Combining bimanual manipulation and pen-based input for 3D modelling

Pedro Lopes, Daniel Mendes, Bruno Araújo & Joaquim A. Jorge

Department of Information Systems and Computer Science
INESC-ID / IST / Technical University of Lisbon, Portugal

Abstract

Multitouch enabled surfaces can bring advantages to modelling scenarios, in particular if bimanual and pen input can be combined. In this work, we assess the suitability of multitouch interfaces to 3D sketching tasks. We developed a multitouch enabled version of ShapeShop, whereby bimanual gestures allow users to explore the canvas through camera operations while using a pen to sketch. This provides a comfortable setting familiar to most users. Our contribution focuses on comparing the combined approach (bimanual and pen) to the pen-only interface for similar tasks. We conducted the evaluation helped by ten sketching experts who exercised both techniques. Results show that our approach both simplifies workflow and lowers task times, when compared to the pen-only interface, which is what most current sketching applications provide.

Categories and Subject Descriptors (according to ACM CCS): User Interfaces [I.3.6]: Interaction techniques—

1. Introduction

The emergence of tabletop technologies has reached many areas in creative work, ranging from music to architectural modelling. We explore how tabletop surfaces can support the sketching workflow, by providing users with a natural combination of bimanual and pen-based input.

While many previous works show how multitouch enabled surfaces can bring advantages to many application scenarios, our proposal explores how they accommodate 3D sketching tasks. This is because bimanual gestural input allows users to explore the canvas through camera operations, including zoom, pan and rotate. Furthermore, pen input maintains a strict correlation to the users’ preferred sketching device (the "pen"). Additionally, we explore synergies between both input classes, allowing gestures to be performed with either pen or touch (finger) input. For users who execute combinations of non-dominant hand gestures with pen movement, that combination might ultimately feel more natural than pen input alone for non-sketching operations, such as camera manipulation.

Our solution allows both finger-touch and pen input across the tabletop surface, providing separate streams to drawing applications. The hardware adaptations can fit any existing optical tabletop, without requiring installation of complicated apparatus or incurring high costs.

The multitouch and pen-based interaction techniques presented here were implemented and tested in our multitouch version of ShapeShop. We carried out tests with ten expert users, ranging from professional 3D modellers to CAD architects and traditional pencil-and-paper illustrators. Results show that multitouch provides added value in simplifying direct manipulation, when users hold the pen-device in their dominant hand. Furthermore, bimanual input supports fluid interaction schemes where the non-dominant hand is focused on secondary tasks, such as camera operations and interface interactions, while the dominant-hand is reserved for the main task, sketching. This translates to 10% less time spent on camera operations, and 44% less time per camera operation. Expert users provided optimistic feedback suggesting that this combination could improve the sketching workflow for multitouch tabletops.

Throughout this paper we discuss related work that also studied pen-input in combination with multitouch. Next, we detail our solution from two perspectives: technological and interaction. We follow through with an expert evalua-
tion carried out with both professional sketchers and modelers, whose results support our hypothesis regarding the adequacy of the bimanual and pen-input combination. Finally, we present conclusions and point out future goals for the project.

2. Related Work

Research on Sketch based Modelling has primarily focused on providing alternatives to WIMP interfaces (Window, Icon, Menu, Pointing device). These interfaces mainly use pen devices as the (single) input modality. One example, the Teddy system [IMT99] allowed users to create 3D shapes by sketching 2D contours directly on a perspective view.

ShapeShop [SWSJ05] proposed a richer set of modelling operations, camera manipulation and geometrical sub-element edition supported by an hierarchical (BlobTree) structure. However, this increased the number of WIMP-based elements (such as toolbars, lists for shape hierarchy, and so forth). To counter the inadequacy of WIMP in calligraphic interfaces. ShapeShop adopted the CrossY [AG04] selection method and seamlessly integrated the 2D user interface on the drawing canvas, thus enabling gesture interpretation to deal with sketching recognition ambiguities. This solution allows the user to create 3D content, and to control the spatial frame of reference visible on the canvas. However, these modelling operations cannot be performed continuously since only one stylus input device is supported at a time. This does not mimic the way people interact with real paper, where both hands are used simultaneously: one to align the canvas and the other for pencil control.

The Guiard model [Gui87] describes how people interact bimanually in daily tasks, proposing guidelines for human-computer interaction techniques. For bimanual tasks, Guiard identifies different rules and actions for each hand depending on whether it is the preferred (also dominant-hand or DH) or non-preferred (also non-dominant hand, or NDH) user hand. It states that while the DH performs fine movements and manipulates tools, the NDH is used to set the spatial frame of reference and is used for coarse movements. This asymmetric model is observed when users sketch on paper; it also shows that people do not explicitly switch between the definition of the spatial frame of reference and tool manipulation. For modelling tasks, [BK09] proposes a 3D user interface applying these principles to improve camera and object manipulation. Regarding traditional pen based techniques, the Guiard model has been partially adopted to create 3D content in the ILoveSketch system [BBS08]. However, this was done using two devices (a pen-based digitizer with function keys and a keyboard) to explicitly switch modes - from recognizing gestures to controlling the virtual camera or to perform additional modelling operations while sketching.

Bimanual gestures have been researched under the scope of multitouch interfaces. In [WMW09] an agreement over user-defined bimanual gestures (that rank high in familiarity) is proposed to support relevant operations such as selection, panning, zooming and dragging. This gesture-based interaction could be combined with sketch based interfaces as presented in [LGHH08]. However this work relies on explicit switching between modalities (from command to free drawing) requiring the user to press a button. Pen-input, unlike touch-based input, generates a single well-defined contact point, with less occlusion than hands, while affording higher precision and allowing arm and palm rest. Furthermore, [BFW+08] states that bimanual operators are familiar tools that leverage users’ experience. On the other hand, they also stated that touch input does not need to follow the Guiard asymmetric model. Another advantage is its gadget-free nature, which results in immediate actions and requires low user attention; therefore, it suits non-drawing tasks such as manipulation and navigation.

As opposed to pen-only interactions, the combination of touch and pen-input does not require separation of drawing strokes from command strokes; that is provided by the two devices on the same interactive surface. To separate pen data from touch input, Leitner et al. [LPB+09] utilizes an Anoto pen on a layer of tracking paper on top of an optical FTIR tabletop (as proposed by [Han05]). Although this device provides high resolution output, it is a costly solution, that also require modifications to the tabletop surface.

To create a pen sensitive area, [Els09] uses tablet PCs (pen-based input devices with on-surface display) positioned on top of the multitouch surface. Although this allows a higher resolution, it requires movement of heavy components on the interface, plus a complex synchronization of visual content. Furthermore, it creates a physical barrier to fluid interaction, since continuous gestures cannot be performed (such as direct touch across both surfaces).

Several authors strived to combine bimanual and pen-input, although none applied it to 3D sketching tasks. In [FHD10] the authors showed that by combining bimanual and sketch-input, mode switching can be performed via hand gestures without interrupting the drawing workflow. Wu et al. [WSR+06] proposed a combination of pen and touch input that can accommodate actions for a publishing application, such as annotation and copy-paste. The authors strive to maintain the touch gestures as simple as possible, so that the user can still hold the pen device when performing them. Also, [FHD10] and [HYP+10a] suggest that combined gestures (touch and pen simultaneously) for zooming and scaling allow users to choose naturally one approach (either just touch or touch and pen) without interrupting flow of work. Indeed, these actions can be performed with touch (one or two hands) or a combination of NDH and stylus motion. In [HYP+10a] the pen is identified through blob brightness, because the LED-pen is brighter than the touch points, since it is directly pointed at the IR camera. Furthermore, the authors describe several usability issues raised by multitouch.
Our proposal provides a synergistic combination of bimanual operations for sketching, camera control and interface interactions for 3D modelling, which has not yet been covered in any of the previous research, including those on simultaneous multitouch and pen-input for tabletops.

3. Our Approach

In order to observe users interacting with pen and touch in a 3D modelling scenario, we modified the ShapeShop modelling system [SWSJ05]. The original version relies exclusively on a single input interface for creating 3D models based on sketched contours. Editing and transformations are supported by a set of operators, while camera control allows users to navigate the scene and visualize models. Our technique seamlessly combines pen and multitouch input on the same interactive surface. Furthermore, we adapted the modelling system interface to take advantage of the synergies between bimanual gestures and stylus input, to support editing shapes more fluidly than existing approaches.

3.1. Combining touch and pen input

In order to provide stylus-like input on a multitouch tabletop, we devised an IR-laser pen prototype, as depicted in Figure 1, with about the same length as a traditional graphite pencil. At the press of a button, the device emits a narrow IR light beam, which is tracked by an IR-sensitive camera inside the tabletop. This solution can fit any optical tabletop regardless of technology (DI, FTIR or LLP), requiring only an additional camera that operates on a different wavelength from the optical multitouch system. Furthermore, this eliminates the need to alter the tabletop surface - a distinct advantage over Anoto pens [BFW*08], that degrades projection quality and requires additional modifications to the surface.

As our stylus-device operates on a different wavelength, we ensure that pen and finger inputs are separated, as depicted on Figure 2. This provides a higher degree of input separation from touch points, whereas other solutions will not separate touch data from pen data, since they operate within the same wavelength and rely on relative brightness to filter stylus from finger data. Furthermore, having separate inputs for each camera allows us to calibrate the pen without interfering with the touch configuration. Also, the stylus grip is positioned ergonomically, so that users can naturally rest their index finger on the button while holding the pen.

3.2. Graphical User Interface

Figure 3 depicts the revised ShapeShop graphical user interface components (numbered from 1 to 5). Component 1 is the drawing canvas, which allows full use of the whole surface for the core task, sketching. Furthermore, the surface is sensitive to both pen and touch inputs. Components 2, 3 and 4 are control toolbars, which users can rearrange at will by dragging them (via touch or pen) to a new position. In detail, component 2 provides users with shape creation functionalities. This toolbar, once a closed contour has been successfully drawn, shows a set of possible actions; otherwise, it allows deletion of the currently sketched line. Once a shape has been selected, component 3, a contextual toolbar, allows the shape’s depth and width parameters to be configured. Component 4 is the toolbar for camera/view manipulations: zoom, pan and rotate (from top to bottom). This toolbar is only available in the pen approach, since it is replaced by touch gestures in the combined approach, as described further down. Component 5 represents a selected shape from the canvas. Finally, supplementary 3D axis widgets are displayed to allow direct manipulation of the shape.

3.3. Interacting with pen and fingers

While the original version of ShapeShop relies only on a sketch based interface adapted to pen stylus or pen enabled computers, our revised version supports both the IR-based pen device and multitouch gestures. Bimanual gestures and fingers control the scene visualization, as depicted in Figure 4. While two fingers can be used to zoom (Figure 4(b)) and rotate around the 3D scene (Figure 4(c)), four fingers
or more pan the camera view (Figure 4(a)). These gestures control the spatial reference using touch, which is adequate for coarse movements. Since the user can hold the pen during these gestures, the tip of the pen can be used as a finger provided its laser beam is not activated.

The main purpose of the pen is sketching, which allows creating 3D shapes on the canvas. Also, either pen or touch can be used to interact with toolbars, as illustrated in Figure 5. This enables users to seamlessly switch between input sources and provides redundancy for selections.

Table 1 summarizes which operations can be initiated by which modality, in what we describe as the combined approach.

<table>
<thead>
<tr>
<th>Actions</th>
<th>Pen</th>
<th>Touch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sketch</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Select Object</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Move Object</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Camera Gestures</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Activate Widgets</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 1: Summary of combined approach.

4. Expert User Evaluation

We set out to evaluate two situations: issuing all input through a single pen device (pen) vs. using bimanual and pen in combination (combined). Tests were structured in four stages: a pre-test questionnaire to establish user profile regarding sketch expertise; briefing about ShapeShop and training session; three sketching tasks; and finally, a questionnaire to get detailed information about interaction experience.

4.1. Apparatus

All tests were conducted in our customized 1.58x0.87m optical tabletop, with multitouch and pen capabilities, as depicted in Figure 6. Behind the tabletop, a whiteboard displayed the task objectives for the users to consult throughout the test. With the users’ permission, tests were videotaped and comments were transcribed from audio recording.

4.2. Participants

Evaluation was carried out by ten users (five female and five male) from which two are left handed, while the remain-
der are right handed. Six users are professionals (designers, illustrators, 3D animators and modellers) while the other four are semi-professional (still engaged in art courses). All users are familiar with 2D digital sketching and seven are familiarized with 3D software solutions, such as CAD software. None of the testers had any previous contact with ShapeShop. Half the users use digital pen tablets on a regular basis. Regarding multitouch experience, six users have daily contact with multitouch phones, but none has expertise with tabletop surfaces or tablet computers.

4.3. Task Description

We asked users to generate 3D models through sketching and shape manipulation. Furthermore, some operations, such as altering the model’s polygon count, were not required since our focus was on sketching. The test was comprised of three tasks, increasing in complexity. For each task the subject was shown the task objective, a 3D model depicted in different views. These models were created by two 3D modellers (not part of the test group) to better judge its complexity and adjust the task beforehand.

Figure 7 shows the target shapes for each task, while the minimum strokes and operations required to complete the task are detailed in Table 2. When users judged each task was complete they would inform us so verbally.

![Figure 7: Shapes for user testing tasks.](image)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Task A</th>
<th>Task B</th>
<th>Task C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapes</td>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Subtractions</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Camera Operations</td>
<td>0</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 2: Minimum operations for task completion.

5. Result analysis

We present three different perspectives on the analysis of the expert user evaluation data that, when combined, validate our hypothesis that a combination of bimanual and pen-input will generate less interruptions on the 3D sketching workflow. Firstly, a quantitative analysis gives us a qualitative measure on how users felt on each approach. Finally, video analysis allowed us to conduct and discuss several observations made throughout the test sessions, that help us validate our hypothesis.

5.1. Quantitative analysis

For each task we accounted the total time needed to create the target 3D shape. We also kept partial scores on elapsed time: sketching (when the user was drawing on canvas); camera (rotate, zoom or panning time); and idle (when the user was performing operations not included in the essential moves required for task completion, such as rearranging the interface, translating shapes and so forth).

An Anderson-Darling Test suggests that all distributions are possibly normal (p-value > 0.05, for each task in either approach). Furthermore, a T-test applied to Task B and Task C suggests a statistically significant difference between the results of combined and pen approach (t(17.43)=1.93, p<0.05 for Task B; t(16.02)=1.97, p<0.05 for Task C). As for Task A the t-test shows no significant difference for any approach (t(9)=1.03, p>0.05). This may be due to its simplicity and zero requirements for camera operations.

In order to validate our hypothesis that the combined approach enables a more fluid interaction for camera operations, a more detailed analysis is required. Figure 8 depicts the average time dedicated to sketching, camera and idle, in relation to the total task time. Figure 8 demonstrates that time spent in camera in the combined approach (blue line) decreased by 10% when compared to the pen approach (pink line). Conversely, time spent in sketching in the combined approach (blue line) increased by 10% when compared to the pen approach (pink line). Both results suggests that users are more proficient with bimanual manipulations for camera operations, allowing them less interruptions to the sketching task, which is considered the main goal.

![Figure 8: Average of elapsed time in each operation for all user tasks (percentage of total task time).](image)
pen approach. This improvement may be explained by subjects’ preference for direct manipulation and taking advantage from it. An additional explanation is that subjects do not lose much time in mode switching - e.g.: locating and acquiring the camera widgets floating bar.

Figure 9: Average time spent in camera (in seconds).

Figure 10 allows us to understand how users took advantage of bimanual support. It shows the distribution of DH, NDH or Both Hands for Pan and Rotate operations in combined approach. The results for combined approach seem optimistic: 91% of the rotation gestures was performed with NDH, while only 9% used the DH. These 9% also suggests that the rotate gesture can seem frail for some users, that felt the need to drop the pen (or swap it to the NDH) to execute the rotation. The pan gesture was performed 78% of the time using NDH and 17% with both hands, showing that some users appreciate moving the canvas with both hands, without needing to drop the pen device (no occurrences). The remaining 3%, which is not significant, executed the pan with the DH only, which suggests there is no preference for issuing this gesture with the preferred hand while holding the pen.

Figure 10: Distribution of hand preference (DH, NDH or both) when performing rotate and pan gestures.

The zoom command is not included in the last graph, since it requires both hands to be performed under the combined approach. Zoom accounted for 37% of pen-touch gestures, which indicates that subjects zoom using one finger from NDH and the pen tip on the DH (as depicted in Figure 4(b)), also showing that some users took advantage of using the pen tip as a touch point.

5.2. Qualitative analysis

Figure 11 shows a summary of users’ responses to our questionnaire. In Figure 11(a), the users’ preferences regarding which input (touching with fingers or pen) is preferred for all actions, is depicted. Regarding interacting with the toolbars (for choosing shape operations), selecting and translating objects in the canvas, 70% of the users preferred touching them directly.

As far as camera operations are concerned, the users manifested a clear preference for bimanual interaction through touch gestures. Indeed, 90% preferred touch for overall camera manipulation, rather than relying on the camera widgets via pen interaction. Looking in detail to each camera manipulation allows us to conclude on the adequacy of our design choices for each gesture. For pan all users agreed that touch gesture (with either one or two hands) is preferable for panning. Furthermore, users described panning as a solid gesture, that allowed them to manipulate the canvas with the NDH hand while looking for the best place to start sketching with the DH; other users said that they felt comfortable performing it with both hands (still holding the pen) because it gave them a stable feeling while dragging the canvas. For the zoom operation, 80% of the users preferred the bimanual gesture (one finger of each hand) over using the pen on the appropriate widget. Also, all users (even those that preferred pen) stated that they felt comfortable performing the zoom without the need to put down the pen. Last, for rotate, 70% of the subjects still preferred touch input (two finger gesture), but, as commentaries suggested, this is the most error-prone gesture and subjects seemingly did not feel comfortable performing it while holding the pen in the DH.

When asked which approach was preferable for overall actions, all subjects were keen to stress the advantages of the combined approach, as depicted in Figure 11(b). Furthermore, 90% of them stated that the combined approach suits their workflow better, providing for a more fluid task, with less interruptions to sketching. This supports our hypothesis that combining bimanual and pen input eases interaction when modelling 3D shapes in a tabletop scenario. Finally, 70% of the subjects considered the pen approach
more stressful and physically demanding, since it often re-
quired more interactions with widgets and rearranging the
interface to place the toolbars nearer to both the shape being
edited and the DH.

5.3. Observations

Figure 12 depicts a subject engaged in the combined ap-
proach. In Figure 12(a) we can observe how subjects natu-
rally use the NDH to interact with the shape creation toolbar.
Furthermore, as Figure 12(b) depicts, even when the NDH is
at rest, it stays closer to the toolbars than the DH, which is
engaged in sketching. Overall this allows for a more fluid
experience, and ultimately could also reduce task duration.

We observed throughout the tests that most users found a
natural pose for executing the camera gestures (pan, rotate
and zoom) without the need to drop the pen, or swapping it
to the NDH, as depicted in Figure 13(a). Furthermore, some
users performed gestures using pen and fingers simultane-
ously, namely the zoom, as depicted in Figure 13(b) without
being prompted to do so.

The rotate gesture showed lesser results, being the only
gesture that drove a few users to swapping the pen to the
NDH, so they could free the DH for the rotation gesture, as
depicted in Figure 14(a). This suggests that for this scenario,
two-finger one-handed gestures, as depicted in Figure 14(b),
are not stable or adequate for the DH while holding the pen
device.

6. Conclusions

Throughout our research we devised a pen device that allows
separate input from finger touches in tabletop scenarios. We
evaluated a modified version of ShapeShop, with multitouch
and pen support, with a panel of expert users. From test re-
results we can conclude that when using both bimanual and
pen-input users tend to work more fluidly, with less interrup-
tions due to the need of manipulating the UI with pen-input.
Multitouch input also proved to be of value to the 3D sketch-
ing context, suiting camera/view manipulations with less in-
terruption time of the sketching task, which in this context
is clearly the main objective. Furthermore, we see that our
conclusions are aligned with Guiard’s bimanual model, be-
cause, we observed that users exhibit a natural tendency to
navigate around the canvas through coarse motions of the
NDH, while keeping the DH focused on the sketching task.
Finally, we can conclude that our gestures minimize inter-
ruptions to the sketching workflow.

© The Eurographics Association 2011.
7. Future Work

The exploration of additional devices can open new dimensions to this research, such as hand pose while gripping the pen [SBG*11], rolling the pen [BMR*08] and pressure sensitivity or haptic feedback [WKK*10]. The expert commentaries allow us to plan further alterations of ShapeShop that will need validation, such as placing contextual menus in key locations of the interface as the work evolves, either near the last pen stroke for the pen approach, or near the user’s NDH for the combined technique. We believe that such small changes which anticipate either changes in focus of attention or steps in the workflow may be key to further gains in fluidity.

8. Acknowledgements

This research was partly supported by FCT through research grant MIVIS PTDC/EIA/EIA/104031/2008 and CEDAR PTDC/EIA-EIA/116070/2009, which we gratefully acknowledge.

A word of appreciation to Ryan Schmidt for the original code of Shape Shop. Lastly, we kindly thank all users and modellers for their commitment in helping our research.

References


[Han05] HAN J. Y.: Low-cost multi-touch sensing through frustrated total internal reflection. In Proceedings of the 18th annual ACM symposium on User interface software and technology (New York, NY, USA, 2005), UIST ’05, ACM, pp. 115–118. 2


© The Eurographics Association 2011.