Multi–Touching 3D Data: Towards Direct Interaction in Stereoscopic Display Environments coupled with Mobile Devices

Frank Steinicke, Klaus Hinrichs
Visualization and Computer Graphics
Westfälische Wilhelms-Universität Münster
Einsteinstraße 62, Münster, Germany
{fsteini,khh}@uni-muenster.de

Johannes Schöning, Antonio Krüger
Institute for Geoinformatics
Westfälische Wilhelms-Universität Münster
Robert-Koch-Str. 26-28, 48149 Münster,
{j.schoening,krugera}@uni-muenster.de

ABSTRACT
In many different application domains the use of 3D visualization is accelerating. If the complexity of 3D data increases often stereoscopic display provides a better insight for domain experts as well as ordinary users. Usually, interaction with and visualization of the 3D data is decoupled because manipulation of stereoscopic content is still a challenging task. Hence, 3D data is visualized stereoscopically whereas interaction is performed via 2D graphical user interfaces. Although such *interscopic* interaction between stereoscopic and monoscopic content is of major interest in many application domains it has not been sufficiently investigated. Recently emerging multi–touch interfaces promise an alternative approach to this challenge. While multi–touch has shown its usefulness for 2D interfaces by providing more natural and intuitive interaction, it has not been considered if and how these concepts can be extended to 3D multi–touch interfaces, in particular in combination with stereoscopic display. In this paper we discuss the potentials and the limitations as well as possible solutions for the interaction with interscopic data via multi–touch interfaces.

1. BACKGROUND AND MOTIVATION
In recent years virtual environments (VEs) have become more and more popular and widespread due to the requirements of numerous application areas. Two–dimensional desktop systems are often limited in cases where natural interfaces are desired. In these cases virtual reality (VR) systems using tracking technologies and stereoscopic projections of three–dimensional synthetic worlds support a better exploration of complex data sets. Although costs as well as the effort to acquire and maintain VR systems have decreased to a moderate level, these setups are only used in highly specific application scenarios within some VR laboratories. In most human-computer interaction processes VR systems are only rarely applied by ordinary users or by experts – even when 3D tasks have to be accomplished [1]. One reason for this is the inconvenient instrumentation required to allow immersive interactions in such VR systems, i. e., the user is forced to wear stereo glasses, tracked devices, gloves etc. Furthermore the most effective ways for humans to interact with synthetic 3D environments have not finally been determined [1, 3]. Even the WIMP metaphor [14], which is used for 2D-Desktop interaction, has its limitations when it comes to direct manipulation of 3D data sets [6], e. g., via 3D widgets [7]. Devices with three or more degrees of freedom (DoFs) may provide a more direct interface to 3D manipulations than their 2D counterparts, but using multiple DoFs simultaneously still involves problems [3]. Most 3D applications also include 2D user interface elements, such as menus, texts and images, in combination with 3D content. While 3D content usually benefits from stereoscopic visualization 2D GUI items often do not have associated depth information. Therefore, interactions between monoscopic and stereoscopic elements, so–called *interscopic interactions*, have not been fully examined with special consideration of the interrelations between the elements.

Multi–touch interaction with computationally enhanced surfaces has received considerable attention in recent years. When talking about multi-touch surfaces we think of surfaces that support multi-finger and multi-hand operation (in analogy to the seminal work by Bill Buxton [5]). Multi–touch surfaces can be realised by using different technologies, ranging from capacitive sensing to video analysis of infrared or full color video images. Recently the promising FTIR (frustrated total internal reflection) technology has been rediscovered by Jeff Han [12]. Its cheap footprint has accelerated the usage of multi–touch in the last two years. If multi–touch applications need to distinguish between different users, the *Diamond Touch* concept from MERL [8] could be used, with the drawback that the users either need to be wired or stay in specially prepared locations. Another benefit of multi–touch technology is that the user does not have to wear inconvenient devices in order to interact in an intuitive way [16]. Furthermore, the DoF are restricted by the physical constraints of the touch screen. In combination with autostereoscopic displays such a system can avoid any instrumentation of the user, while providing an advanced user experience. However, the benefits and limitations of using multi–touch in combination with stereoscopic display have not been examined in-depth and are not well understood. Our experiences make us believe that mobile devices

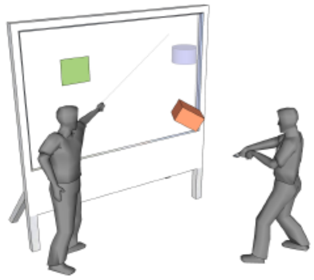

Figure 1: Illustration of two users interacting with stereo- as well as monoscopic content.

with multi–touch enabled surfaces, such as the iPhone/iPod touch, have great potential to support and enrich the interaction with large scale stereoscopic projection screens or even in immersive virtual reality. In this position paper we discuss challenges of such user interfaces for stereoscopic display setups and in particular the role multi–touch enabled mobile devices could play in those environments.

The paper is structured as follows: In section two we discuss issues related to the parallax-dependent selection and direct manipulation of 3D objects as well as issues related to navigation in 3D data sets. These issues have to be taken into account when designing a multi–touch user interface for 3D interaction. In addition, we will illustrate how the combination of a mobile multi–touch device and a stereoscopic multi–touch wall can enrich the interaction and solve existing interaction problems. Furthermore, we discuss application areas that show the potential for the interaction with stereoscopic content via multi–touch interfaces, in particular multi-touch enabled mobile devices. Section 3 concludes the paper.

2. MULTI–TOUCHING 3D DATA

In this section we discuss aspects which have to be taken into account when designing a multi–touch user interface for interscopic interaction.

2.1 Parallax Paradigms

When stereoscopic display is used each eye of the user perceives a different perspective of the same scene. This can be achieved by using different technologies, either by having the user wear special glasses or by using special 3D displays. Although the resulting binocular disparity provides an additional depth cue, in a stereoscopic representation of a 3D scene it may be hard to access distant objects [3]. This applies in particular if the interaction is restricted to a 2D touch surface. Objects might be displayed with different parallax paradigms, i. e., negative, zero, and positive parallax, resulting in different stereoscopic effects. Interaction with objects that are displayed with different parallaxes is still a challenging task in VR–based environments.

2.1.1 Negative Parallax

When stereoscopic content is displayed with negative parallax the data appears to be in front of the projection screen (see orange-colored box in Figure 1). Hence, when the user wants to interact with data objects by touching, s/he is limited to touch the area behind the objects since multi–touch screens capture only direct contacts. Therefore, the user virtually has to move fingers or her/himself through virtual objects, and the stereoscopic projection is disturbed. Consequently, immersion may get lost. This problem is a common issue known from two-dimensional representation of the mouse cursor within a stereoscopic image. While the mouse cursor can be displayed stereoscopically on top of stereoscopic objects [18], movements of real objects in the physical space, e. g., the user's hands, cannot be constrained such that they appear only on top of virtual objects.

2.1.2 Zero Parallax

If stereoscopic content is displayed with zero parallax an object appears to be aligned with the projection screen (see green-colored rectangle in Figure 1). Hence both eyes perceive the same image which causes a two-dimensional impression. As mentioned in Section 1, for such a situation multi–touch interfaces have considerable potential to enhance the interaction process, in particular when 2D manipulations are intended.

2.1.3 Positive Parallax

When stereoscopic content is displayed with positive parallax the data appears to be behind the projection screen (see purple-colored cylinder in Figure 1). These distant objects can not be accessed directly via virtual touch since the projection screen limits the reach of the user's arms. This is a common problem known from VR-based environments, and several approaches address this issue [15, 3]. Some of these approaches, in particular image plane techniques, are even applicable with multi–touch displays. When using image-based approaches, the interaction is performed on the projection screen analogous to a 2D mouse. Selection can be performed by casting a ray from the dominant eye of the user through the touch position on the screen (see Figure 1). The first object hit by the ray is the active object the user can select, e. g., by pressing a button. On a multi–touch screen even a pinch gesture can be used to perform the selection of the object underneath the fingers.

Possible Solution of Parallax Problems

One solution might be to allow a user to interactively change the parallax of objects by using a mobile device attached to the user's body as a "soft slider". If the touch–surface is portable the screen can be moved through the VE (analog to a 3D window metaphor) until desired objects are displayed with zero or negative parallax and interaction can be performed as described in Sections 2.1.2 and 2.1.3. An interesting alternative proposed by Zadow et al. [20] recognizes the positions of the user's hands not only on the surface but also above it.

2.2 3D Manipulation

3D manipulation (*selection*, *translation*, *rotation* and *scaling*) of objects on stereoscopic displays is a complex task. A major goal when performing direct 3D interaction is to manipulate an object in terms of its position and orientation in space. For two–dimensional manipulation multi–touch has proved to be a very powerful interface paradigm. Objects can be manipulated by means of a single or multiple fingers or with the palm or edge of one or both hands; even different levels of pressure can be applied [5]. However, when the user's interaction is restricted to a two–dimensional touch surface the specification of six DoF gets non-intuitive and complicated gestures may be required [9, 10, 21].

2.2.1 Selection

Before a user can interact with virtual objects the desired targets have to be identified. This task has to be accomplished by an interaction technique itself. When interaction is restricted to a 2D surface, selection can be implemented by using image-plane techniques [15, 2, 19] (see Section 2.1).

2.2.2 Translation

When a 3D object is selected and translation is intended, a movement in the plane parallel to the surface can be implemented easily. For example, contacts on the projection screen's local x and y direction can by mapped one-to-one to the virtual object. Translations are constrained to the orientation of the touch–screen surface. Since perspective projection is usually applied when stereoscopy is used, this mapping may be disadvantageous because distant objects appear to move more slowly in image-space than objects close to the projection screen. Therefore, different mapping strategies may be applied, for instance, a projected distance can be mapped [19]. However, when translation along the z direction of the screen's coordinate system is desired, different approaches have to be considered. For instance, gestures can be used to specify a translation along the depth axis, but users need to learn different non-intuitive gestures.

2.2.3 Rotation

Rotation in 2D can be implemented very naturally. For instance, one touch point determines the center of rotation, while the amount of rotation is specified by circular movements around the center. Thus objects can be rotated via two contacts only. In 3D the center of rotation and the rotation axis have to be determined by means of a 3D point and a vector. Since a touch surface constrains the user's action to 2D, rotations in 3D are difficult to realize.

2.2.4 Scaling

While scaling in 2D can be implemented very intuitively using a multi–touch interface, for example, by means of one– or two–handed pinch gestures, scaling in 3D is complicated. In particular, if non–uniform scaling is intended, an intuitive specification of the scaling vector to be applied to the virtual object is a challenging task. Even in VR-based environments non–uniform scaling is often implemented via indirect approaches, e. g., GUI widgets.

Approaches for Multi–Touch 3D Manipulation

Position and orientation of mobile multi–touch surfaces can be tracked very accurately and could therefore be used for specifying fine-grained input data. The orientation of the device could be used to provide precise data, in particular 3D vectors which could otherwise not be specified by the rather coarse multi–touch input alone. Such separation between precise and coarse interaction performed with the dominant and non-domination hand, respectively, is also applied in 2D multi–touch interfaces [4]. Likewise translation in space can be implemented by using the mobile device's orientation that determines the axis along which a translation of an object is to be performed. In the same way the device's orientation can define a rotation axis or a non-uniform scaling vector.

2.3 Navigation

Since navigation is the most common interaction task in VEs it is essential to provide intuitive mechanisms to enable users to explore large and complex virtual environments. Essentially navigation is similar to performing 3D object manipulation, whereas when exploring a VE manipulations are applied to the virtual camera. Current navigation techniques exploiting multi–touch devices are limited to simple panning, zooming or rotation approaches [12]. Usually, the non-dominant hand poses a predefined gesture that determines the navigation mode, while the dominant hand specifies the amount of movement. Since the touch is only used to define certain modi multi–touch is actually degraded to single touch. It has not been examined how multi–touch, for instance by using the entire hand surface, can enhance this process.

Possible Solution of Navigation problems

For single touch interfaces there are already intuitive mechanisms to implement certain camera movements [11]. Such traveling metaphors can be realized by means of specifying direction, speed, velocity, etc. of the virtual camera. As mentioned above mobile devices equipped with orientation sensors may be exploited to define the orientation of the virtual camera. All movements of the camera may be determined by the touch interface of the mobile devices. This is especially beneficial for presentation scenarios, where the presenter is using a mobile device to guide other viewers through a VE. Alternatively the touch surface itself can be used as a navigation device. Camera movements can be initiated by virtually pushing the touch screen. For instance, pressing the touch screen on the right side yields a camera rotation to the left, touching the screen at the top moves the camera downwards and vice versa. Furthermore, these concepts can be combined such that an omni-directional flying metaphor can be implemented. Hence the user gets the impression of navigating a vehicle via the window to the virtual world.

3. CONCLUSIONS

In this position paper we have discussed problems and potentials related to the use of multi–touch interfaces for the interaction with interscopic data. Figure 2 summarizes, from our point of view, the potentials of multi–touch interfaces for the interaction with stereoscopic content as well as the possibilities when using multi-touch enabled mobile devices. We are working on the realization of such a system for formal evaluation. The icons indicate whether we believe that an interaction in this combination is beneficial (green/smile), possible (yellow/neutral), or impracticable (red/sad). Of course, not all problems are covered or can be solved with such a device setup. We have mentioned some problems which might occur in such scenarios.

3D widgets can be used to integrate solutions for desktop-based environments [7]. If direct interaction is not required, users can specify 3D manipulations by means of constraint-based techniques. These widgets provide several interaction handles which themselves support different interaction tasks such as translation, rotation or scaling. The available DoFs are reduced with respect to the degrees provided by the input device. Currently multi–touch walls are horizontally or vertically mounted. VR-based display devices such as the responsive workbench allow to turn the display from horizontal to vertical. In contrast to vertical multi–touch surfaces, horizontal ones provide the possibility to place physical objects on the surface [13]. In order to enhance the perception

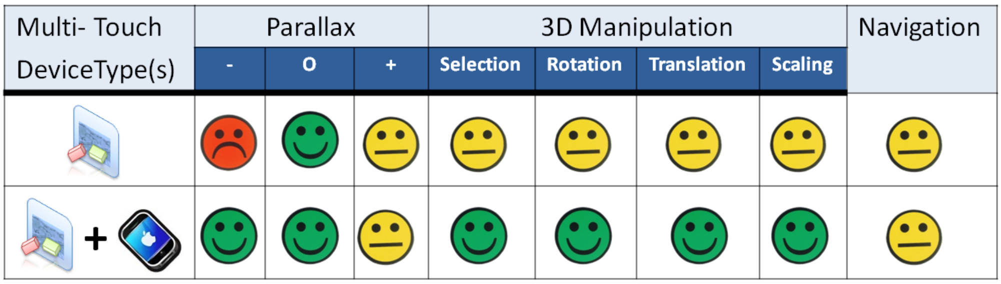

Figure 2: Potentials and limitations as well as possible solutions for using multi–touch interfaces with/without mobile devices (having a multi–touch enabled surface) to interact with stereoscopic content.

of spatial data 3D multi–touch or at least 2.5D projection screens can be exploited [17].

ACKNOWLEDGEMENTS

We thank the Deutsche Telekom Laboratories, which partially funded this research.

4. REFERENCES

[1] R. Balakrishnan. A Grant of 3D. Keynote speach Symposium on 3D User Interfaces, 2006.

[2] D. Bowman. *Interaction Techniques for Common Tasks in Immersive Virtual Environments: Design, Evaluation, and Application*. PhD thesis, Georgia Institute of Technology, 1999.

[3] D. Bowman, E. Kruijff, J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison-Wesley, 2004.

[4] P. Brandl, M. Haller, J. Oberngruber, and C. Schafleitner. Combining and Measuring the Benefits of Bimanual Pen and Direct-Touch Interaction on Horizontal Interfaces. In *Proceedings of Advanced Visual Interfaces*, page accepted, 2008.

[5] W. Buxton and B. Myers. A study in two-handed input. *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 321–326, 1986.

[6] W.-S. Chun, J. Napoli, O. S. Cossairt, R. K. Dorval, D. M. Hall, T. J. P. II, J. F. Schooler, Y. Banker, and G. E. Favalora. Spatial 3-D Infrastructure: Display-Independent Software Framework, High-Speed Rendering Electronics, and Several New Displays. In *In Stereoscopic Displays and Virtual Reality Systems XII*, volume 5664, pages 302–312, 2005.

[7] D. B. Conner, S. C. Snibbe, K. P. Herndon, D. C. Robbins, R. C. Zeleznik, and A. van Dam. Three-Dimensional Widgets. In *Symposium on Interactive 3D Graphics*, 1992.

[8] P. Dietz and D. Leigh. DiamondTouch: a multi-user touch technology. *Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 219–226, 2001.

[9] J. Epps, S. Lichman, and M. Wu. A study of hand shape use in tabletop gesture interaction. In *CHI '06: CHI '06 extended abstracts on Human factors in computing systems*, pages 748–753, New York, NY, USA, 2006. ACM.

[10] C. Forlines and C. Shen. Dtlens: multi-user tabletop spatial data exploration. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 119–122, New York, NY, USA, 2005. ACM.

[11] M. Hachet, F. Decle, S. Knödel, and P. Guitton. Navidget for Easy 3D Camera Positioning from 2D Inputs. In *Proceedings of IEEE Symposium on 3D User Interfaces (3DUI)*, pages 83–89, 2008.

[12] J. Y. Han. Low-cost multi-touch sensing through frustrated total internal reflection. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 115–118, New York, NY, USA, 2005. ACM.

[13] Microsoft Surface. `http://www.microsoft.com/surface/`, 2008.

[14] B. A. Myers. A taxonomy of window manager user interfaces. *IEEE Comput. Graph. Appl.*, 8(5):65–84, 1988.

[15] J. Pierce, A. Forsberg, M. Conway, S. Hong, R. Zeleznik, and M. Mine. Image Plane Interaction Techniques in 3D Immersive Environments. In *ACM Symposium on Interactive 3D Graphics*, pages 39–44, 1997.

[16] J. Schöning, B. Hecht, M. Raubal, A. Krüger, M. Marsh, and M. Rohs. Improving Interaction with Virtual Globes through Spatial Thinking: Helping users Ask "Why?". In *IUI '08: Proceedings of the 13th annual ACM conference on Intelligent User Interfaces*, New York, NY, USA, 2008. ACM.

[17] R. Smith, B. Thomas, and W. Piekarski. Techn Note: Digital Foam. In *Proceedings of IEEE Symposium on 3D User Interfaces (3DUI)*, pages 35–38, 2008.

[18] F. Steinicke, T. Ropinski, G. Bruder, and K. Hinrichs. Interscopic User Interface Concepts for Fish Tank Virtual Reality Systems. In *Proceedings of the Virtual Reality*, pages 27–34, 2007.

[19] F. Steinicke, T. Ropinski, and K. Hinrichs. VR and Laser-based Interaction in Virtual Environments using a Dual-Purpose Interaction Metaphor. In *VR2005 Workshop Proceedings on New Directions in 3D User Interfaces*, pages 61–64. IEEE, 2005.

[20] U. von Zadow. http://www.multi-touch.de/.

[21] M. Wu and R. Balakrishnan. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. *Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 193–202, 2003.