EUROGRAPHICS Symposium on Sketch-Based Interfaces and Modeling (2009)
C. Grimm and J. J. LaViola Jr. (Editors)

Multi-touch Focus+Context Sketch-based Interaction

Uwe Hahne,[1] Jonas Schild[2], Stefan Elstner[1] and Marc Alexa[1]

[1]Computer Graphics, TU Berlin, Germany
[2]Entertainment Computing Group, University of Duisburg-Essen, Germany

Abstract

In this paper we present a Focus+Context screen for combined pen and touch interaction. A tabletop display presents contextual information and enables multi-touch detection for navigation through frustrated total internal reflection (FTIR). A high resolution pen enabled display is continuously localized on the tabletop. It is used as a movable focus display and pen input device. The resulting system has a large virtual resolution for both display and pen interaction. We demonstrate the advantages of Focus+Context for pen-based interaction in combination with multi-touch navigation in applications that make use of sketch-based interfaces.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Sketch-based interfaces—multi-touch interaction

1. Introduction

As pointed out by Olsen et al. [OSSJ09], recent sketching interfaces lack natural behavior – the "everybody can draw" paradigm is only valid for pen and paper interaction allowing users to manipulate details with high spatial accuracy. On the other hand, interaction with the now common large models or data sets benefits from large display devices providing contextual information and intuitive navigation at coarse scales. In this work we describe how to mimic a workspace that successfully combines these requirements, inspired by the human working style of placing or moving a piece of paper on a desk: the desk is modeled as a large multi-touch table with the paper being a commercially available pen tablet display.

Multi-touch sensing is based on frustrated total internal reflection (FTIR), offering natural interaction metaphors for coarse scale navigation [Han06]. Pen interaction uses electro-magnetic resonance (EMR), enabling accurate interaction and sketching. Both types of interaction devices are scalable to a certain degree, however, larger size and/or higher resolution require significantly more elaborate and expensive technology (e.g., high-resolution projectors, tiled displays, large capacitive arrays).

The combination of these technologies in this configuration yields an interactive Focus+Context system that offers users both, visualization and manipulation of details as well

Figure 1: *A user interacting with the pen display and table.*

as an overview of the data and navigation using the (non-dominating) hand. Brandl et al. [BFW*08] show that the combination of pen and multi-touch leads to very intuitive application design. Moreover, the Focus (=pen tablet) being free to move allows the user to bring objects into focus by either moving the tablet display or panning the content of the Context (=projection of the multi-touch table). This leads to a virtual display with a resolution of over 24 megapixels that is capable of pen input everywhere on the large tabletop. It also addresses the problem of self-disclosure [LaV07]: it is preferable that the workspace for sketching is empty. In our setup we can use the Context region to display annotations and keep the workspace free from disturbing menus. Figure 1 shows a user working with our setup.

Our multi-touch display uses an extension of FTIR that makes it easy to track the pen tablet (the setup is described in section 3): active infrared LEDs on the backside of the pen tablet can be tracked using the camera of the multi-touch table. Pulsed illumination of the panel in the touch table makes it easy to distinguish touch events from the pen display. We have developed several example applications, demonstrating the opportunities of interactive and movable Focus+Context systems (see section 4).

2. Related work

Our system design creates an intuitive tabletop workspace by combining several other approaches. One may classify these approaches as follows: first, our setup is an implementation of a classic magic [BSP*93] or sigma lens [PA08] system, however, with the lens also being an interaction device through the pen input. Second, multi-touch interaction combines natural bi-manual and pen input in an intuitive way. And third, the physical setup is similar to Focus+Context screens, with the addition that the Focus is physically movable.

There are several recent setups combining pen and multi-touch interaction. C-Slate [IAC*07] uses stereo vision to enhance a pen-based display with multi-touch sensing. This is in a sense inverse to our approach, only that pen-based displays are difficult to increase to the size of projection screens. Another system combining pen and multi-touch interaction is FLUX [LPB*09]. It is based on the Anoto™ technology for pen interaction and needs a special screen layer printed with the Anoto™ pattern on the panel, which can still use FTIR for detecting touch events. The first commercially available sensor combining capacitive multi-touch with a digitizer pen sensor is N-Trigs DuoSense™ technology. While DuoSense™ relies on palm rejection algorithms to nullify touch interference from pen interactions, our system separates the two technologies to Focus and Context physically. This separation allows us to work at a much higher virtual resolution and hence allowing more precise input.

From the many Focus+Context hardware setups [BGS01, AR05, UI97, FB04, SH06, BHB07] that mainly follow the ideas of Furnas and Bederson [FB95] about multi-scale interfaces, we want to focus on Ubiquitous Graphics, developed by Sanneblad et al. [SH06]. Here, a tablet PC can be held in front of a wall sized projection area. The position of the tablet PC is determined through ultrasonic tracking to show a high resolution image of the part being covered. From their users' reviews and own experiences with the Wacom pen display that it is cumbersome holding such a device all the time, we came up with our tabletop solution including touch interaction in the Context. For further information, we refer to Cockburn et al. [CKB08] who give a great overview. They show that Focus+Context setups do support the user by addressing its peripheral vision and further outperform zooming interfaces on tasks that require a quick overview.

In our Focus+Context setup we use two independent displays. The multi-touch table displays Context information while the pen tablet on its surface displays the Focus region in higher resolution. This overcomes many of the disadvantages of previous approaches:

- Direct input is possible in the Context as well as in the Focus.
- The Context, displaying an overview without the need of precise input, allows multi-touch input for navigation.
- The pen display is capable of highly accurate pen input that is needed when working on details like in selection and manipulation tasks.

In contrast to many of the previously mentioned works the pen display and thus the Focus can be moved around in the Context area. Moreover, there is no need for distortion or blending, as introduced by fish-eye views [Fur86] and scales and dimensions are kept between Focus and Context.

3. Setup

Figure 2: *The multi-touch table*

3.1. Context: Multi-touch Table

Our setup is designed as an extension of an FTIR based multi-touch table following Han [Han05]. Standard equipment includes an ultra short throw projector and a camera with fish-eye lens, both chosen to avoid mirrors, making the setup more robust. The projector has a resolution of 1024 by 768 while the camera works at VGA resolution (640 by 480). The light inside the panel is generated with LEDs at 850nm (near infrared) and the camera is equipped with a bandpass filter for the same wavelength.

The LEDs illuminating the panel can be switched. Alexa et al. [ABB*08] continuously alternate between images with and without panel illumination and use the images without FTIR effect as references for the suppression of ambient light. We have generalized this approach and can control the illumination in the panel, i.e. generate reference frames only when necessary.

We also use the on-demand reference images for easy and robust detection of the pen tablet: the tablet is identified based on IR LEDs that are on also when the panel illumination is turned off, making it easy to distinguish touch events and the contact points of the panel. The components and procedures are detailed in the following.

3.2. Focus: Pen display

We use the Wacom Cintiq 12WX, a pen display with a native visual resolution of 1280 by 800 pixels and and a pen touch resolution of 5080 lines per inch at 1024 pressure levels. It measures 40.5 by 27 cm, has a thickness of 1.7 cm and weighs about two kilograms, making it uncomfortable to lift. The dimensions of the Wacom pen display lead to a virtual resolution of the entire system of 5885 by 4417 pixels. The screen is surrounded by a border with a width of approximately 7 cm on the left and right side and approximately 2 cm on the top and 8.3 cm on the bottom. This border interferes with the seamless integration of Focus and Context, however, it simplifies moving the pen display and allows users to rest their hands for using the pen. A single cable connects the pen display to a base station. We chose this pen display for its precise and low latency input, the display resolution, and the single cable, which is less obtrusive than other systems using different connections for input, output, and power supply. We have glued plates of felt to the rear side of the pen display to prevent the multi-touch table from damage and support smooth motions.

In order to render the Focus on the Wacom display its position and orientation relative to the table has to be determined continuously. It is natural to use the camera in the multi-touch table for this task. A straightforward approach is based on tracking the contact points on the touch table similar to other objects (mostly fingers) touching the table. We have found several minor practical difficulties: because of imbalance in the pressure put on the pen display most often only three out of four (for stability reasons) contact points are detectable. Matching these three points to the pen display (i.e. discriminating them from user input) requires algorithmic effort, which increases because of the uncertainty which contact points are visible. We designed our setup enabling a simpler active solution. For this we attach three infrared LEDs to the pen display, shown in Figure 3. These LEDs are powered continuously. Next we explain how they can be easily detected.

(a) top

(b) bottom (active marker magnified)

Figure 3: *Wacom Cintiq 12WX Pen Display and active markers on the bottom.*

3.3. Active tracking of the Focus tablet

As described before, the LEDs illuminating the touch panel are switched on and off, providing blobs for the touch input in one frame and a reference image of ambient light in the next. The LEDs mounted to the pen display are clearly visible in both frames. Thus, differencing two frames not only identifies touch points, but also the contact points, namely as very bright spots that are removed by taking the difference.

The distances between each pair of contact points of the tablet (containing the LEDs) are mutually different. Thus, they form a triangle A, B, C with normalized side lengths a, b, c with $a + b + c = 1$. Measuring the distances between the bright spots in the (undistorted) camera image allows matching the edges based on their relative lengths. Matched edges yield matches between vertices. The 2D positions of the vertices in the camera image (and the knowledge of their positions on the tablet) immediately yield the necessary coordinate transformation for the tablet. Note that two points would be sufficient for computing the coordinate transformation, but distinguishing the vertices would be impossible without further technical setup, thus only specifying the transformation up to a rotation by 180 degrees.

4. Applications

Applications using the Focus+Context setup need to drive two independent displays. To ease the development of applications, we designed and implemented a compact library that handles the use of multiple displays. It mainly provides the viewport for the Focus relative to the Context. For 3D applications the view frustum of the Context has to be computed. Since this depends on the projection matrix, which is specific to the application, minimal communication between the library and the application cannot be avoided.

In the following, we present three applications that have been created or customized for our new setup, showing both 2D and 3D visualization and interaction.

4.1. Proof of concept: Fish tank

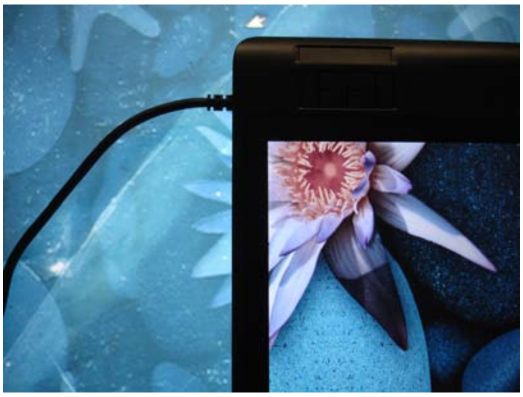


Figure 4: *Detail view of Fish tank application*

A fish tank visualization has been implemented as a proof of concept. It combines the magic lens and Focus+Context approach. Users see a fish tank including a small fish from above. They can interact with the environment by tipping on the surface, virtually dropping food and attracting the fish. In order to get a clear view of the regions under water, the user has to move the pen display and look *through* it. Below the water plane lies a ground plane which is textured with an image of 4096 by 4096 pixels. While it is visible in the Context, it is necessarily scaled down and thus only visible with low detail. The pen display, the Focus, shows life under water in more detail. Since the water surface is not displayed, the pen display can also be interpreted as a magic lens as it displays different information. Interaction in the Focus is similar to the Context. Through the display, the fish is seen enlarged since the point of view sits under the water plane. The virtual resolution of the system is even higher than the size of the ground texture. As can be seen in Figure 4 the difference in detail between Focus and Context is obvious.

The application demonstrates the following objectives of the Focus+Context setup:

- It shows that the Focus environment provides more detail than the Context.
- A different view of the fish tank is presented in the Context and Focus displays.
- Every part of the Context environment can be viewed in full detail by just moving or rotating the pen display.

This application is not meant to demonstrate novel interaction techniques with multi-touch and thus, not making any use of the multi-touch capabilities beside localizing the pen display.

4.2. Map exploration

Figure 5: *GoogleEarth™ application*

To describe the way of exploring geographical data in our setup, we developed a small GoogleEarth™ application. On one hand, we implemented a standard multi-touch application for panning and zooming the earth map on the Context surface. On the other hand, for the Focus display a second virtual camera is set that displays exactly that part of the globe lying under the tablet (see Figure 5). This Focus part is not only displayed in a higher resolution and thus containing more details, it is additionally possible to display extra information like infrastructure or 3D models only in this part while keeping the Context area less textured as it is only recognized in the peripheral view of the user. Otherwise, moving or blinking objects in the Context could distract users from their center of attention in order to alert them.

Exploring maps in this application users can either pan the map by touching the surface with their finger or moving the pen display to any intended position on the table. While rotating the pen display has no impact on the Context, the Focus region is redrawn. Using the two finger pinch gesture on the surface rotates and zooms the Context maps and hence the Focus is again adapted respectively. On the pen display the user is able to set landmarks or draw sketches in order to highlight specific regions on the map.

4.3. Sketching

In order to demonstrate the interaction possibilities of our setup, we adapted the sketch-based 3D modeling application FiberMesh [NISA07]. FiberMesh allows users to create freeform 3D models from scratch. To do this, a number of so called control curves have to be sketched in 2D. These control curves define the model and stay on the model to serve as handles which can be pulled later to change the geometry. The application has been customized to be controlled by multi-touch and pen input. Figure 6 depicts FiberMesh as Focus+Context application.

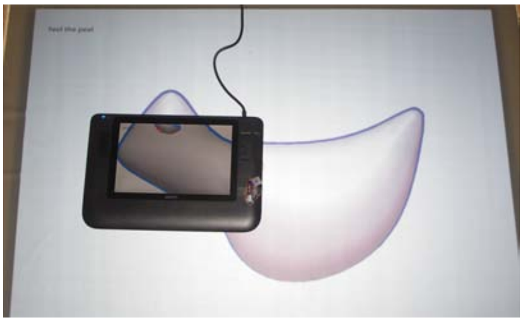
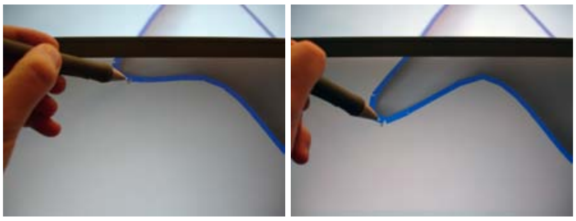


Figure 6: *FiberMesh displaying a curve mesh*

The Context displays the overall model. Users navigate by multi-touch interaction with their fingers. Operations include panning, zooming, rotation around a user defined axis and rotation around the normal vector of the table's surface. Panning can be accomplished by simply using one finger. Since one finger may also be used to sketch, the panning tool has to be selected first. The selection is realized with a simple pie-menu like widget. We decided to implement mode switching because it happens rarely and the advantage of drawing and sketching with one finger overcomes the drawback of utilizing a menu instead of a gesture. In contrast, zooming and rotation is done with two touch points and always possible.

As the Focus display is usually placed in the user's center of attention, the panning has to be performed in the periphery of the user's hand's scope. It is not possible to pull an object directly into the Focus region. Informal user studies have shown that this is no problem as most users are used to perform panning tasks indirectly as with the mouse. Rotating the object around an axis defined by two fingers also turned out to be very practicable in the Context region while zooming and rotating around the normal vector of the table's surface is preferably done without any offset, hence most users use both hands for zooming.

The pen can be used to create precise curves, in contrast to sketching with the fingers, which is less accurate. Additionally, existing curves can be modified only on the pen display. To do this, the pen has to be put directly on a curve, grabbing the curve handle, as shown in Figure 7a. Then, the handle can be moved into the desired direction to alter the geometry of the curve and the shape. Figure 7b shows the resulting deformation. Carrying out extrusions and cuts is also possible by drawing strokes and gestures with the pen.



(a) Before (b) After

Figure 7: *Pulling curve handles on the pen display*

Our adapted sketch-based 3D modeling application benefits in several ways from the Focus+Context environment. First, navigation is made intuitive using the capabilities of multi-touch input. In contrast to mouse input in a standard desktop environment, users can easily rotate and zoom without clicking into specific areas or pressing hotkeys. In comparison to working with a single pen display alone, the navigation becomes much more similar to natural pen and paper interaction.

Second, sketching curves directly with a finger or pen is much more intuitive than sketching with an indirect input device as the mouse. Using a sketch pad, for example, allows also intuitive sketching using a pen, but then navigation has to be initiated by using buttons or widgets. With multi-touch, this disadvantage can be overcome. While manipulating a thin curve handle with a fingertip on the multi-touch table is very hard, this can be done easily using the precise pen input. Furthermore, pen interaction is already possible when its tip is not actually touching the display. This derives from the hover capabilities of the Wacom display that recognizes the pen tip being in proximity to the display. Whether the tip is touching the surface or not is distinguished with high accuracy by the pen display, allowing the application to give valuable visual feedback to the user. In the case of dragging an existing curve, the curve is highlighted when the pen is in proximity to the display without touch.

Third, handling large models can now easily be done by moving the pen display on the surface of the table. The users get an overall impression of the object while sketching details. This is harder to achieve using a single monitor setup or a sketch pad. In addition working with fingers and a pen is quite natural as the non-dominating hand controls the position and orientation of the object, and the dominating hand holds the pen in order to work on the object directly. This behavior is common for working with pen and paper.

5. Conclusions

The objective of our work was to create a workspace for sketch-based interaction mimicking the way humans place a piece of paper on a desk. We created a Focus+Context environment by combining a pen enabled display with a multi-touch table. Therefore the pen display was tracked on the surface of the table. Several prototypes have been presented showing objectives of sketching in Focus+Context environments. We implemented a simple fish-tank application as a proof of concept. An informal user study showed that the two displays are perceived as a single screen with an arbitrary region of high resolution. GoogleEarth™ and the 3D sketching application FiberMesh have been customized to handle the Focus+Context display. For FiberMesh several navigation techniques have been implemented, exploiting multi-touch. The Context screen showed an overall view of the model, while sketching and curve deformation could be done on the pen display.

This Focus+Context system overcomes many drawbacks that came along with previous systems. Neither distortion nor blending was needed to coat the transition between Focus and Context area. User interaction is possible on the Context display as well as on the Focus display. Furthermore, the Focus can be moved around intuitively, maintaining a direct visual connection to the Context.

5.1. Future Work

Expanding the multi-touch table with direct illumination (DI) seems to offer more possibilities for object detection. Image based markers could be used to track the pen display and other objects as well. Using an arbitrary TFT panel instead of the pen display could offer an even less visible transition between Focus and Context while keeping the high resolution of the Focus. Touch input on such a panel might be achieved using the either existing FTIR or DI techniques or EMR sensors as the Wacom pen display.

Network transmission of the Focus image or distributed computing could be used to get rid of cables to the Focus display. Therefore, a tablet PC, Smartphone, iPhone or any other mobile and wireless device could be tracked on the touch table. Its occupying region on the table could be sent to the device over WLAN or similar network interfaces. In turn, pen and touch events could be sent back to the application computer. A less heavy weighted Focus display could be used as an additional navigation input device. Tilting and lifting the device could control a virtual camera.

We further think about extending the setup to a multi user setup where several users share a common context. The pen displays already physically solve the problem of mutual exclusions for shared workspaces. Here, the user is forced to move the pen display in order to change his personal Focus. Hence, we are planning a user study in order to find out when users prefer virtual navigation via multi-touch and when physical navigation by moving the pen display is more suitable.

References

[ABB*08] ALEXA M., BOLLENSDORFF B., BRESSLER I., ELSTNER S., HAHNE U., KETTLITZ N., LINDOW N., LUBKOLL R., RICHTER R., STRIPF C., SZCZEPANSKI S., WESSEL K., ZANDER C.: Continuous reference images for ftir touch sensing. In *SIGGRAPH Posters* (2008), p. 49.

[AR05] ASHDOWN M., ROBINSON P.: Escritoire: A personal projected display. *IEEE MultiMedia 12*, 1 (2005), 34–42.

[BFW*08] BRANDL P., FORLINES C., WIGDOR D., HALLER M., SHEN C.: Combining and measuring the benefits of bimanual pen and direct-touch interaction on horizontal interfaces. In *AVI '08: Proceedings of the working conference on Advanced visual interfaces* (New York, NY, USA, 2008), ACM, pp. 154–161.

[BGS01] BAUDISCH P., GOOD N., STEWART P.: Focus plus context screens: combining display technology with visualization techniques. In *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology* (New York, NY, USA, 2001), ACM, pp. 31–40.

[BHB07] BORING S., HILLIGES O., BUTZ A.: A wall-sized focus plus context display. In *PERCOM '07: Proceedings of the Fifth IEEE International Conference on Pervasive Computing and Communications* (Washington, DC, USA, 2007), IEEE Computer Society, pp. 161–170.

[BSP*93] BIER E. A., STONE M. C., PIER K., BUXTON W., DEROSE T. D.: Toolglass and magic lenses: the see-through interface. In *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1993), ACM, pp. 73–80.

[CKB08] COCKBURN A., KARLSON A., BEDERSON B. B.: A review of overview+detail, zooming, and focus+context interfaces. *ACM Comput. Surv. 41*, 1 (2008), 1–31.

[FB95] FURNAS G. W., BEDERSON B. B.: Space-scale diagrams: understanding multiscale interfaces. In *CHI '95: Proceedings of the SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 1995), ACM Press/Addison-Wesley Publishing Co., pp. 234–241.

[FB04] FLIDER M. J., BAILEY B. P.: An evaluation of techniques for controlling focus+context screens. In *GI '04: Proceedings of Graphics Interface 2004* (School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2004), Canadian Human-Computer Communications Society, pp. 135–144.

[Fur86] FURNAS G. W.: Generalized fisheye views. In *Conference Proceedings* (New York, NY, USA, 1986), vol. 4-17, ACM, pp. 16–23.

[Han05] HAN J. Y.: 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* (New York, NY, USA, 2005), ACM, pp. 115–118.

[Han06] HAN J. Y.: Multi-touch interaction wall. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Emerging technologies* (New York, NY, USA, 2006), ACM, p. 25.

[IAC*07] IZADI S., AGARWAL A., CRIMINISI A., WINN J., BLAKE A., FITZGIBBON A.: C-slate: A multi-touch and object recognition system for remote collaboration using horizontal surfaces. *Horizontal Interactive Human-Computer Systems, 2007. TABLETOP '07. Second Annual IEEE International Workshop on* (Oct. 2007), 3–10.

[LaV07] LAVIOLA JR. J. J.: Sketching and gestures 101. In *SIGGRAPH '07: ACM SIGGRAPH 2007 courses* (New York, NY, USA, 2007), ACM, p. 2.

[LPB*09] LEITNER J., POWELL J., BRANDL P., SEIFRIED T., HALLER M., DORAY B., TO P.: Flux Ű a tilting multi-touch and pen based surface. In *ACM CHI Interactivity* (2009). [in press].

[NISA07] NEALEN A., IGARASHI T., SORKINE O., ALEXA M.: Fibermesh: designing freeform surfaces with 3d curves. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), ACM, p. 41.

[OSSJ09] OLSEN L., SAMAVATI F. F., SOUSA M. C., JORGE J. A.: Technical section: Sketch-based modeling: A survey. *Comput. Graph. 33*, 1 (2009), 85–103.

[PA08] PIETRIGA E., APPERT C.: Sigma lenses: focus-context transitions combining space, time and translucence. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2008), ACM, pp. 1343–1352.

[SH06] SANNEBLAD J., HOLMQUIST L. E.: Ubiquitous graphics: combining hand-held and wall-size displays to interact with large images. In *AVI '06: Proceedings of the working conference on Advanced visual interfaces* (New York, NY, USA, 2006), ACM, pp. 373–377.

[UI97] ULLMER B., ISHII H.: The metadesk: models and prototypes for tangible user interfaces. In *UIST '97: Proceedings of the 10th annual ACM symposium on User interface software and technology* (New York, NY, USA, 1997), ACM, pp. 223–232.