical versus digital media on interactive surfaces. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1157–1166, New York, NY, USA, 2007. ACM.
16. D. Wigdor, C. Shen, C. Forlines, and R. Balakrishnan. Perception of elementary graphical elements in tabletop and multi-surface environments. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 473–482, New York, NY, USA, 2007. ACM.
17. A. Withana, M. Kondo, G. Kakehi, Y. Makino, M. Sugimoto, and M. Inami. Impact: enabling direct touch and manipulation for surface computing. In *Adjunct proceedings of the 23nd annual ACM symposium on User interface software and technology*, UIST '10, pages 411–412, New York, NY, USA, 2010. ACM.