A Wall-sized Focus plus Context Display

Sebastian Boring, Otmar Hilliges, Andreas Butz
University of Munich, Media Informatics, Amalienstrasse 17, 80333 Munich, Germany
{sebastian.boring, otmar.hilliges, andreas.butz}@ifi.lmu.de

Abstract

In this paper we present a wall-sized input system with high accuracy input in the center and lower precision tracking on the remaining parts of the wall. This work complements the concept of focus plus context displays where output quality in the center is better than in the periphery. Our contribution is to realize a similar concept for the input direction. A high precision tracking unit in the center provides a workspace with high interactivity, while four additional cameras placed in the corners of the wall provide lower precision tracking and hence much coarser interactivity on the entire wall. In our prototype this multi-precision input solution is combined with a focus plus context display, thus providing two levels of interactivity in both the input and output direction. With this setup, we have implemented two example applications to demonstrate the benefits of variable precision input.

1. Introduction

Large display technologies become increasingly available followed by a cumulative need for input solutions which scale to larger areas as well. While the use of very large displays in advertising is commonplace today, research investigates their use as public information displays [1] or as shared interaction spaces [22]. Particularly in co-located collaboration settings, interaction is a key requirement for their efficient usage and is hence central issue.

Large displays allow users – individuals and groups alike – to organize large information sets and many media objects in such a way that they are always visible. Distributing digital objects in 2D space on a large display is an effective means to organize work [15][20]. A typical work setup on traditional tables consists of a central area where users manipulate documents and a peripheral area where documents are stored in the meantime [25] (see Figure 1). Moving documents from the center to the periphery and back indicates in a collaborative process what item is currently worked on. Having the other items visible and easily accessible in the surrounding space eases the work process. Similar approaches can be observed when people work on interactive walls. Tasks in which several independent objects need to be coordinated (e.g. juggling) are handled by time-sharing of attention or, in other words, rapidly switching the focus between objects. In an environment with large displays, this switching can be done by moving along the display.

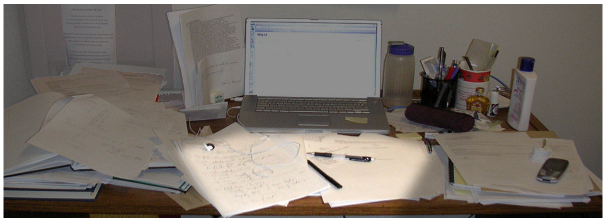

Figure 1. A typical traditional desktop arrangement. The light area in the center denotes the working area.

Baudisch et al. [3] present a focus plus context display to support the perceptual phenomenon of focus on a technical level. It contains a detailed display in a central focus area, while providing coarser displays in the outer context area. This scheme is very adequate for human perception while at the same time preserving technical resources, such as rendering power.

While Baudisch's focus plus context display is uniform regarding input, because it simply uses mouse and keyboard, we present a large focus plus context display, which has interactive areas with varying input accuracy. We argue that this systematic extension of the original concept can easily be applied to large interactive surfaces, such as wall displays. Another aspect of Baudisch's work is that it uses only one input position at a time. The considerable size of wall-sized displays allows multiple users standing in front of them. To allow collaborative work on such displays, we present a system that is capable of tracking multiple fingers – in the focus region as well as in the outer areas – at the same time. A side-effect of this possibility is to allow two-handed input.

Reducing the amount of information can actually support users in many tasks. Adaptive focus plus context displays provide a solution, because they allow users to keep a limited set of objects in their focus. Users can retrieve highly detailed information from those objects and interact with them. At the same time other objects are kept in the context region in lower detail. Our system is designed to allow multiple users a fast, but coarse-grained access to objects, which are "parked" in the periphery, but provides detailed input for fine-grained interaction with objects in the center. Switching the object, which is currently in focus and used for manipulation, with objects in the outer areas should be as easy as moving sheets of paper to the border of a traditional workspace and back to the center.

In this paper we present a low-cost, multi-precision input system. We demonstrate the combination of high resolution tracking technologies in focus regions with low accuracy webcams used for optical finger tracking in the periphery. To avoid confusion we present techniques for adaptive system feedback to inform users about currently possible interactions, i.e. fine or coarse grained. Finally, this paper illustrates two demo applications to show how multi-precision input and feedback can be used in instrumented environments.

2. Related Work

Baudisch et. al. [2][3] describe the principle of focus plus context displays. They found that users working with a large amount of information mostly appreciate having lots of space. This can be done using multiple screens or having a projector to increase a single display's size. The projected image offers lower physical resolution. Their work considers focus plus context displays only in the output direction for single users with a mouse and a keyboard as input devices. The interactivity on the context part of the screen is also realized with mouse input resulting in lower physical resolution as well. Further research on large displays concentrates on interaction techniques which just transfer the desktop metaphor to a bigger size [4][6][26]. In this work constant physical tracking accuracy is assumed across all involved displays. Others try to devise novel interaction concepts which are better suited to the particular interaction situation [5][18]. One of our demo applications – BrainStorm – is strongly influenced by Igarashi's work on sketch-based interfaces [11][19].

Previous work has also investigated how displayed information on large screens can be made interactive. The HoloWall [18] leverages back-projected infrared light which is reflected by the users' fingers or objects. This light is then captured by an infrared camera to identify the location of each finger. In addition, they show how hovering is enabled by using the brightness of detected points. In their work they consider back-projected displays instead of front-projected solid walls. Han's [10] multi-touch system applies frustrated total internal reflection (FTIR). It leverages the property of light being refracted to a certain extent if traveling between two different media. In his work, the camera receiving infrared light signals is again mounted behind the projection surface. The Digital Vision Touch (DViT) [29] also employs infrared light emitted parallel to the display surface. Four cameras mounted in the corners of the surface recognize objects as black blobs which can be positioned using triangulation. In addition, the system is able to detect objects hovering over the surface. This system does not allow objects to be pinned onto the tracked surface. Furthermore, it distinguishes two different positions at the same time. While this number may be sufficient for tabletop interfaces it is inadequate for large wall-sized displays.

The DiamondTouch system [7] employs capacitive sensing in analogy to today's touch screens and touch pads. In this work users are active coupling devices for the system. The antenna arrays are built into the front-projected tabletop surface whereas the receivers are attached to the users' chairs. The DiamondTouch also facilitates determination of different users working on the table at the same time. However, back-projected displays hardly work with built-in antennas unless they are implemented transparently. The MultiSpace [8] combines horizontal as well as vertical displays which span an interactive workspace with the DiamondTouch table as central hub. Their work additionally employs an interactive, vertical surface which accepts single-user input. This setting strongly influenced the setup of our second demo application BrainStorm.

Large and wall-sized displays have been used in instrumented rooms, such as iWork [13] or Roomware [22][30] with a DynaWall [9] and ConnecTables [31]. They mostly use off-the-shelf SmartBoards to allow interaction with the displayed information. Thus, this work does not provide any interaction possibilities physical parts of the environment such as a concrete wall with projected information. In addition, Rekimoto et al. describe the concept of continuous workspaces [24] which allow users to treat different display surfaces as continuous screen. This system also considers the spatial arrangement of these individual screens.

Hand- or finger recognition is another issue in research. In general, there are four ways to segment a human hand which have been summarized in [12]. First, it is possible to use a controlled and undisturbed background within the captured image. Others try to use a known background for later image subtraction. In addition, it is possible to apply motion segmentation or the color segmentation to captured images. All of these techniques are dependent of the environment and scenario they are used in. Color segmentation is hard to employ if the tracked finger might have arbitrary colors due to indirect lighting. Furthermore, using known backgrounds might not work when the fingers' will have the background color due to display radiation.

3. The Concept of Variable Precision Input

Our goal is to create an interactive surface in analogy to the working surface of a desk (see Figure 1) comprising a main working area for precise activities such as writing or sketching. In addition, users can employ the remaining space as storage clipboard for other, currently unused documents usually arranged in stacks. In these areas, a person would only use rough interaction such as moving or grabbing documents.

Precise activities such as writing need high precision as well as fast response times for interaction. The outer regions only require low-resolution input since interaction there is less precise. For example, moving a document can be realized by grabbing it at any position whereas writing needs pixel-level accuracy. Thus, large interactive surfaces at least require a high-resolution input area in the center as well as lower accuracy in the periphery (see Figure 2). This ensures interaction throughout all surface areas.

3.1. Real-Time Finger Tracking

Interactive surfaces need to offer fast visual feedback for users. Thus, input generated by fingers touching the surface needs to be processed real-time as close as possible. While the center region will be tracked using an off-the-shelf high precision technology, the periphery is observed using low-cost webcams.

In order to achieve a reasonable precision of about 10-20 millimeters in the context areas of our display, we chose triangulation as the base tracking technology. Several cameras determine the position of a finger in the image and from this derive the angle between their optical axis and the connection line between camera and finger. This technique is well-known from other products using four cameras such as the SmartBoard from SmartTech [29].

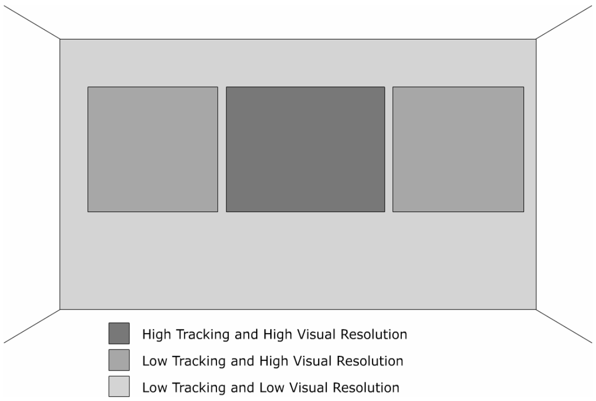

Figure 2. A possible setup of an interactive surface utilizing variable precision input and focus plus context output.

Using two cameras at known positions and orientations, the intersection of those lines is the final position. In order to track multiple fingers simultaneously, optical tracking systems need to employ additional cameras. For example, using a four camera setup with modifications and extensions to the base method allows the detection of up to four points. In addition, this setup increases the system's robustness and reliability.

3.2. Detecting Multiple Points

Multiple fingers touching the surface lead to a number of detected lines by each camera. Thus, the number of intersections of all lines is higher than the number of actual objects touching the surface. Hence, we introduce several modifications to the base tracking technology to finally gather the correct positions on the surface. Due to the limited resolution of low-cost webcams, exact intersections of four involved lines will not exist. We can address this issue by allowing points to be close (i.e. below a certain threshold) to other intersections detected. Figure 3 (left) shows one situation described above.

However, four cameras can only track multiple fingers correctly, if a finger is not occluded by another finger. Further problems arise when fingers are on a virtual line connecting two cameras. In this case, lines that are involved in an intersection are (nearly) parallel. Leveraging four cameras we can reduce this problem as long as the finger is not close to the center point of the tracked surface. In this case the intersection would have two parallel pairs of lines and it remains unclear whether this is a valid point or not (see right side of Figure 3).

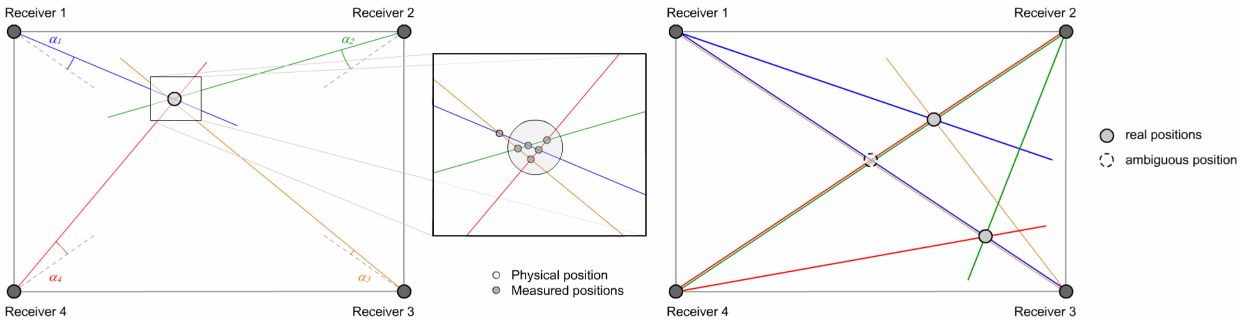

Figure 3. Left shows the positioning with four cameras. The magnification illustrates the tolerance area for a single position. Right shows the detection of two fingers and an ambiguous position that cannot be resolved with the basic algorithm of triangulation.

One of our goals is to minimize these issues caused by optical tracking. Thus, we defined further criteria which must be fulfilled by an intersection to be an accepted position:

- *Clearness of involved lines.* An intersection must provide at least three lines without having a pair of them nearly parallel. In particular, the angle between involved lines must not be less than a certain threshold. Otherwise, the two lines will become one involved line regarding this intersection.
- *Unambiguous mapping of lines to intersections.* Two intersections must have at least three different lines involved. If two lines are contained in two intersections the more accurate intersection will be taken while ignoring the second one. Due to occlusion problems it is important to have at most one line involved in multiple intersections.

The second criterion will obviously prevent finger tracking in the center region of our interactive surface. However, this is nicely compensated by the overall setup of our focus plus context display with the high resolution input area in the center.

3.3. Input Fusion Architecture

The major goal is to have multiple tracking systems with variable accuracy combined into one single input stream. As long as fingers are not crossing the boundaries of variously tracked regions this is a trivial task. Once a user crosses these boundaries, the system should still be able to associate the previous input stream with the new input stream coming from a different source. Thus, we introduce an abstraction layer to merge various streams into one uniform input stream. This layer also ensures the correct association of positions to fingers which allows users a seamless transition between independently tracked regions on the surface.

The input abstraction layer listens for positions detected by all registered tracking technologies. Whenever positions have been detected and sent to the input layer, it will immediately store these as new positions. Subsequently, the system will now match the new positions with previously detected ones using a simple algorithm called "Dead Reckoning" (Deduced Reckoning) as described in [15]. In addition, the system is now able to determine the fingers' action:

- *Finger down.* This position could not be associated with an old one. Hence, it describes a finger that has not been detected during the last five detection cycles.
- *Finger move.* The position can be matched to a previously detected finger. Thus, the finger has been on the wall during the last detection cycle.
- *Finger up.* All positions that do not have a match during the past five detection cycles. These fingers must have left the wall completely.

Fingers moving fast across the surface might not be detected in each detection cycle. Thus we introduced a value indicating the time (i.e. the number of detection cycles) the finger has been undetected. This allows having a finger occluded for a short amount of time.

Table 1. Final input event sent by a surface using variable precision input

Event type	Fields	Class types	TTL [ms]
NewPositions	FingerID	Integer	100
	Action	Enumeration [DOWN, MOVE, UP]	
	PositionX	Integer	
	PositionY	Integer	
	RawX	Double	
	RawY	Double	
	Precision	Double	
	Time	Long	

Once all fingers have been updated, the input layer sends the unified position data to all registered applications. As the input layer receives tracking data from both tracking devices in high frequency, it produces a steady stream of position data and sends it using the Event Heap infrastructure [14]. Connected applications can use this stream analogue to the operating system's mouse input events (see Table 1).

3.4. Performance and Limitations

Our implemented system provides pixel-accurate input in the central region as well as lower precision input in the periphery with a speed of approximately 15 positions per second. The accuracy in the periphery depends on the scale of the interactive surface (width w_i and height h_i), the width of captured images w as well as the camera's angle of aperture α. Equation 1 gives the maximum error ε (millimeter) between real and calculated positions. Using this equation one can also determine the size of an interactive surface regarding the cameras' resolutions or vice versa.

$$\varepsilon = d_i \cdot \tan\left(\frac{\alpha}{w} \cdot 1px\right), d_i = \sqrt{w_i^2 + h_i^2};$$

Equation 1. Maximum error of the tracking system in millimeter

In our system the transition between both tracking subsystems is seamless and a listening application does not need to know from which input region positions have been calculated. Nevertheless, by evaluating the precision parameter provided in the input stream, applications might decide to offer only certain operations in lower precision areas. In addition, the interactive surface can track up to six fingers simultaneously (four in the outer regions plus two in the central region).

Other objects can be tracked on the surface as well except very dark ones as they will disappear in the captured image due to black background. Finally, a major problem is the system's sensitivity to light conditions. Especially direct light hitting the lens of a camera inhibits tracking entirely. This is a usual limitation for visual tracking systems.

4. Multi-Precision Feedback

One of the most important aspects of such systems is to give adequate feedback to users working with it. Intentionally, they might not know that tracking in the outer regions is not as accurate as in the center. Also, only a limited or different set of possible interactions should be offered regarding the resolution of the tracking technique. Chia Shen et al. [28] describe the concept of occlusion-aware visual feedback. This also influences the feedback for multi-precision input systems as it gives feedback in different resolutions. We have designed three possible techniques that allow multi-precision feedback to the user in several ways. All identified techniques also work in environments with more than two levels of precisions.

4.1. Accuracy of Interaction

Our system is designed to have lower resolution and thus lower accuracy in the peripheral regions of the interactive surface. Since coarser precision can be a limitation for certain interactions, the user needs to be aware that his or her finger might be detected on a slightly different position. One possibility is to give hints to users that show an aura in which their finger will be detected. With this region, users are able to see which parts of the surface might be affected by their action. In addition, this technique is also able to visualize the chance of an action taking place at a specific position.

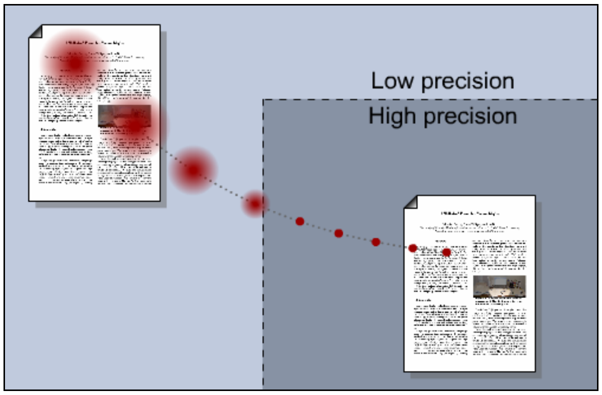

Figure 4. User feedback of inaccuracy occurring in regions tracked with low precision. The circle's size determines the inaccuracy of measured positions

On high precision regions this is usually a small point without an aura because the action will be activated at the touched point. On lower precision areas this region would be a circle. The circle's radius is dependent on the position's accuracy. Furthermore, the transparency of it will increase at points farther from the circle's center. Before a user touches the wall, this technique could give an impression of where the action might be. This can be realized using a hovering mode available by most optical tracking technologies.

Figure 4 sketches the idea of this technique on high precision as well as on low precision tracked surfaces: The user crosses the border between surfaces tracked by different technologies and hence gets an immediate feedback of how accurate the detected position is.

4.2. Different Levels of Detail for Interaction

Another option we investigated is to provide different interaction levels considering the accuracy of the measured position. The number of different accuracies determines the number of levels regarding possible interactions. As Input events sent by the input abstraction layer provide a precision value, applications can use these events within their event bubbling system. Elements in the event hierarchy can define a minimum precision. Thus interaction behavior and visual feedback can be adapted dynamically.

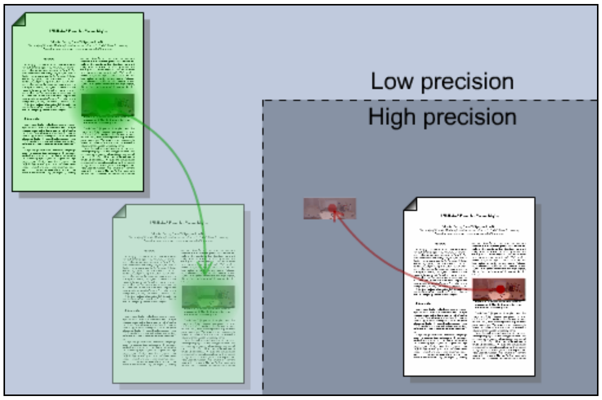

Figure 5. Different levels of detail regarding the precision of input in a certain region.

One scenario might be that several people are working with textual documents which may contain pictures or charts. Interaction with and within these documents can be split into several granularities. Some examples of dealing with textual documents are:

- *Moving, scaling or modifying the whole document.* The entire document may be moved around the whole display by grabbing it inside.
- *Moving or modifying embedded objects.* Pictures or other objects within the document can be dragged to another place in the document or to another application to further modify them. In addition, their size and/or orientation can be changed in place.
- *Selecting text.* Text within the document can be selected in order to copy it to the clipboard or to drag it to another position in the document.
- *Writing text.* Text can be added to a document by simply writing it with a pen or a keyboard. In case of using the pen, the system needs to have handwriting recognition as well as high input resolution.

As shown above, even text documents offer a variety of interaction levels that can be employed by multi-precision input systems. For example, moving a document in high precision regions can be realized by grabbing it on a position which does not contain any information (e.g. the margins). This ensures that a user is still able to modify information included in the document such as pictures or tables by tapping on them.

In contrast, this level of interaction is not needed in the peripheral regions. These regions are designed to offer an interaction granularity that allows moving and sorting them. Thus, grabbing a document at any interior position facilitates moving the entire document while access to included objects is denied. This technique does not require a touch on an exact position to drag and hence rearrange documents in the periphery.

As shown in Figure 5, visual feedback for users can easily be done by highlighting the area in which the interaction will take place: A user touches the document on an included image. Dependent on the tracking accuracy the system can decide whether the entire document (*low precision*) or just the selected image (*high precision*) will be dragged.

4.3. Size of Interaction Area

With decreasing tracking accuracy fine-grained interactions get more and more difficult if the objects' sizes remain the same in all regions. One solution to bypass this issue is to scale objects dependent of the area's accuracy the objects are currently located in. This technique allows the same level of interaction on both peripheral and central region but necessitates more space on surfaces employing low accuracy input technologies. Thus, sorting documents in outer regions cannot be used as with other techniques described before. One solution is to shrink a document after it has been inactive for a certain amount of time. Once a user touches an object again it will be scaled to its old size.

5. Implementation Details

In this section, we describe the general steps to implement our wall-sized focus plus context display. First, we introduce the setup for variable precision input technologies followed by a detailed description of the setup of different display technologies we have built into our instrumented environment.

5.1. General Input Setup

Users should be able to interact on the wall with their bare fingers. In the wall's center, i.e. on the middle screen, we wanted the input to be very exact and fast in order to support detailed interaction in a focus area. This was achieved by a SmartTech SmartBoard [29]. The surrounding areas of the wall as well as the side displays, serving as context areas, were meant to have a lower input resolution. We decided to use optical tracking for this, as it would allow bare finger interaction and interactions on non-display surfaces. As our system is wall-sized, multiple users can stand in front of the display and interact simultaneously. Hence the system needs to support simultaneous input. The wall should appear as one interactive display to the user as well as the programmer. Thus none of them should be aware of the tracking technology currently used. This requires seamless transitions between the technologies while users interact with the wall display. The combination of different tracking systems together with the simultaneous multi user input bears several technical challenges.

The setup of the wall-sized tracking system mimics the SmartTech DViT technology [29] by arranging four cameras in the four corners of the wall. In order to cover a large part of the wall, we used Logitech's QuickCam Fusion [17] with a resolution of 640 x 480 pixels, 30 frames per second and a diagonal field of view of 72 degrees. Other cameras had either a lower resolution or a lower field of view. To ensure a fast detection close to real-time processing, the tracking system runs on an Intel Pentium 4 workstation with 3.0 GHz and 1 GB of RAM. This computer is dedicated to the camera-based tracking system. As the webcams have a field of view of 72 degrees we needed to decide how they will be arranged on the wall to match the following requirements: First, Every position on the wall's surface needs to be observed by at least two cameras. Second, the coverage area of three or four cameras should be maximized. Out of several possible arrangements (see Figure 6) we chose to use setting (c). The area that is observed by four cameras covers 74.2 % of the total surface area.

The cameras needed to overlook the entire wall with their optical axis parallel and close to the wall's surface, but space did not allow them to be embedded into the wall. Therefore we used a mirror construction with cameras facing directly towards the wall. This allows observing a small rectangular area over the wall's surface. Initially, we modeled the ideal mounts in a 3D rendering program that allowed us to test our formulas and parameters. The actual mounts have been built from wood and hold the mirrors. We have attached black foamed rubber to avoid that the cameras capture too much indirect light emitted by various sources in our instrumented environment.

5.2. Display Setup

Our main goal was to make a wall entirely interactive in the input as well as the output direction, so that it would appear as one big logical screen to users as well as programmers. At the same time, it should support different degrees of interactivity in different areas. In order to provide high resolution output at regular working heights, we embedded three back projection displays (center: 147 x 112 centimeters, side displays: 120 x 112 centimeters) into the wall which has overall dimensions of 4.50 x 2.40 meters. The total resolution of these three screens together is 3072 x 768 pixels, which corresponds to a spatial resolution of about 0.8 pixels per millimeter.

The remaining part of the wall is covered by a steerable projector – the Beamover 40 from publitec [23] – mounted at the ceiling, which creates a movable display area of about 60x45 centimeters at a spatial resolution of about 1.7 pixels per millimeter without image correction. The projected image is rectified using a technique similar to the one described in [21] resulting in a reduction of the spatial resolution. The parameters of the projector (e.g. *pan*, *tilt*, *focus* and *zoom*) can be set in real-time using a USB to DMX interface. This makes the entire wall above and below the screens a time-multiplexed low resolution display. In addition to projected images, physical objects may be stuck onto the walls surface. Hence, users are able to attach post-it notes, printed pages or physical photographs.

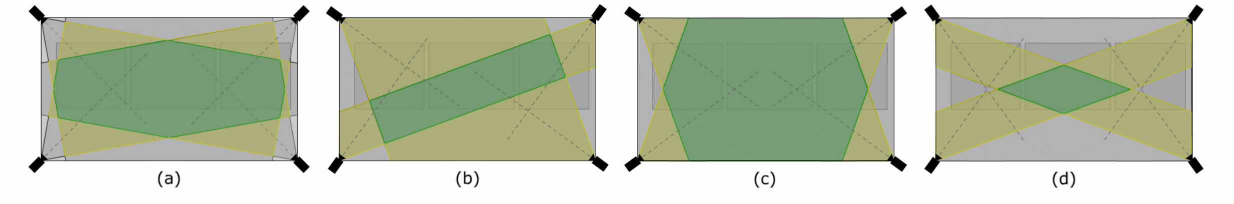

Figure 6. Different orientations of four cameras. Green (darker) areas represent the coverage of all four webcams while yellow (brighter) areas show the coverage of three cameras.

6. Example Applications

We implemented two example applications that make use of all the system's input and output components to demonstrate its capabilities. These applications use the steady input stream from the input layer as if they were system mouse events.

6.1. WallDraw

Our initial test application enables users to draw on the wall using their fingers. The three screens form one single display, but are tracked by different input systems. The steerable projector provides a tool palette upon request, which is displayed in the context area on the wall outside of the drawing area. This saves space otherwise wasted by placing tool bars in the drawing area. Furthermore, users do not need to move to the palette physically since it can be displayed everywhere. As mentioned before, the wall's size allows multiple users in front of it drawing simultaneously on the displays. Thus, the tool palette provides every user with their own choice of tools (line, curve, rectangle and eraser) independently.

While the displays do not consider the different tracking technologies at all, the tool palette has been enlarged to ensure fast and reliable interaction with it. Thus, the palette implements the concept of scaling the interaction area and does give feedback after the user has selected another tool. In this case, scaling the palette allows users to touch the wall in a certain area within the desired visualization of a tool. Since users will still touch the tool close to its center this technique hence ensures the correct selection.

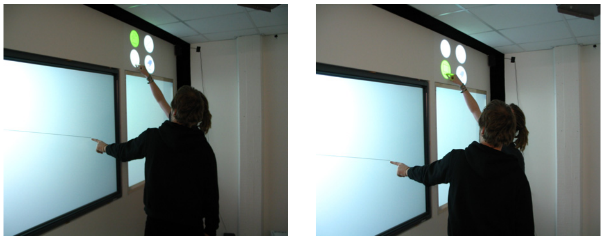

Figure 7. Left shows the displayed tool palette before selecting a new tool. Right shows it after the selection.

Users can "call" their palette by tapping on the wall outside of the displays. The palette will then be displayed at the detected position showing the currently selected tool for this user (associated via the nearest drawing position). Now the user can select another tool by simply touching it (see Figure 7). While a user is selecting a tool, others are still able to use the application independently. *WallDraw* runs on two different machines which control the display wall (drawing) and the steerable projector (tool palette) respectively. It is not aware of the different input technologies which are used to enable drawing across the entire wall.

6.2. BrainStorm

Our second and more complex demo is a brainstorming application for a team consisting of up to 4 users. The process of brainstorming is divided into different phases. In the first phase, all users comfortably sit around an interactive table and create notes of ideas by scribbling sketches or keywords on the desk's surface, which then turn into virtual post-it notes. Users interact on their side of the desk and write notes oriented correctly for them. When the notes are created, they appear simultaneously on the wall's focus display, where they are all oriented upright. Their relative positions on the display still match their positions on the table, so users can easily find their own notes again on the wall display through a direct spatial mapping.

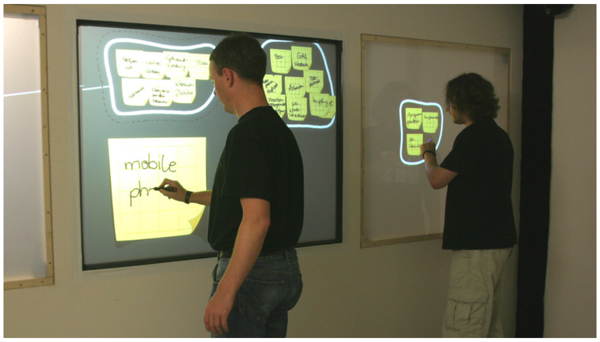

Figure 8. Our variable precision input system with one user doing fine-grained interaction and another using coarse tracking for organization.

Users can now stand up and move notes around the entire wall, group them, form clusters by drawing a line around a set of related notes, and create connections among clusters, between clusters and notes by drawing a connecting line (see Figure 8). Notes or clusters, which are decided to be of secondary importance in this phase, can be moved to the context area of the display, i.e. the two side displays. Notes or clusters which are discarded altogether in the second brainstorming phase can be deleted from the wall by dragging them over the border of a back-projection display, i.e., out of the working area. In this setup, the inner focus display represents a group focus. It is still large enough to accommodate multiple user interaction, but only on this middle display, fine-grained manipulation takes place. The two adjacent context displays serve as a visual clipboard, where clusters can be stored when they are currently not being dealt with. This frees space in the focus area, but maintains visibility of these clusters, so that the collection of notes and clusters can still be overviewed when users take a step back.

This application employs different levels of interaction. Post-its can be enlarged in the central area of the display to add text by simply touching them at their center. In the periphery this is not possible as the tracking does not provide accuracies that allow fine-grained interaction such as writing. Thus, post-its and clusters can only be dragged in order to rearrange them. This second structuring phase eventually results in a mind map of related and grouped concepts and visually represents the result of the brainstorming.

7. Conclusions and Future Work

We have presented a wall-sized focus plus context display using different tracking technologies and have shown two example applications of it. The display implements the focus plus context concept in two directions. In the output direction, it contains the back projection displays, which provide a simultaneous resolution of about 0.8 pixel per millimeter, and the rest of the wall which can display content using the steerable projector at a resolution of about 1.7 pixel per millimeter (not rectified), but most importantly, only in one limited area of about 60x45 centimeters at a time, i.e. time-multiplexed. In the input direction, the wall is split differently: the middle one of the back projection displays has pixel-exact low latency finger input, which is provided by the commercial SmartBoard hardware. The adjacent, back-projected displays and the wall above and below, has less fine-grained input at a resolution of about 10-20 millimeters, depending on the exact location and a tracking rate of 15Hz independent on the number of fingers. This divides the focus plus context display into three primal areas:

- high input and output resolution: center display
- low input and high output resolution: side displays
- low input and output resolution: rest of the wall

In addition we have shown three techniques for multi-precision feedback to inform the user what kind of accuracy s/he is tracked with. Each of those considers different aspects of interaction such as visualizing tracking accuracy, prohibiting certain levels of interaction or rescaling of objects according to the area's tracking accuracy. We have illustrated how event bubbling mechanisms are able to use the input stream of our system to decide the level of interactivity an object will have according to the tracking precision.

Our demo applications describe a close co-located collaboration between several users on a single large display. Users share a common area of attention and interact simultaneously in it. Particularly in this situation, the distinction between different levels of focus makes sense because of the wide viewing angle we have to cover to see the full display. When users are close enough to interact, the context area is mostly out of their physical reach. This suggests that fine-grained interaction in the outer areas is not needed in this situation. Nevertheless, users can easily increase or decrease the perceived section of the display by stepping back or moving towards the display.

Our current implementation still suffers from a slight jitter due to timing problems with the four USB cameras. When a finger is moving and the camera images are not taken at the exact same time, the computed positions are slightly wrong. We hope to increase tracking accuracy and reduce jitter by applying temporal filtering to the signals. Although we haven't formally evaluated our large focus plus context display, the first impressions and user feedback are mostly positive. We are currently trying to increase tracking accuracy and reduce jitter in the outer tracking areas by applying temporal filtering to the signals. After this, we will formally evaluate BrainStorm.

Acknowledgments

This work was funded by "Deutsche Forschungsgemeinschaft" (DFG). We thank Lucia Terrenghi for designing and David Kim for implementing substantial parts of BrainStorm. Further, we would like to thank Albrecht Schmidt for his valuable feedback.

References

[1] Ballagas, R., Rohs, M., Sheridan, J. and Borchers, J., *Sweep and Point & Shoot: Phonecam-Based Interactions for Large Public Displays*, in CHI '05 extended abstracts, ACM Press, 2005, p. 1200-1203

[2] Baudisch, P., DeCarlo, D., Duchowski, A., and Geisler, B., *Focusing on the Essential: Considering Attention in Display Design*, in Communications of the ACM 46(3), March 2003, p. 60-66

[3] Baudisch, P., Good, N., Stewart, P., *Focus plus Context Screens: Combining Display Technology with Visualization techniques*, in Proceedings of UIST 2001, p. 31-40

[4] Bezerianos, A., Balakrishnan, R., *The Vacuum: Facilitating the Manipulation of Distant Objects*, in Proceedings of CHI 2005, p. 361-370

[5] Cao, X., Balakrishnan, R., *VisionWand: Interaction Techniques for Large Displays using a Passive Wand Tracked in 3D*, ACM CHI Letters, 5(2)

[6] Collomb, M., Hascoet, M., Baudisch, P., and Lee, B., *Improving drag-and-drop on wall-size displays*, in Proceedings of GI 2005, p. 25-32

[7] Dietz, P., Leigh, D., *DiamondTouch: A Multi-User Touch Technology*, in Proceedings of UIST 2001, p. 219-226

[8] Everitt, K., Shen, C., Ryall, K., Forlines, C., *MultiSpace: Enabling Electronic Document Micro-mobility in Table-Centric, Multi-Device Environments*, in IEEE International Workshop on Horizontal Interactive Human-Computer Systems, IEEE Xplore, p. 27-34

[9] Geißler, J., *Shuffle, throw or take it! Working Efficiently with an Interactive Wall*, in Proceedings of CHI 1998, ACM press, p. 18-21

[10] Han, J., *Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection*, in Proceedings of UIST 2005, ACM, p. 105-108

[11] Igarashi, T., *Freeform User Interfaces for Graphical Computing*, in Proceedings of 3rd International Symposium on Smart Graphics, Springer LNCS 2733, p. 39-48

[12] Jennings, C., *Robust Finger Tracking with Multiple Cameras*, in Proceedings of the International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems 1999, p. 152-160

[13] Johanson, B., Fox, A., Winograd, T., *The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms*, in IEEE Pervasive Computing Magazine 1, 2002, p. 71-78

[14] Johanson, B., Fox, A., *The Event Heap: A Coordination Infrastructure For Interactive Workspaces*, in Fourth IEEE Workshop on Mobile Computing Systems and Applications (WMCSA 2002), IEEE Computer Society, p. 83-93

[15] Kirsh, D., *The Intelligent Use of Space*, in Artificial Intelligence 73(1-2), 1995, p. 31-68

[16] Küpper, A., *Location-based Services: Fundamentals and Operation*, John Wiley & Sons Ltd., Weinheim, Germany, 2002

[17] Logitech Europe S.A., http://www.logitech.com, 2006

[18] Matsushita, N., Rekimoto, J., *HoloWall: Designing a Finger, Hand, Body, and Object Sensitive Wall*, in Proceedings of UIST 1997, ACM press, p. 209-210

[19] Mynatt, E. D., Igarashi, T., Edwards, K., LaMarca, A., *Flatland: New Dimensions in Office Whiteboards*, in Proceedings of ACM CHI 1999, ACM Press, p. 346-353

[20] Nova, N., *A Review of How Space Affords Socio-Cognitive Processes during Collaboration*, in Psychology 3(2), 2005, p. 118-148

[21] Pinhanez, C., *Using a Steerable Projector and a Camera to Transform Surfaces into Interactive Displays*, in Proceedings of CHI 2001, ACM, p. 369-370

[22] Prante, T., Streitz, N. A., Tandler, P., *Roomware: Computers Disappear and Interaction Evolves*, in IEEE Computer, Dec. 2004, p. 47-54

[23] Publitec GmbH, http://www.beamover.com, 2006

[24] Rekimoto, J., Saitoh, M., *Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments*, in Proc. of CHI 1999, ACM press, p. 378-385

[25] Rensink, R.A., *Internal vs. External Information in Visual Perception*, in Proceedings of 2nd International Symposium on Smart Graphics 2002, p. 63-70

[26] Robertson, G., Czerwinski, M., Baudisch, P., Meyers, B., Robbins, D., Smith, G., Tan, D., *Large Display User Experience*, in IEEE Computer Graphics & Application, July/August 2005, p. 44-51

[27] Scott, S. D., *Territory-Based Interaction Techniques for Tabletop Collaboration*, in Proceedings of UIST 2003, ACM press, p. 17-20

[28] Shen, C., Ryall, K., Forlines, C., Esenther, A., Vernier, F. D., Everitt, K., Wu, M., Wigdor, D., Ringel Morris, M., Hancock, M., Tse, E., *Informing the Design of Direct-Touch Tabletops*, in IEEE Computer Graphics and Applications, Sept./Oct. 2006, p. 36-46

[29] SMART Technologies Inc., http://www.smarttech.com/

[30] Streitz, N. A., Geißler, J., Holmer, T., Konomi, S., Müller-Tomfelde, C., Reischl, W., Rexroth, P., Seitz, P., Steinmetz, R., *i-LAND: An interactive Landscape for Creativity and Innovation*, in Proceedings of CHI 1999, p. 120-127

[31] Tandler, P., Prante, T., Müller-Tomfelde, C., Streitz, N. A., Steinmetz, R., *ConnecTables: Dynamic Coupling of Displays fort he Flexible Creation of Shared Workspaces*, in Proceedings of UIST 2001, p. 11-20