H. Reiterer & O. Deussen (Hrsg.): Mensch & Computer 2012
München: Oldenbourg Verlag, 2012, S. 173-182

Towards Bimanual Control for Virtual Sculpting

Marc Herrlich, Andreas Braun, Rainer Malaka

Research Group Digital Media, TZI, University of Bremen

Abstract

In 3D modeling, the work piece should be at the very center of attention. Many 3D modeling tools work rather directly on the mathematical representations, such as polygonal meshes. In contrast, virtual sculpting abstracts the 3D modeling process by providing virtual tools not unlike those used by artists for sculpting physical objects. In this paper, we investigate how virtual sculpting can benefit from bimanual interaction on interactive surfaces. We present our interface design and the results of a conducted case study. We compared bimanual interaction on interactive surfaces to bimanual interaction using a stylus together with the keyboard. We found that interactive surfaces have greater affordance for bimanual interaction and we suggest that virtual sculpting could benefit from a combination of stylus and multi-touch interaction in the future.

1 Introduction

In 3D modeling, the work piece itself should be at the very center of attention, not the interface or the mathematical model behind it. However, state of the art tools for 3D modeling and animation work rather directly on mathematical representations, e.g., polygonal meshes, as do the tools, which even require users to numerically specify the details of some operations. In recent years, virtual or digital sculpting has become more popular as an alternative method to traditional mesh oriented modeling tools. Virtual sculpting (VS) abstracts from the mathematical model by providing the user with a set of virtual tools (often called "brushes"), which to some extend mimic physical tools used by traditional sculptors and artists. For example, in most VS systems there are brushes to add material, to flatten, smooth, or crease the surface (figure 1). From an input device perspective, VS in established software tools can be performed using the mouse or a stylus/pen in combination with the keyboard, the latter combination being the favored one. Although, VS frees the artist to some extend from thinking about the underlying mathematical models, instead allowing him to focus more on the details of the work piece, VS compared to physical sculpting is still an awkward and tedious process.

Figure 1: Screenshot illustrating virtual sculpting and showing the modeling environment presented to the user. The target object is located at the center of the screen; on the left the user can select different tools ("brushes") for sculpting. Sculpting is performed by applying the active brush to the target object repeatedly at different locations.

In part, this results from the lack of a physical representation and tactile feedback but in part this can also be attributed to the workflow resulting from the typical combination of input devices, especially regarding effective bimanual interaction. The benefits of bimanual interaction have been demonstrated by numerous studies in the past (Leganchuk et al. 1998; Hinckley et al. 1998a; Hinckley et al. 1998b; Gribnau & Hennessey 1998; Balakrishnan & Hinckley 1999; Buxton & Myers 1986; Guiard 1987; Balakrishnan & Kurtenbach 1999) and indications have been found that bimanual interaction can be superior compared to unimanual interaction not only in terms of efficiency and speed but also in terms of task understanding. However, the same studies also suggest that the benefits of bimanual interaction are highly task and interface dependent. Therefore, in addition to studying general abstracted tasks, it is also important to investigate specific task/input device combinations within the context of their application domain. In this paper, we present our interface design for bimanual VS on interactive surfaces and we report on a case study, we conducted in order to investigate unimanual and bimanual interaction patterns for VS on interactive surfaces compared to the established stylus/pen input devices. The results of our study suggest that interactive surfaces can indeed be used successfully for VS and they have a stronger affordance for bimanual interaction than stylus/pen in this context. Guiard's kinematic chain model (Guiard 1987) for asymmetric division of tasks in bimanual interaction holds for VS regarding modeling and camera control when using multi-touch (MT), although not as strongly as expected. We could also observe that bimanual interaction takes a substantial amount of training and that the stylus is better suited for the manipulation part, which strongly suggests a combination of stylus-based interaction with MT for this application area.

2 Related Work

Different approaches to VS have been proposed in the literature. Early systems allowed the user to draw simple polygonal shapes on the screen that were automatically extruded to 3D and intersected with other shapes to generate the final object (Parent 1977). Other works improved upon the mathematical model, employing different voxel-based representations (Galyean & Hughes 1991; Wang & Kaufman 1995) or models based on implicit functions (Raviv & Elber 2000; Ferley et al. 2001) using different adaptive sampling schemes to extract isosurfaces for rendering. Many systems include typical tools for adding or removing material or smoothing the surface. Some systems work rather directly on polygonal meshes (Bill & Lodha 1994) while others employ constructive solid geometry (Mizuno et al. 1998). Parviainen et al. investigated aiding geometric shapes such as boxes, planes, and lines to improve the user's recognition and understanding of the 3D world in a sculpting system (Parviainen et al. 2004).

Many interface designs and much research in this area is based on studies of Guiard, who investigated interaction patterns for bimanual interaction. He found that even for tasks that superficially might be considered as unimanual tasks, often the non-dominant hand (NDH) supports the actions of the dominant hand (DH) in specific ways, leading to increased overall performance. He called his model the "kinematic chain" (Guiard 1987). As different input devices exhibit different capabilities with respect to bimanual interaction, researchers have tried to formalize and group input devices based on their interaction model (Hinckley et al. 1998a). Other studies investigating the potential benefits of bimanual interaction have found that users can benefit on different levels from bimanual interaction. Depending on the specific scenario, bimanual interaction can lead to increased speed, precision as well as understanding and an improved mental model of the task (Buxton & Myers 1986; Leganchuk et al. 1998; Balakrishnan & Hinckley 1999). Balakrishnan and Kurtenbach explored the benefits of distributing manipulation and camera control between the DH and NDH, using two mice as input devices (Balakrishnan & Kurtenbach 1999).

Frisch et al. investigated using combined MT and pen input for diagram editing. In contrast to our findings they report no clear user preference for choosing pen over MT with regard to specific operations and their participants scarcely employed combined MT and pen interaction (Frisch et al. 2009).

3 Interaction Design and Gestures

Following a user-centric design approach, the interaction design is based on a careful examination of video tutorials for VS in order to identify the most important tools and how they are used. The goals of the video analysis were two-fold: Identifying the fundamental set of tools needed for VS (provided such a set actually exists) and identifying common usage patterns. A corpus of video tutorials from the web was selected for analysis by the following criteria: A focus on VS, i.e., no "mixed" tutorials including other 3D modeling techniques, coverage of the complete modeling process from start to finish by the tutorial, and the level of expertise/quality of the instructor and video tutorial, respectively, judged by taking user ratings into account.

In total 12 video tutorials (approx. 5h:30m) were analyzed. For the analysis all operations related to sculpting (including menu control for changing properties and tools) were counted. Operations/commands related to general window/application control were excluded. The results of the video tutorial analysis showed that the fundamental set of operations needed for VS consisted mainly of only seven actions: application of current tool/menu use (38.27%),

	**One finger** is used for the most common operations, i.e., application of the current tool or selection from the menu.
	**Two fingers** if used on the target object are mapped to the smoothing tool. If used off the object two fingers can be used to pan the view.
 	**Three fingers** can be used to either rotate the view or change the tool size depending on their relative angle. If they approximate a 90-degree angle changing the tool size is selected otherwise the view is rotated.
	**Four fingers** are used to zoom the view.
	**Five fingers** when performing a short tap are used to execute the undo operation.

Table 1: Multi-finger gestures for sculpting

view rotation (21.94%), view zooming (12.33%), smoothing (9.89%), view panning (8.22%), tool/brush size adjustment (2.44%), and undo (1.2%). Other actions comprised less than 1.2% of the total number of actions observed. The application of the current tool and the use of the sculpting menu were counted together because both operations were performed similarly from the user's point of view, i.e., simply touching the tablet surface with the pen, while other operations required additional modifier keys.

If the set of operations is relatively small, multi-finger gestures provide an effective means of mapping general operations to MT gestures (Matejka et al. 2009; Walther-Franks et al. 2011) and they allow for different types of bimanual interaction. In addition to using the number of fingers, we used parameters, such as if the fingers are on the target object or the relative position of fingers to extend the possible number of mappings. Based on the results of the video analysis multi-finger gestures were defined for the set of fundamental sculpting operations (table 1).

4 Case Studies

A user study was conducted to explore interface design issues and usage patterns for bimanual VS with MT displays. The concrete goals of the study were as follows: Demonstrate the principle applicability of MT control to VS. Verify if Guiard's kinematic chain for bimanual interaction applies to VS, possibly identify alternative bimanual/unimanual usage patterns. Compare usage pattern/interface issues to the interface combination of pen and keyboard.

Many existing studies in the field of MT interfaces strive for a "walk up and use" application and evaluation. However, we claim that 3D modeling is inherently difficult and requires some training. Furthermore, several studies suggest that bimanual interaction on MT displays might be an expert technique not naturally performed by novice users (North et al. 2009; Terrenghi et al. 2007; Walther-Franks et al. 2011). Therefore, it was not the goal of the study to demonstrate any kind of naturalness or walk up and use criteria. A case study with fewer subjects but with each subject having more time than usual and exploring the interface together with an instructor therefore seemed a good approach to investigation for this application.

In order to compare MT and pen/keyboard input, participants were asked to try both types of interfaces. The pen/keyboard controls were based on the default settings of the 3D modeling application Blender[1] and simplified to focus on the same fundamental set of operations provided by the video tutorial analysis described above (table 2).

[1] http://www.blender.org

Use tool / menu operations	Stylus tip
Rotate view	Stylus button + stylus tip
Zoom view	Ctrl + stylus button + stylus tip
Smoothing	Stylus eraser
Pan view	Shift + stylus button + stylus tip
Change tool size	F key + stylus tip
Undo operation	Ctrl + Z key

Table 2: Control mappings for stylus/keyboard interface condition. Left: Operation to be executed Right: Stylus/keyboard action required to perform the respective operation

4.1 Setup and Procedure

Four participants (2 male / 2 female) took part in the experiment with an average age of 20.75 years (SD 2.86), all right-handed. Participants had heterogeneous pre-experience regarding MT/pen interaction and 3D modeling. Each participant tested both interface conditions (MT vs. pen/keyboard) for a substantial amount of time (approx. 1.5 hours per participant). Because of the very substantial time investment, participants were paid a small (15 EUR) amount of money as compensation for their time. Participants were encouraged to take breaks whenever they felt exhausted.

For modeling a special version of the open source 3D modeling tool Blender was employed, which was modified by us in order to process MT input according to the control mappings described above. As MT input device a 22” 3M monitor (M2256PW) was used set to its native resolution of 1680x1050 pixels. The monitor also acted as non-interactive display with disabled MT functionality for testing pen input. For the pen input we used a Wacom tablet (Intuos3 A3 wide) along a standard keyboard. Besides the tip, the employed pen featured two buttons and the butt end of the pen could be used like an “eraser”. To avoid confusion both buttons were set to the same functionality for the test.

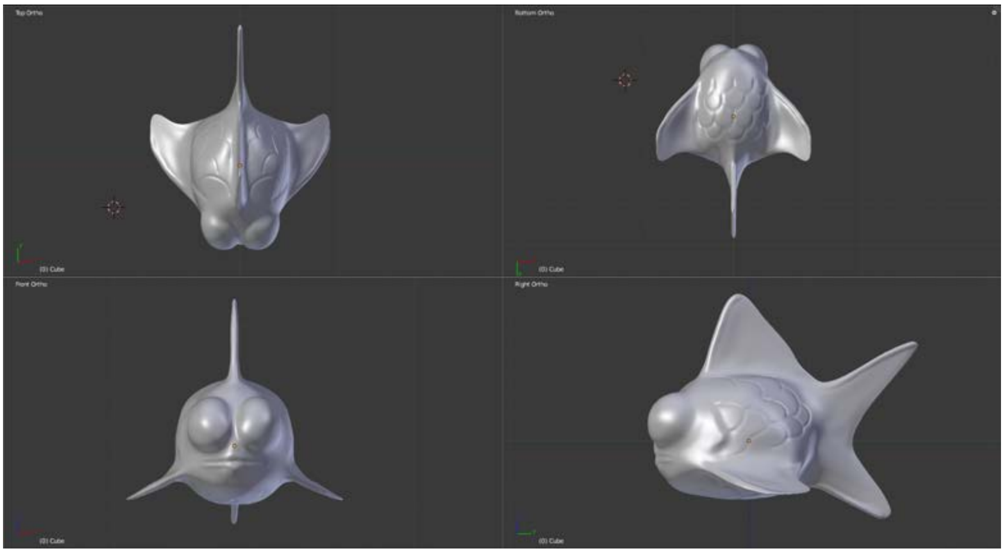

Figure 2: Different views of the target model used to instruct the participants during the experiment.

The complete procedure consisted of the following steps for each interface condition: Participants were introduced to the experiment and demographic data was collected. Then participants were instructed by a short tutorial video (approx. 3 minutes) that demonstrated each control and the basic workflow of VS. Following the video, participants had time to acquaint themselves with the controls and ask questions. When the participants affirmed to feel comfortable with the controls, they were provided with images of different views of a target model (figure 2) and they were asked to re-model it using the particular interface as accurately as possible, starting from scratch. The goal model was taken from one of the tutorials intended for beginners and was chosen because it utilizes the complete fundamental set of operations. Participants were provided with a printed one-page summary of the interface mappings. We considered this step finished, when the participants were satisfied with their created model. During the experiment, participants were asked to adjust the monitor and input devices to their comfort and encouraged to try both a sitting and standing position. The final step consisted of collecting subjective feedback. Participants were asked to fill out the System Usability Scale (SUS) questionnaire followed by a semi-structured interview. We collected video and audio data of each session for further analysis. The order of interface conditions was randomized across participants.

4.2 Results

Participant 1 had experience with mobile MT but no experience with 3D modeling, desktop MT, or pen. The participant was generally able to perform all MT gestures successfully. At first, the participant exclusively used the DH. Later, the participant started using the NDH more often, mostly for asynchronous camera control. Synchronous bimanual control could only be observed a single time. Although the illustrations from the one-page manual (table 1) did not depict any use of the thumb, the participant used the thumb for most gestures. Several times we observed the participant trying to use the "pinch" gesture for zooming although not supported by our system. After the first half of the MT condition, the participant switched from a sitting to a standing position. During the pen interface condition only unimanual control was observed. When asked about it, the participant noted to prefer the pen interface to MT.

Participant 2 had considerable experience with MT on mobile devices but no experience with desktop MT, pen or 3D modeling. For the pen condition, only unimanual control with the DH was observed at first. Later, the participant started using the keyboard with the NDH. Although the pen worked in absolute mode, the participant often tried to use it like a mouse, i.e., lifting the pen several times while panning, which seemed to confuse the participant. The participant interacted mostly unimanually with the DH during the MT condition and made extensive use of the reference sheet. During the experiment, the participant switched from a sitting to a standing position. Towards the first half of the MT condition, the participant started to keep both hands slightly above the surface and often used the NDH asynchronously for camera manipulation, while the DH was used for modeling. At the end of the experiment, the participant started to use both hands for modeling and camera control based on distance to the point of interaction. The participant rated MT to "provide a better feeling for what you are doing".

Participant 3 had some experience with MT on mobile devices but no prior experience with 3D modeling, desktop MT, or pen. The participant almost exclusively interacted unimanually with the DH in both conditions. In general, the participant successfully used all MT gestures. The participant only interacted in a sitting position but changed his posture often during the experiment. Similar problems with using the pen in a mouse-like manner as noted above were observed. The participant rated the pen condition to be a "little less straining" than MT.

Participant 4 had some experience with desktop MT, pen and 3D modeling. In the pen interface condition only unimanual interaction with the DH could be observed. In the MT condition, the participant employed bimanual interaction right from the start, using the NDH strictly for camera manipulation and the DH for everything else. The participant only interacted from a sitting position. While commenting very positively on the bimanual interaction possibilities of the MT interface, pen was rated to be "slightly less straining".

The SUS scores were 24.0 (SD 5.24) for MT and 32.25 (SD 1.92) for pen. The frequency of operations was roughly equal for both conditions (on average approx. 15 actions per minute), MT coming out slightly ahead. Regarding bimanual interaction, we summarize the results as follows: After a learning and acclimatization phase of about 30-40 minutes, all participants occasionally used bimanual interaction in the MT interface condition. One participant with considerable pre-experience demonstrated bimanual interaction right from the start. When observed, bimanual interaction was almost exclusively restricted to asymmetric and asynchronous interaction, using (mostly) the NDH for camera control and the DH for everything else. In the pen condition, only very rare bimanual interaction (holding modifier keys with the NDH) could be observed if at all.

5 Discussion

Regarding our initial goals as stated in section 4, we conclude from the results that we could demonstrate the principal applicability of MT and bimanual interaction to VS. We could confirm results of earlier studies (Terrenghi et al. 2007; North et al. 2009) regarding the great impact of the mouse pre-experience on the user's mental model concerning MT interaction. Because of the case study approach with more time per individual user we could clearly observe an increase in bimanual interaction over time. Guiard's model seems to hold for asymmetric division of labor between the DH and NDH in the case of manipulation and camera control. However, it must be noted that the model is not strong enough to derive heuristics for detection as users still occasionally switch the roles of the DH and NDH or use the hands symmetrically. Regarding the comparison between pen and MT interaction patterns, it must be stated that pen is more established and refined, which we see as one reason why it performed slightly better. The participants preferred the pen for the manipulation part, however, in the current combination with the keyboard, users do not leverage the potential for bimanual interaction compared to MT. We conclude that a combination of pen for manipulation with MT for camera control might be a good combination. In a first iteration of our system, we considered continuous operation switching by lifting and putting down fingers "on-the-fly" but participants initiated gestures in a very distinctive and sharply separated manner, taking their hands completely off the screen between different actions.

6 Conclusion and Future Work

In this paper we presented our interface design to support bimanual VS. We reported results of a user study showing that in principle VS works successfully on interactive surfaces. Our results show that our interface design affords a division of labor according to Guiard's kinematic chain for VS on interactive surfaces, however, bimanual interaction required a substantial learning phase and thus is not well suited for walk up and use scenarios. In comparing MT interaction patterns to pen/stylus interaction, we could show that MT can be easier to understand than pen interaction. While both input devices often seem to be affected by the strong influence of mouse-based mental models. The pen was preferred for the manipulation part, while MT was preferred for camera manipulation. We conclude that VS might benefit from a combination of pen and MT. While our results have been encouraging they have to be verified and extended through additional studies in the future. Furthermore, we would like to extend our system to incorporate simultaneous pen and touch interaction.

References

Balakrishnan, R., & Hinckley, K. (1999). The role of kinesthetic reference frames in two-handed input performance. In *Proc. UIST '99*, (pp. 171–178). New York, NY, USA: ACM.

Balakrishnan, R., & Kurtenbach, G. (1999). Exploring bimanual camera control and object manipulation in 3D graphics interfaces. In *Proc. CHI '99*, (pp. 56–62). New York, NY, USA: ACM.

Bill, J. R., & Lodha, S. K. (1994). Computer sculpting of polygonal models using virtual tools. Tech. rep., Santa Cruz, CA, USA.

Buxton, W., & Myers, B. (1986). A study in two-handed input. In *Proc. CHI '86*, (pp. 321–326). New York, NY, USA: ACM.

Ferley, E., Cani, M.-P., & Gascuel, J.-D. (2001). Resolution adaptive volume sculpting. *Graphical Models*, *63*(6), 459–478.

Frisch, M., Heydekorn, J., & Dachselt, R. (2009). Investigating multi-touch and pen gestures for diagram editing on interactive surfaces. In *Proc. ITS '09,* New York, NY, USA: ACM.

Galyean, T. A., & Hughes, J. F. (1991). Sculpting: an interactive volumetric modeling technique. *Comput. Graph.*, *25*(4), 267–274.

Gribnau, M. W., & Hennessey, J. M. (1998). Comparing single- and two-handed 3D input for a 3D object assembly task. In *Proc. CHI '98*, (pp. 233–234). New York, NY, USA: ACM.

Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor Behaviour*, *19*, 486–517.

Hinckley, K., Czerwinski, M., & Sinclair, M. (1998a). Interaction and modeling techniques for desktop two-handed input. In *Proc. UIST '98*, (pp. 49–58). New York, NY, USA: ACM.

Hinckley, K., Pausch, R., Proffitt, D., & Kassell, N. F. (1998b). Two-handed virtual manipulation. *ACM Trans. Comput.-Hum. Interact.*, *5*(3), 260–302.

Leganchuk, A., Zhai, S., & Buxton, W. (1998). Manual and cognitive benefits of two-handed input: an experimental study. *ACM Trans. Comput.-Hum. Interact.*, *5*(4), 326–359.

Matejka, J., Grossman, T., Lo, J., & Fitzmaurice, G. (2009). The design and evaluation of multi-finger mouse emulation techniques. In *Proc. CHI '09*, (pp. 1073–1082). New York, NY, USA: ACM.

Mizuno, S., Okada, M., & Toriwaki, J.-i. (1998). Virtual sculpting and virtual woodcut printing. *The Visual Computer*, *14*(2), 39–51.

North, C., Dwyer, T., Lee, B., Fisher, D., Isenberg, P., Robertson, G., & Inkpen, K. (2009). *Understanding Multi-touch Manipulation for Surface Computing*, vol. 5727 of *LNCS*, (pp. 236–249). Berlin, Heidelberg: Springer.

Parent, R. E. (1977). A system for sculpting 3-D data. In *Proc. SIGGRAPH '77*, (pp. 138–147). New York, NY, USA: ACM.

Parviainen, J., Sainio, N., & Raisamo, R. (2004). Perceiving tools in 3D sculpting computer human interaction. vol. 3101 of *LNCS*, (pp. 328–337). Berlin, Heidelberg: Springer.

Raviv, A., & Elber, G. (2000). Three-dimensional freeform sculpting via zero sets of scalar trivariate functions. *Computer-Aided Design*, *32*(8-9), 513–526.

Terrenghi, L., Kirk, D., Sellen, A., & Izadi, S. (2007). Affordances for manipulation of physical versus digital media on interactive surfaces. In *Proc. CHI '07*, (pp. 1157–1166). ACM.

Walther-Franks, B., Herrlich, M., & Malaka, R. (2011). A multi-touch system for 3D modelling and animation. In *Proc. Smart Graphics '11*, (pp. 48–59). Berlin, Heidelberg: Springer.

Wang, S. W., & Kaufman, A. E. (1995). Volume sculpting. In *Proc. I3D '95*. New York, NY, USA: ACM