MAD Boxes: A Plug-And-Play Tiled Display Wall

Running head (shortened title): MAD Boxes

Title: MAD Boxes: A Plug-And-Play Tiled Display Wall

Authors: Ryan Schmidt, Eric Penner, Sheelagh Carpendale

Affiliation: Interactions Lab, Department of Computer Science, University of Calgary

Full address for correspondence, including telephone and fax number and e-mail address:

Ryan Schmidt
Department of Computer Science, University of Calgary
ICT 602, 2500 University Drive NW, Calgary, Alberta
Canada, T2N 1N4
email: rms@cpsc.ucalgary.ca
phone: 403 210 9499
fax: 403 284 4707

Abstract.

While interest in large displays is growing rapidly, they are still not common-place. Significant technical knowledge is required to construct and maintain current display wall systems. Our goal is to make large tiled-projector displays essentially ‘plug and play’. We want a design that can be incrementally expanded and reconfigured at will. We want a software environment that is identical to a standard desktop computer, with no need for rendering clusters and special libraries. We have designed a display wall solution that meets our needs. With our Modular Ambient Display (MAD) boxes, a variety of high-resolution large display configurations can be quickly assembled. By integrating interaction hardware into each box, we have created a stand-alone interactive large display component. Our system permits experimentation not only with the wall software, but the physical wall configuration as well.

1 Introduction

Over the last ten years computers have made significant advances into our everyday living and working environments. The primary way computers convey information to us is through visual display. Interest in display surfaces beyond the traditional desktop monitor has been growing. This is in part because we have new needs and tasks and in part because we want them to blend into rooms, halls, and furniture. New displays range from the very large to the very small and from the totally portable to those specified in the design of new buildings.

Few of the large-scale display systems available provide any sort of flexibility in regards to physical configuration. Usually researchers have to decide on a configuration or two that relate to their research. In turn, once built, the chosen configuration’s display parameters will influence the research. As we work towards discovering how to best merge information technology into our everyday environments it is difficult to decide a priori what the right configuration will be. Because of the significant cost, floor-space, and up-front planning required to construct a display wall, research institutions rarely have the option to re-design.

As a research group interested in interaction, collaboration and data visualization, large displays held much promise for increased information display-space and group interaction. These research interests established several of our design criteria. As our understanding of display wall construction and maintenance grew, we quickly realized that the existing display wall solutions did not meet our requirements. New design criteria were introduced, and we focused on designing a flexible and accessible wall. A key technical requirement was that the wall be as easy to use as a standard desktop monitor. In addition, we did not wish to be tied to our initial design or the limitations of our current budget.

To minimize any limitations we will later discover, we have attempted to support different types of large display configurations in a manner similar to the construction toy Lego™. With Lego™ a child can build trucks and boats and bridges that are quite good though less than perfect. The advantage is that a truck can become a boat. With this in mind we have built Modular Ambient Display (MAD) Boxes.

2 Related Work

Since MAD Boxes support the creation of many different types of displays, there is a considerable amount of related work. A significant amount of research has been directed towards building large-scale high-resolution display systems. The typical configuration is large, fixed display surfaces illuminated by an array of data projectors [4,5], these systems are known as *tiled display walls*.

A key problem with tiled display walls is image alignment. Early display wall projects were manually positioned to avoid gaps in the output image. Overlapping the projected images simplifies alignment and improves the uniformity of the composite image, but requires either physical or software-based edge blending techniques. Recent work has been directed towards avoiding manual calibration entirely. Geometric and photometric registration is done entirely in software, based on feedback from camera input [2,11].

Another issue inherent in multi-projector displays is application support. Standard PC operating systems support a limited number of display outputs. Existing systems generally use a cluster of rendering machines, one per projector. Rendering is controlled by applications on the host a computer, either using parallel rendering libraries such as Chromium for OpenGL (http://chromium.sourceforge.net/), or by forwarding system API calls to the cluster nodes [5].

Interaction with large displays, whether tiled or not, is a current research issue. Direct wall interaction is necessary to mimic the traditional whiteboard or blackboard environment. Several commercial systems are available. The

SMARTBoard, available from SMART Technologies (http://www.smarttech.com/) is capable of recognizing a single touch and hence is unsuitable for collaborative work. The recent SMART DViT system supports a maximum of two simultaneous touches, which is somewhat limiting for a large display wall. The touches are not identifiable, and small blind spots can occur. Identifiable touch input is possible with the MERL DiamondTouch system (http://www.merl.com/projects/DiamondTouch/), however the screen size is very limited and does not support rear-projection. The Polhemus (http://www.polhemus.com/) FASTRAK system provides identifiable 3D tracking with no blind-spot issues. However, the stylus is tethered, range is limited to approximately 5 feet, and any metal components will interfere with sensor input. Full 3D motion-tracking systems are promising, although quite expensive. The sensing environment must be instrumented and hence is essentially fixed. Users must also be instrumented; generally some sensors or tracking balls must be worn.

Several research solutions are being developed, including a 3D colour based input [1] and laser-pointer tracking systems [3,10]. Laser spots are identified using frequency-modulation techniques. Many of these systems place the cameras in front of the display surface, in the same space as the user. In this configuration, occlusion problems inevitably arise when multiple users interact close to the display surface.

3 MAD Box System

The principle goals for the MAD Box project are:

Modular. To create a modular high-resolution display system that allows us to experiment with the physical configuration of the display surface. We require desire a flat display surface with minimal gaps between display modules, ruling out tiled plasma and LCD displays.

Extensible: We are interested in truly large displays walls and believe that questions regarding exploding the visual frame, as being investigated by Shedd [9], are of considerable interest for data visualization. However, our moderate budget will not support this; hence we would like to build an extensible wall that we can expand when feasible.

Movable: Large display walls – not surprisingly – require a large amount of space. Space is required for the screen support structure, projector mounting, and the projector throws. Existing tiled display walls consume a large and relatively permanent amount of floor space. Since our existing lab structure did not support a dedicated footprint for our wall, we needed it to be moveable and, in the extreme case, removable.

High resolution: Pixel resolution of a single LCD projector is currently limited to 1400x1050. Multiple projectors must be tiled over a large display surface to achieve higher resolutions. Our SMARTBoard provides 1024x768 pixels over a 72” display surface, approximately 18 dpi. We require at least double that resolution.

Regular Software Environment: Researchers must be able to run their software written for single-display machines on our multi-display configurations with no modification. By reducing the software barrier-to-entry, we hope to encourage novel display wall research projects.

Interactive: Simultaneous, close-proximity direct interaction with the display wall is necessary for collaborative use. Ideally our interaction system will be easily portable, requiring minimal set-up and configuration.

Figure 1 Box Frame

Figure 2 Projector Mount

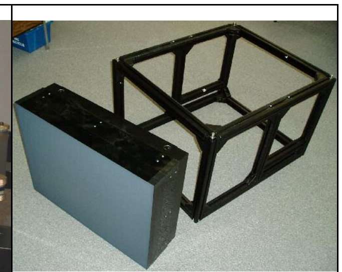

Figure 3 Floating Screen

3.1 Display Hardware

Our tiled display wall is assembled from modular, self-contained stackable display units, which we call *MAD Boxes*. Each MAD Box consists of an LCD projector, an aluminum frame, and a floating rear-projection screen. Several of our design goals are met by the MAD Boxes. Wall assembly is a matter of stacking the boxes in the desired configuration and inserting some stabilizing brackets. As a result, the wall is movable, removable, extensible and reconfigurable. The rectangular design affords a variety of configurations – a wall, a tower, several towers, a cave, etc (Figure 8). Note that while it is physically possible to position MAD Boxes so the display surface is horizontal (Figure 8), this currently voids most projector warranties and is potentially a fire hazard without extra projector cooling facilities.

The aluminum frame is 29 inches wide by 22 inches high, providing a 36 inch diagonal display surface. The frame is 34 inches deep, and includes enough space behind the projector for the ends of power and video cables. The projector is mounted on an adjustable alignment platform (Figure 2) that is bolted inside the frame. Our current alignment solution provides 6 degrees of freedom. Translation in the X and Z axis, as well as Y-rotation, can only be accomplished by moving the entire platform. Four bolts provide Y-translation, and limited X/Z rotation.

Despite this restricted system, physical alignment of the MAD Boxes is relatively simple. Alignment is limited to a single box, and is greatly assisted by the presence of the screen frame. Using a full-screen black image with a one-pixel wide white border, an operator can stand behind the box and manipulate the alignment controls until the white border is in the correct position. The projectors are mounted in essentially the correct physical location, so only small adjustments are necessary, primarily to correct for keystoning. Because alignment is local to the MAD Box frame, alignment errors in one box do not cascade to the rest of the wall. After alignment, the entire system is fixed in place by locking nuts and is very stable even when moving the box. It is possible for the projector alignment to drift internally, however we have not actually observed this in practice.

Each box contains one NEC MT1060-R LCD projector with a short throw lens. These projectors provide 2600 ANSI lumens, resulting in a very bright image that is easily visible in a well-lit room. The projectors are quite uniform in color and brightness. The projectors are programmable through both serial and wireless interfaces, providing easy access for software-based photometric alignment systems. Each projector runs at a resolution of 1024x768 pixels. Projected across a 36-inch diagonal screen at 4:3 aspect ratio, we have approximately 36 dpi, satisfying our resolution requirement.

A 36-inch diagonal acrylic back-projection screen is affixed to a removable floating mount that hangs off the front of the frame. The floating mount is approximately 9 inches deep (Figure 3). Because each screen is small and supported by a mount, they are sufficiently rigid that one can push on a screen, or rest a hand on it, and it will not flex. Individual screens can also be replaced if they are damaged. The screen is attached to the mount using tape, glue, and plastic sheets less than 1mm thick. This results in a very small inter-screen gap. As shown in Figure 4, when the screens are properly aligned the gap is approximately 2mm. In our experience this very small gap is barely noticeable. This may parallel evidence gathered in [6] showing that removing the gap pixels largely mitigates any discomfort for users.

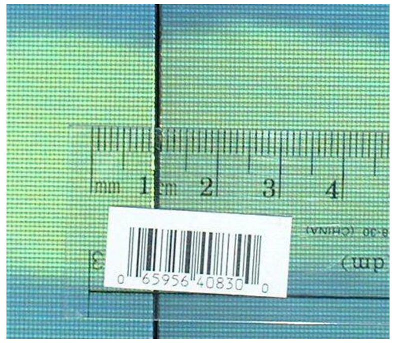

Figure 4 The physical gap between MAD boxes is less than 2mm.

3.2 Display Software Environment

Since it is our intention that the multi-display environment be transparent to as many user applications as possible, we have looked for solutions that use a single computer to render the display. Several off-the-shelf display adapters exist that support 4 separate video outputs on a single PCI card. Using two of these cards we can drive 8 MAD Boxes with one computer (Figure 5). Recent versions of Microsoft Windows can be configured so that all 8 outputs are merged into a single desktop. Existing applications can be stretched across the entire 8-output desktop and used normally.

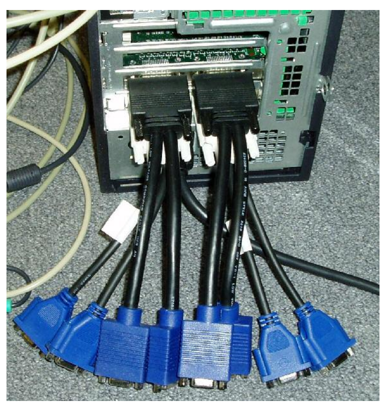

Figure 5 One machine drives all 8 MAD boxes in our wall.

This off-the-shelf hardware solution has some limitations. Configuring a standard Windows XP PC to function with multiple display adapters is simple. However, many display operations common in current applications - such as video playback and fast 3D graphics – require hardware support. The display hardware manufacturer must provide drivers that are capable of coordinating multiple boards to play back a single video stream. Because this is not a common mode of operation, few manufacturers provide this support.

Initially we used display adapters with NVidia Quadro4 NVS 400 chips. These 4-output adapters cannot support video playback that spans an arbitrary number of displays. In addition, both Direct3D and OpenGL are not hardware accelerated in a window stretched across more than one display. To support Direct3D 9 we used a technique similar to OpenGL cluster-rendering systems, but applied to a single machine. We intercept Direct3D calls using a custom Direct3D DLL. These intercepted calls are forwarded to individual rendering contexts created for each display that the full Direct3D window covers. This technique is sufficient to support hardware acceleration for many Direct3D applications in both full-screen and windowed modes. However, a similar solution for OpenGL using multiple rendering contexts does not work with the available drivers for this board.

We have had much better results using the Matrox QID Pro video adapter. This board supports OpenGL and Direct3D with hardware acceleration across all 4 outputs. We use two of these cards and obtain hardware-accelerated display across eight display boxes. With this configuration we can play DVD video across 8 boxes at full frame rate. Software video formats can be stretched across the wall, however the framerate varies with resolution. The fundamental issue here is fill rate, and also occurs with 2D graphics in some cases. We have not determined whether the current limit is due to the CPU or the display hardware. Hardware-accelerated OpenGL is supported across the entire display. Framerates similar to those for a single display are achieved for a variety of applications. The graphics boards use the PCI bus and hence are primarily geometry-limited for interactive 3D graphics. We have also run the PC game Quake 3 Arena across 4 boxes at 100 frames-per-second.

Our single-machine configuration provides a standard Windows XP environment that is identical to a single-head machine. No special libraries are required to display output on the wall. Every windows application we have tried runs without any display or refresh-rate issues. Some of the applications we have tested include web browsers, Microsoft Word, PowerPoint, AutoCAD, Maya, Visual Studio, and a variety of student software.

Our wall has no graphics cluster. This significantly reduces the technical challenges inherent in set-up and maintenance of the display wall. No special libraries are necessary and no expertise with distributed rendering is needed. However, there are scalability limits. At this point in time a maximum of four 4-head cards can be supported in a single machine. This translates to a maximum of sixteen screens. Since the graphics adapters are PCI-based, there is a limit on OpenGL geometry throughput. In addition, the cards lack programmable vertex hardware and have limited texture memory. Fill-rate for 2D and video applications may also be restricted in some cases. Increased scalability can be achieved by resorting to a cluster-rendering system [4,5].

3.3 Interaction Systems

Our primary interaction goal is direct 2D wall interaction. Touch input has appeal in that it is natural and requires no additional objects, and can be implemented using the system described in [8]. However, consider the following scenario. Several collaborators are considering a visualization of their data set. One touch activity is to use one's finger the trace characteristics within the data. Usually this should not be an active touch in that is should not commandeer one of the available cursors and it should not solicit an interactive response. We would like an equivalent to touch input that does not eliminate casual touching and indicating and does not prevent resting one's hand on the displays surface. Whiteboard pens provide a familiar and appropriate metaphor, so we have been exploring interaction systems that would work in this physical form.

We have developed an experimental interaction system for our display wall that is based on tracking of light. Our system is inspired by the camera-based tracking of frequency-modulated laser pointers discussed in [3,10]. We also use a computer-vision system. We currently use color to identify individual pointers, although this could be combined with frequency-modulation to increase the range of pointer identities.

We base our input system on low-cost commodity USB 2.0 webcams. These cameras produce a 640x480 pixel uncompressed video stream at 30 frames per second. The operating system hides hardware-specific details behind a standard camera API. We use the open-source OpenCV vision library (http://sourceforge.net/projects/opencvlibrary/) to perform computer vision and image processing functions. The cameras use CMOS sensors, which are sensitive to infrared (IR) light. By blocking non-IR light and using only invisible IR light, novel input systems can be designed, such as [8].

Occlusion and shadowing effects are an issue if the camera is placed in front of the screen. Instead we put cameras behind the screen, inside each MAD Box. Two cameras are needed per box because the field-of-view of a single camera is too low to see the entire screen. The field-of-view of the camera can be increased with aftermarket wide-angle lenses, however this is at odds with our off-the-shelf system goal. This raises a new issue – vision processing requires significant CPU power. With two cameras per box, we now have 16 cameras for the entire wall. A single machine cannot handle this many cameras. Hence, we mount a small computer in each box to process the camera output (Figure 6). These computers forward tracking data to the main computer, which can process the tracking data or forward it to other software. Because the cameras and computer are mounted inside each box, the modularity of the system is maintained. Each box is still a self-contained unit, providing display and interaction support.

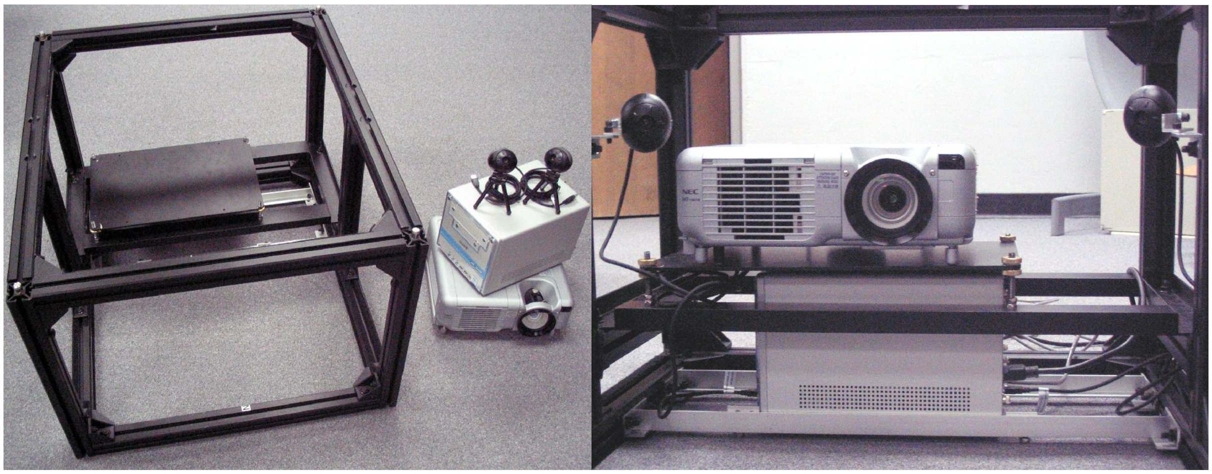

Figure 6 Two webcams are mounted inside each MAD Box. A small PC processes camera output and sends the results to a host computer over a LAN. Each unit is self-contained, the only external cables are power, video input, and network.

Our current tracking algorithm simply identifies bright spots, or "blobs", in each image. The blobs are created by laser pointers and LED lights. Laser pointers are usable at a distance, however LED lights only work within a few inches of the screen. Because of this proximity limitation with LED lights, our rear-camera solution is required. By reducing the brightness of the LED we can require that the user physically touch the LED to the display surface. This restriction may be desirable, as it is closer to the whiteboard metaphor.

To assign an identifier to a tracked blob we attempt to recognize the color of light the input device produces. Commercially-available laser pointers that are safe for the human eye are limited to red and green. This limits the number of identifiable inputs to two. RGB-LED lights can produce a large range of colors and in theory provide a significant number of identifiable inputs. However, the actual number of identifiable colors is heavily dependent on the quality of the image sensor in the camera. Currently we have only experimented with two colors, but have found good results.

It is possible to modulate the brightness of the laser pointers to create identifiable input [10]. However, hardware modifications and microcontroller programming are required to add brightness modulation to a standard laser pointer. This is again at odds with our goal of using easily-acquired hardware. Camera framerate limits the number of pointers that can be identified, although faster cameras are available at significantly increased cost. We have not attempted to introduce brightness modulation into our system at this time, although combining the technique with color recognition would significantly increase the range of identifiable pointers. In addition, using programmable RGB-LEDs, the color of an individual LED can also be modulated.

A key technical issue with blob tracking is being able to avoid external bright spots that can confuse the tracker. One source of bright spots is the reflection of the on-screen image itself. However, the cameras support variable exposure, and simply by turning down the exposure we can completely hide the on-screen image. Light from the lasers or LEDs coming directly into the box is much brighter than the reflection and stays visible (See Figure 7).

Another source of unwanted light is the reflection of the projector bulb itself, which is only visible when viewing the projection screen from the rear (See Figure 7). This reflection is very bright, white light, completely obscuring all other incoming light. Background subtraction schemes are not possible because the light is white. It is possible to block out this bright spot, and in fact all projected light, by placing a circular polarizer over the projector lens and another circular polarizer over the camera lens. However, all the circular polarizers we have found that are large enough are plastic, and they melt due to heat from the projector light. They also dim the projected image significantly, which makes them an unappealing solution.

Our solution makes use of the two cameras mounted in each box (Figure 6). We position each camera so that it views the opposite side of the screen. Since the location of the reflected bright spot is dependent on viewing position, it is not visible to either camera if they are properly oriented. Because the cameras are viewing a skewed image, a portion of the camera pixels do not cover screen pixels and are in a sense “wasted”. However, with two 640x480 cameras in each box and only 1024x768 screen pixels, we can afford to sacrifice some pixels. Our tracking algorithm performs a weighted average of the segmented blob pixels, achieve sub-pixel accuracy. The tracker output is very accurate.

Figure 7 View from inside a MAD Box. Central bright spot is reflection of projector bulb. Top right red dot is laser pointer spot. Bottom left spots are white and red LED lights**.**

4 Conclusions and Future Plans

The MAD Boxes have satisfied most of our design goals. They are modular, stackable and reconfigurable. Several configurations are shown in Figure 8, including an 8 box wall, two 3 box towers (one starting at floor level and one elevated and used as a bulletin board at an art gallery showing), a 4 box table, and a 4 box counter. One could imagine such a counter in one’s breakfast nook and eating cereal while surfing current news.

The price we pay for modularity is the inter-screen gap. It is possible that the1-2mm (~2-3 pixel) gap could be further reduced with more complex manufacturing processes. It is uncertain that this gap has a different effect than the slight brightness and color variation visible in other display wall projects. A wall-length whiteboard commonly has several seams. Many people have commented on how “you hardly notice the edges”. We have noticed an occasional tendency to group content on individual screens, however the same behavior is commonly observed on whiteboards [7]. In addition, evidence gathered in [6] shows that removing the gap pixels largely mitigates any discomfort for users. We have yet to try this approach.

There are some issues with our current MAD Box design. Due to assembly errors the floating screen mounts do not always sit in the same position if they are removed and replaced. Also, the current aluminum frames weigh approximately 70 pounds each, making reconfiguration at minimum two-person job. Both of these issues will be dealt with in future design iterations.

Our color-recognition-based input system is currently functional and improving. Much work remains to be done on the host-computer side. Currently, the Microsoft Windows operating system does not support multiple cursors. By integrating our tracker with libraries such as the SDG Toolkit (http://grouplab.cpsc.ucalgary.ca/software/SDGT/), multiple cursor functionality can be made readily available in client applications.

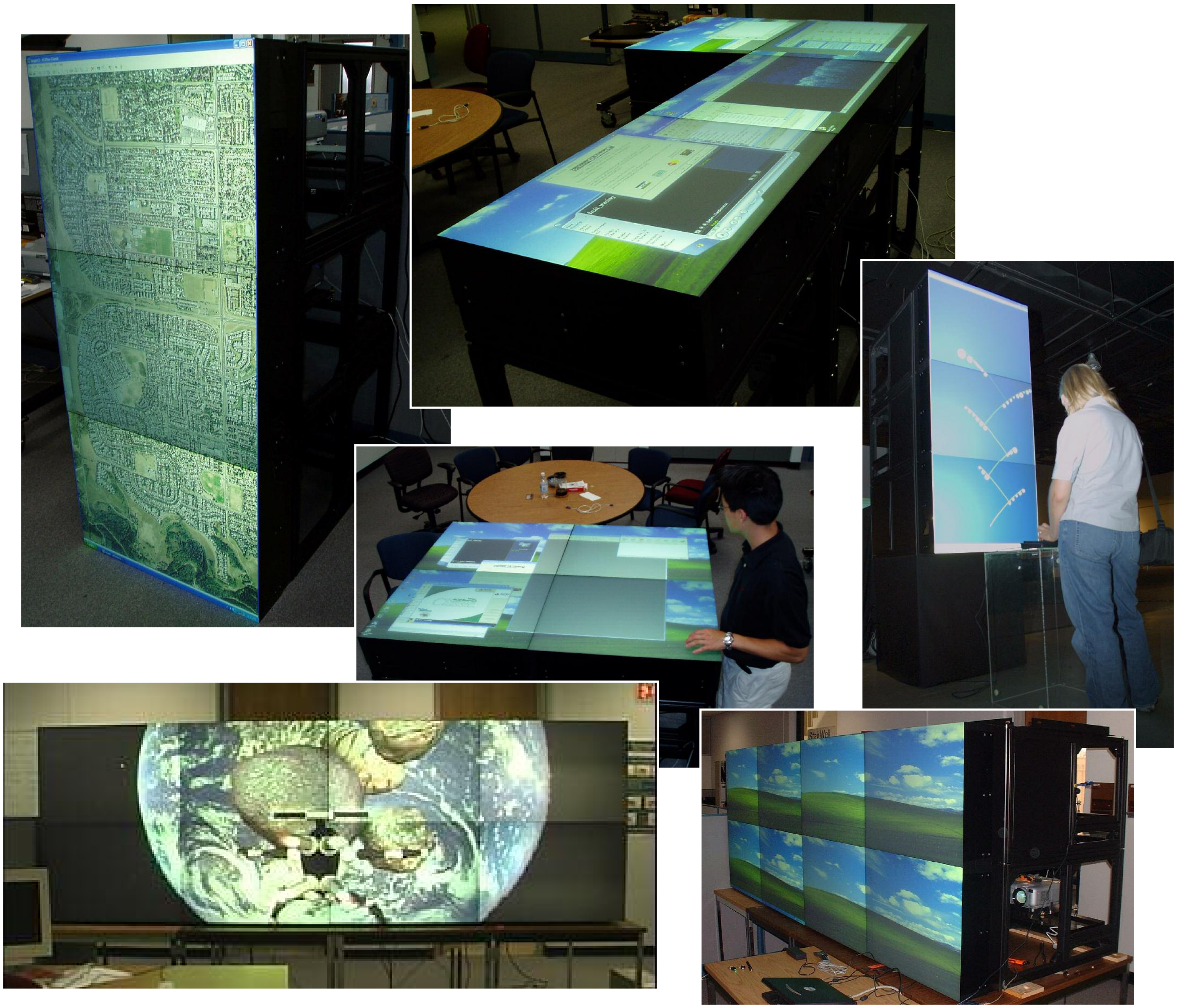

Figure 8: A variety of display configurations. Note that using the boxes upright requires additional cooling for projector safety.

Acknowledgements

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Canadian Foundation for Innovation (CFI).

References

1. Cao, X., and Balakrishnan, R.: VisionWand: Interaction techniques for large displays using a passive wand tracked in 3D. In *Proceedings of UIST 2003 – the ACM Symposium on User Interface Software and Technology*, 2003, pp. 173-182.
2. Chen, H., Sukthankar, R., Wallace, G., Li, K.: Scalable Alignment of Large-format Multi-projector Displays Using Camera Homography Trees. In *Proceedings of IEEE Visualization*, 2002.
3. Chen, X., Davis, J.: LumiPoint: Multi-User Laser-Based Interaction on Large Tiled Displays. *Displays*, **23**(5), 2000, pp. 205-211.
4. Hereld, M., Judson, I.R., Stevens, R.L.: Introduction to Building Projection-Based Tiled Display Systems. *IEEE Computer Graphics and Applications*, **20**(4), 2000, pp. 22-28.
5. Li, K., et al. Building and Using a Scalable Display Wall System. *IEEE Computer Graphics and Applications*, **20**(4), 2000, pp. 29-37.
6. Mackinlay, J.D., Heer, J.: Wideband Displays: Mitigating Multiple Monitor Seams. In *Extended Abstracts of the ACM Conference on Human Factors and Computing Systems*, 2004, pp. 1521-1524.
7. Mynatt, E., Igarashi, T., Edwards, W., LaMarca, A.: Designing an Augmented Writing Surface. *IEEE Computer Graphics and Applications*, **20**(4), 2000, pp. 55-61.
8. Ringel, M., Berg, H., Jin, Y., Winograd, T.: Barehands: Implement-Free Interaction with a Wall-Mounted Display. *CHI 2001 Extended Abstracts*, 2001, pp. 367-368.
9. Shedd, B.: Exploding the Frame: Designing for Wall-Size Computer Displays. In *Proceedings of IEEE Symposium on Information Visualisation (INFOVIS '03)*, 2003, 32.
10. Vogt, F., Wong, J., Fels, S.S., Cavens, D.: Tracking Multiple Laser Pointers for Large Screen Interaction. In *Extended Abstracts of ACM UIST*, 2003, pp. 95-96.
11. Wallace, G., Chen, H., Li, K.: DeskAlign: Automatically Aligning a Tiled Windows Desktop. *IEEE International Workshop on Projector-Camera Systems* (PROCAMS '03), 2003.