Body-Centric Interaction Techniques for Very Large Wall Displays

Garth Shoemaker[1], **Takayuki Tsukitani**[2], **Yoshifumi Kitamura**[3], **Kellogg S. Booth**[1]
[1]Department of Computer Science
University of British Columbia
{garths, ksbooth}@cs.ubc.ca
[2]Human Interface Engineering Lab
Osaka University
tsukitani.takayuki@ist.osaka-u.ac.jp

[3]Research Institute of Electrical Communication
Tohoku University
kitamura@riec.tohoku.ac.jp

ABSTRACT
We examine the design space of interaction techniques for very large wall displays by drawing from existing theory and practice for reality-based interfaces and whole-body interfaces. We also apply insights drawn from research in psychology about the human cognitive mechanisms that support sensorimotor operations in different coordinate spaces, as well as research in sociology examining how people manage coordination and privacy concerns in these spaces. Using guidelines obtained from these analyses, we designed and implemented a novel suite of body-centric interaction techniques. These were integrated into a map browsing and editing application for a very large (5m×3m) wall display. The application was then used to gather user feedback to guide the further development of the interaction techniques.

Author Keywords
Embodied interaction, gesture-based interaction, multimodal, reality-based interaction, post-WIMP interfaces, proxemics.

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation: Input devices and strategies

INTRODUCTION
Most computing systems designed in recent decades support one scenario of use: a single user sitting at a desk with a mouse and keyboard viewing a display approximately 50cm in size. Many aspects of human-computer interaction, including input devices and the WIMP (windows, icons, menus, pointer) model of interaction, have evolved to support this scenario. However, we are on the cusp of a new era in computing, where many different form factors will support a wide variety of interactive scenarios, supplementing traditional computer systems. Very small handheld computers are already becoming commonplace, and soon very large displays will be affordable enough that they too will be widely available. Unfortunately, interaction techniques designed for traditional computers are not always appropriate for use with other form factors, especially for very large displays. In order to benefit fully from the deployment of very large wall displays, we must overcome the challenge of designing interaction approaches that will take full advantage of the properties specific to these new form factors.

Figure 1. Users performing text and sketching input with the system. The system was used to explore body-centric interaction techniques.

In this paper we describe a *body-centric* model of human-computer interaction applied for use with very large wall displays. This model of interaction extends two general design philosophies: reality-based interaction [14] and whole body interfaces [15]. We distilled these into specific design guidelines using additional insights obtained from the psychology and sociology literature. Employing this framework, we explored the design space through a suite of prototypes for new interaction techniques. We evaluated the new techniques in the context of an interactive mapping application (Fig. 1).

Our major contributions are as follows: We first formulate a new body-centric design approach specific to very large wall displays. We then define an implementation framework that supports body-centric interactions using available hardware. Next, we describe a prototype application illustrating our body-centric approach through a suite of interaction techniques. We conclude with an evaluation of the system that informs subsequent design iterations.

NordiCHI 2010, October 16–20, 2010, Reykjavik, Iceland.

RELATED WORK

Before designing new interaction techniques for large wall displays it is helpful to first understand how large physical work surfaces are used. Then we can consider general design philosophies that we can draw from, and specific interaction techniques that other designers have developed. Only then should we synthesize a design framework to guide the development of new interaction techniques.

Understanding Large Work Surfaces

Large physical surfaces such as whiteboards and tables play important roles in everyday life. For example, large surfaces are ubiquitous in the field of education. It has been argued by Buxton that the widespread adoption of classroom blackboards in the early 19^{th} century was a critical advancement in educational technology [2]. He notes the irony that while blackboards replaced personal slates because their larger surface areas were deemed superior for teaching groups of students, there is currently a movement to replace large shared blackboards with smaller, single-user laptops.

Large surfaces are commonly used for a wide variety of other tasks, including managing personal information storage [28], supporting brainstorming activities [3], and supporting casual engagement in public spaces [27]. The properties of large surfaces that lend themselves to these kinds of tasks have been widely discussed in the literature. Rogers and Lindley compared the strengths and weaknesses of large horizontal (table) and vertical (wall) displays [21]. They found that wall displays allow collaboration of dynamically changing groups, support presenting to an audience, and properly orient information for all viewers.

Despite the benefits promised by large wall displays, public adoption is lagging that of other new form factors, notably handheld devices. As observed by Rogers and Rodden, this is not due to any inherent fault of large displays, but is due to the limitations of current hardware [22]. In anticipation of better hardware, we should strive to establish effective interaction approaches, so that when these systems are widely deployed users can immediately gain full benefit.

General Design Philosophies

Reality-Based Interaction (RBI) is an emerging paradigm in interaction design. Jacob et al. [14] identify a number of important themes within RBI, including *naive physics*, *body awareness & skills*, *environment awareness & skills*, and *social awareness & skills*. These themes provide a basis for interaction shared by many people from many cultures. Jacob et al. further argue that as interfaces have progressed from command line to WIMP-based models they have moved much closer to everyday interaction in the real world, but that there is still ample opportunity to adopt additional real world characteristics for the virtual world. They caution, however, that a virtual interface must retain some artificial or unrealistic features in order to be of value. An interface that mimics reality exactly will provide no benefit beyond what reality offers. Making the tradeoffs between the familiarity of reality and the power of "more than real" interfaces is a decision that must be made by designers.

A particularly relevant subset of reality-based interfaces is that of whole body interfaces. Klemmer et al. [15] note that our bodies possess a rich set of abilities that transcend what we can express symbolically. For example, we are able to perform the many complex operations involved in riding a bicycle, but can only explain them in general terms. These abilities are largely ignored by traditional keyboard and mouse interfaces, which make use only of the fine motor skills of the fingers and hands. Klemmer et al. identify five themes: *thinking through doing*, *performance*, *visibility*, *risk* and *thick practice*, that they believe are relevant to interactive systems. They challenge designers to draw on these themes in the realization of interactive systems, so we can better integrate the physical and computational worlds.

Specific Body-Based Interaction Techniques

The ideas just discussed lead naturally to an examination of how the human body can function as mediator in human-computer interaction. In this section we examine various ways that the body has been utilized in interactive systems.

Artists have been at the forefront of exploring whole body interaction. An early example is VIDEOPLACE, by Krueger et al. [16], which supported interaction using a virtual shadow of the user's body on the display. More recently, Lozano-Hemmer has explored shadows of various forms, including his "Shadow Box" series, and his "Under Scan" installation [17]. These works share the goal of breaking down the barrier between one's personal space and the shared space upon which the shadow is cast. This is a form of *expressive embodiment* as defined by Gutwin and Greenberg [10].

Researchers developing interactive systems have also made use of shadows in various ways. The "Shadow Communication" system used shadows to facilitate remote collaboration [19], whereas "Shadow Reaching" used shadows for co-located collaboration [23]. The motivation for both systems is similar to those of the artistic installations, with the additional requirement that they support traditional computing tasks. Each implementation does so differently, either supporting awareness only, in the case of Shadow Communication, or supporting awareness in addition to object pointing and manipulation in the case of Shadow Reaching.

There has been work that has utilized other aspects of the body in interaction. For example, Strömberg et al. developed a group game that sensed users' location in a room through pressure sensitive floor tiles [25]. In contrast, Harrison et al. developed a new sensing technology that employs a user's skin as an input surface [12].

Common to all of these systems is that they each leverage a small subset of body properties to support interaction. The shadow systems make use of a body contour, while the other systems model either user location in the room or touch on skin. None of the systems make use of a whole body model of the users. In order to fully respond to Klemmer's call to leverage embodied engagement we must capture the broad benefits made available by the body; the systems described are a good start, but there remains unexplored potential.

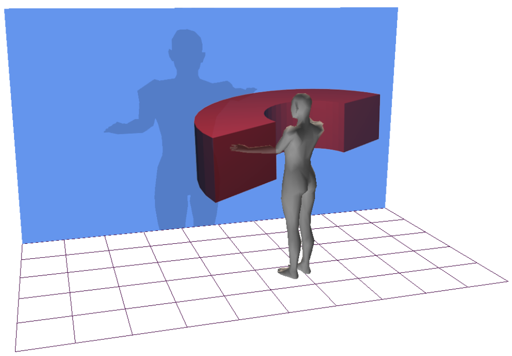

Figure 2. Peripersonal space (red volume) is that which is in reach of the arms, and often does not intersect with a large display during use.

DESIGN PRINCIPLES AND GUIDELINES

Our design framework extends reality-based interaction, as introduced by Jacob et al. [14]. In particular, we emphasize the themes of *body awareness & skills* (BAS) and *social awareness & skills* (SAS). We augment this with the whole body approach described by Klemmer et al. [15], and the five themes they explored. The combined themes are then extended and codified using some specific insights on human sensing and manipulation mechanisms obtained from the fields of psychology and sociology. This synthesis leads to our guidelines for designing body-centric interaction techniques for large wall displays.

Unifying Interaction Spaces

Neuropsychologists discovered that the brain builds multiple representations of space to coordinate sensorimotor operations [5]. Three particular representations are of interest for our design context: *personal space*, *peripersonal space*, and *extrapersonal space* [13]. Personal space is that occupied by the body, peripersonal space is that which is within easy reach of the hands (Fig. 2), and extrapersonal space is that which is outside of one's reach. Although on a conscious level we don't always distinguish between these spaces, the brain possesses separate mechanisms for operating in each of them. These mechanisms result in distinct performance characteristics when operating in the different spaces. In particular, interaction in personal (i.e. body) space is highly optimized, because that is where interaction in the physical world is performed.

One important design implication relates to the "binding" of spaces. A binding of spaces occurs when the brain's mechanisms for operating in one space are able to operate in a second space. Studies have found that in the physical world the brain is naturally able to bind personal and peripersonal space [29]. This allows us to efficiently reach out and grasp an object in our immediate vicinity. However in interactive computer systems it is also desirable to bind extrapersonal and personal spaces, because these systems support interaction beyond physical reach using laser pointers or other devices. If we can bind these spaces we might leverage the brain's highly optimized mechanisms for interaction in personal space.

Researchers have found two methods of binding personal and extrapersonal space. First, Pavani and Castiello have shown that human body shadows bind the two spaces [20]. They conclude that a person's "body schema" extends to include the body's shadow. They note that this can enhance a person's ability to interact in virtual environments. It has also been shown that a mirror can serve to bind extrapersonal space to personal space [18]. From this analysis of binding we are able to formulate our first design guideline:

D1 Where a large display system supports interaction at a distance, the interaction should be mediated through a representation that binds personal and extrapersonal space.

Not all interaction need be performed in the space of the display. The human body itself can play an important role. Proprioception is a person's sense of their own body in space, using information gathered from muscles, skin, and joint receptors [9]. Cocchini et al. showed, using a "fluff test" of experiment participants removing stickers from their own body, that the brain has a separate mechanism for governing proprioceptively-guided self-touching [4]. It has also been shown that "eyes-free" proprioceptive reaching can outperform vision guided reaching [7]. We conclude that proprioceptively guided reaching in personal space can augment parallel observation in extrapersonal space, and formulate a second design guideline:

D2 Leverage the sense of proprioception by allowing some operations to be performed in the user's personal space without reliance on visual feedback.

Supporting Natural User Inter-Relations

Humans follow complex social rules that coordinate interrelationships. Our framework explicitly recognizes the need to leverage how users naturally coordinate with other users.

One important aspect of inter-user coordination is how people position themselves relative to one another during work. As Felipe and Sommer explained, there is a universal cross-cultural concept of *private space*[1] [8]. Every person has a region of private space circumscribed around their body outside of which they attempt to keep other people. It is only during direct collaboration that a person will comfortably allow another to enter into their private space. As described in a review by Sundstrom and Altman, however, the concept of private space is more complex and fluid than the simple dichotomy of private/non-private [26]. In their model, the acceptable distance between two people is dependant on the shifting factors defining the interpersonal relationship. Using knowledge of "private space," a computing system can use the distance between users to draw conclusions regarding coordination, including whether or not users are directly collaborating. We thus have a third design guideline:

D3 Interaction techniques should respect user models of private space, and when possible take advantage of them.

[1] "Private space" in this context is sometimes referred to in the literature as "personal space." We call it "private space" to disambiguate from the other definition of "personal space" used here.

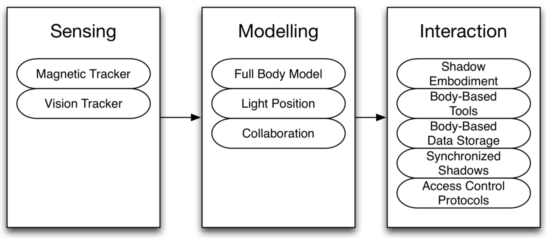

Figure 3. Various sensing, modelling, and interaction components explored in this paper.

Inter-relationships are much more complex than what can be captured by proximity alone. It has been shown that cues such as eye contact, body lean, and smiling are all important in communicating trust and managing coordination [1]. These cues can be difficult for a computing system to capture, due to limitations in sensing. They can nevertheless be leveraged by developing interaction techniques that incorporate direct user-user interactions. We thus have a fourth design guideline:

D4 Where possible allow users to make direct use of body cues such as facial expression and posture in order to help manage coordination.

CAPTURING THE BODY MODEL

In order for a system to implement the full "body-centric" design philosophy, or follow the specific design guidelines from the previous section, the system must maintain a geometric model of where users are in the workspace, ideally including limb poses. The more detailed and more accurate this model, the richer are the interaction techniques that can be built using it. We describe here the technical details of one implementation for developing a virtual scene model comprised of users in a room and relevant displays. This approach does not produce a perfect model, but it is practical to implement, and provides enough detail to support the development of some novel interaction techniques.

We divided our system into three modules: *sensing* and *modelling* components that maintain such a virtual model, in an approach similar to that described by Shoemaker et al. [23], and *interaction* components that define the behaviour of specific interactions (Fig. 3). Segmenting the architecture into components introduces modularity into the design, allowing for an arbitrary combination of different components.

We describe first a few of our implementations for *sensing* and *modelling* components (Fig. 4). These components offer superior robustness, increased flexibility, and ultimately produce a much richer model than was explored in previous work. We then describe specific interaction techniques that were built using these modules.

Components in the Sensing Module

Sensing components produce the raw data used for generating a virtual model of the scene. A typical implementation interfaces with a capture device (e.g. cameras or sensors) in order to generate data, such as 3D coordinates, that can be processed by a modelling component. We have developed two sensing modules that perform real-time measurement of the three-dimensional locations of body joints.

Magnetic Tracking Component

This component utilizes magnetic position sensors (Polhemus Liberty Latus). These sensors do not suffer from occlusion problems, because no line-of-sight is required. The effective range from each tracking station is approximately 2 meters, but a number of tracking stations can be placed over an area to increase coverage. The main disadvantage of magnetic markers is management of the active markers. They must be calibrated and batteries must be changed roughly every half hour.

Vision Tracking Component

The vision component tracks coloured balls attached to the user's joints. Multiple fixed cameras triangulate the position of each ball. The main strength of this approach is that the markers are passive, with no upkeep required. The cameras can be set to run continuously, and a user can start using the system without any calibration or initialization. Two weaknesses of this approach relate to occlusion and lighting. When a marker is hidden from the camera it can't be tracked, and varying lighting conditions change the colour of the marker as seen by the cameras, making identification difficult.

Components in the Modelling Module

One type of modelling component builds a virtual model of users and displays, using as input data from one or more sensing components. In our implementation the locations of the hands and shoulders, along with length and rotation constraints for limbs and joints, are used with an inverse kinematic (IK) solver to derive a complete skeleton of the user. We have found that inputting only hand and shoulder positions into the IK solver produces an adequate approximation to support our interaction techniques. For example, an approximate elbow location is usually accurate enough to be not noticeably different from its actual position.

Displays in the environment are assumed to be fixed, and thus we have not needed real-time updating of these models. Their locations are measured beforehand and are modelled as static rectangles in the room. If the displays were mobile they would need to be tracked in a manner similar to users.

For generation of user shadows and shadow-based interactions, a model of the light sources in the room is also maintained. Our implementation models one virtual light source for each user, which moves according to one of several lighting behavior models. The lighting behavior models take input from all other models in the scene (i.e. displays, users, lights) and output the location of the light source associated with each user. Because individual light behavior models are associated with particular interactions, the specifics of the behaviors will be discussed later.

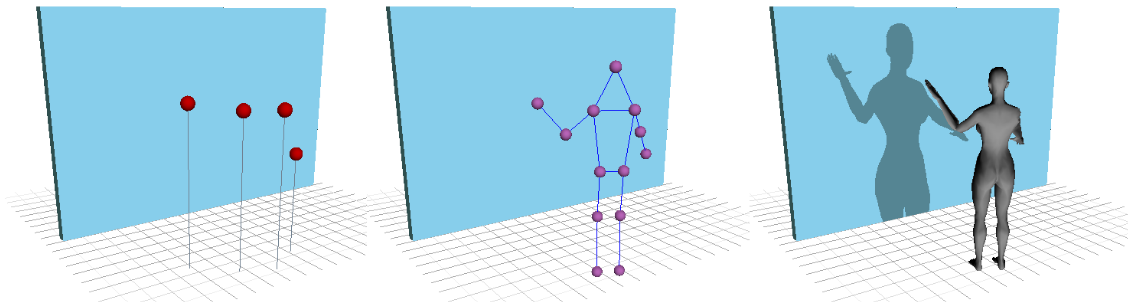

Figure 4. The sensing/modelling/interaction pipeline. From left to right: sensed locations of key joints, construction of an approximate skeleton based on known joint locations, and rendering of a human mesh and corresponding shadow. The images shown are from real data displayed in a test suite.

APPLICATION CONTEXT

The display used in our exploration is $5m \times 3m$ in size, rear projected by a 4×3 array of 800×600 pixel projectors. Neighbouring projected images are blended by 160 pixels, for a total display resolution of 2720×1480 pixels. There are no physical bezels and with proper calibration there are minimal visual artifacts in the blending regions.

To test our interaction techniques we implemented a map viewing and editing application. Click events to the application are performed using two Nintendo Wiimote controllers, one held in each hand. The application supports a number of features that help reveal how our interaction techniques interoperate, and how they function in the context of a real task.

Panning + Zooming

Users can pan the map and perform smooth zooming operations. We chose not to implement rotation because, unlike tabletop displays, orientation has special meaning on vertical surfaces. For maps, north is typically in the "up" direction.

Sketched Annotations

A user can perform free-form sketched annotations. The annotations are geo-referenced, and grow or shrink as the user zooms in or out.

Text Annotations

Text annotations can be entered using a soft keyboard. Text locations are also geo-referenced, but the text remains a constant size in order to remain legible.

SINGLE USER INTERACTION TECHNIQUES

As previously described, the system-maintained scene model, including body models of all users, supports the development of body-centric interaction techniques. We describe here several interaction techniques that make use of these models. Each technique functions by querying the state of the body models. For example, a technique can query the 3D location of a user's shoulder, the orientation of a user's body, or the distance between two users. Several of these techniques were previously demonstrated in video format [24]; we describe them more fully here.

Virtual Shadow Embodiment

A virtual shadow of each user is generated in order to provide expressive user embodiment and to bind personal and extrapersonal space, as recommended by design guideline **D1**. The shadow also supports direct interaction with on-screen elements. Interaction is performed using a cursor attached to each shadow hand, triggered by buttons on the corresponding handheld input device. The shadow is generated using the 3D geometry of a human mesh mapped to the 3D joint locations of the body model. A virtual light source can be placed at any location in the room, allowing an accurate shadow to be cast from any perspective.

The shadow embodiment component is capable of rendering several different shadow types, including: sharp shadows, soft edged shadows, outline shadows, and realistic transparent models. It can also render a Magic Lens style visualization. Renderings are implemented using the High Level Shader Language (HLSL).

Body-Based Tools

Body-based tools are virtual tools that are stored at real physical locations on the user's body (Fig. 5). To enter a mode or select an option in an application, the user places a hand at the corresponding body location and presses a button. This technique follows design guideline **D2**, allowing interaction in the user's personal space and leveraging the proprioceptive sense. Compared to traditional toolbars and tool palettes this approach has several benefits. First, the user can select known tools without having to perform a visual search and targeting operation. Second, a user's tools automatically follow the user and are always available, but don't clutter the display. Third, in collaborative scenarios there is no confusion regarding who controls what tool, because each tool clearly corresponds to a single user's shadow. These advantages can simultaneously improve tool selection performance and reduce confusion.

In our implementation, body tools are normally not visible, but their visibility can be triggered through the press of a button on the controller. The user can then hover over a tool and press a second button to select the tool. In cases where the user knows where a tool is located the user can select it without making it visible.

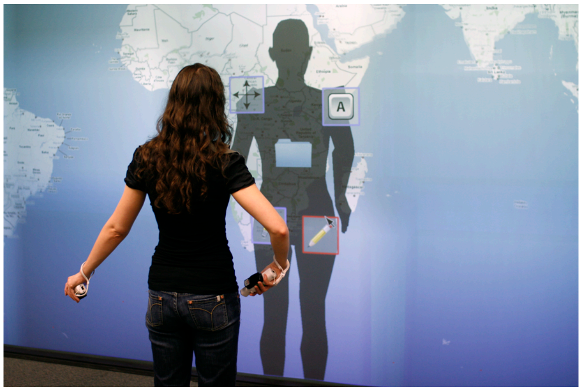

Figure 5. A user reaches her right hand towards her right hip to access a tool. This mechanism allows for immediate eyes-free tool selection regardless of user location in the room, and leverages the proprioceptive sense. Confusion in a collaborative scenario is also minimized.

Body-Based Data Storage

Body-based storage allows for convenient access to a user's personal data (Fig. 6). There are many situations in which a user may want to retrieve personal data, such as a PDF file or photo, and then show it on the shared display. Body-based data storage provides a body-centric metaphor and mechanisms for accessing and sharing this information, consistent with design guideline **D2**.

Each user's torso serves as a virtual container, from which personal data files can be accessed. This virtual storage is mapped to a user's computer or network drive. A user can use his or her hands to open, expand, and search through files virtually stored in the torso. When the desired file is found the user can extract the file from their torso and drag it to the shared space. This approach has many of the same benefits of body-based tools. First, personal files are always in close proximity and readily accessible to the owner, and second, there is little possibility for confusion regarding who "owns" which storage area.

There are several other advantages that are specific to the torso storage technique. Centering the navigation on the torso also centers it between the user's arms. This makes it easy for the user to interact with the data, which is important because navigation through a complex file space is not a trivial task. We also note that the torso is simultaneously the most massive part of a person's body, and the center of the person's body. The mass of the torso lends itself to being a metaphorical container for vast amounts of information. The fact that it is central to the body also makes it a personal part of the body, which associates well with the private nature of the data being accessed, and follows design guideline **D3**.

Visual feedback is provided through a data browsing widget in the form of a familiar hierarchical file browser shown in a grid layout. This is a suitable general purpose solution, however, if the application deals with only specific kinds of personal data, such as photos, a special-purpose widget could be designed.

Figure 6. A user accesses her personal files in her body-based data store. The user can search for and pull files of interest into the shared workspace. Navigation is managed by referring to the large display.

Body-Based Control Surfaces

Adjusting numeric values is a common task in any interactive system. In traditional UIs this is often done using 1D sliders or 2D widgets. Body-based control surfaces combine traditional easily understood widgets with a body-centered proprioceptive approach, following design guideline **D2**.

We implemented two different control surfaces (Fig. 7). The first is a body-based 1D slider. The ends of the slider are connected to specific body joints. The joints chosen are usually connected by a body part (e.g. elbow and hand connected by forearm). The user can adjust a single numeric value by sliding a hand between the two joints. Feedback is shown on the display, but using proprioception the user can avoid relying on the feedback. In our application we implemented a slider that adjusts the darkness of the user's shadow.

A 2D control surface can connect three or more joints. The surface visually connects the joints, and the user can adjust a multi-dimensional value by moving a hand over the surface. We implemented an RGB colour selector for adjusting the colour of sketch annotations.

Dynamic Light-Source Positioning

A single virtual light source is associated with every user, and the shadow cast of the user from the light source location onto the plane of the display is used to support interaction. Supporting dynamic light-source positioning can impact interaction in several meaningful ways. First, changing the projection of the shadow can allow the user to reach arbitrary locations on the screen. Moreover, altering the location of the light can be used to adjust the control-display (C/D) input gain, which can have a significant impact on pointing performance and error rates. C/D gain is a smoothly varying function dependent on light (l) and user (u) distances to the display ($gain = \frac{l}{l-u}$). We have developed several different light behaviours that govern how a light source moves (Fig. 8), based on the scene model.

User Following

This light behaviour allows for easy manipulation over the entire surface of a very large display, without requiring the

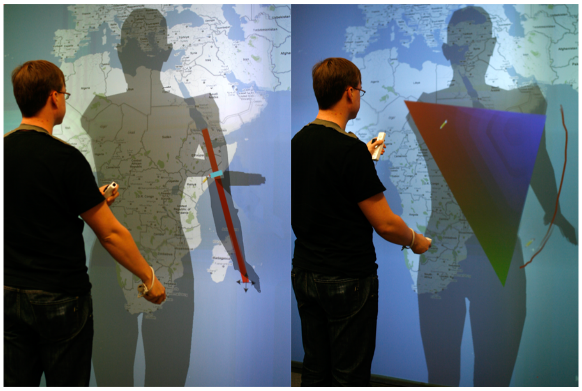

Figure 7. Left: A user adjusts an arm mounted slider. Right: A user selects a colour with one hand, and draws with the other hand.

user to walk around. Based on the known location of the user's shoulders, the behaviour places the light-source directly behind the user at a given distance. The result is that the user's shadow moves as the user turns, so that it is always directly in front of the user. This allows the user to perform continuous operations (such as dragging) across the entirety of a very large display, simply by turning his or her body.

Orthographic
This behaviour depends on the location of the user, and on the position of the display. The light source is placed at a very large distance directly behind the user, in a direction defined by the surface normal of the display. The result is a near-orthographic projection of the shadow onto the display.

The purpose of this behaviour is to provide a shadow mode of minimal distortion, with little risk of confusion. Confusion is minimized because the shadow is at the location on the display closest to the user. Close proximity minimizes the chance that the shadow will interfere with other users who are located elsewhere. The shadow also does not move when the user turns, which can also minimize confusion.

Manually Positioned
At times users may wish to manually position a light source. The user may, for example, wish to optimize the shadow for interaction in a particular region on a very large display. A manually positioned light also provides a very stable projection, which can ease detailed work.

A variety of approaches can be taken for supporting user control of the light source. In our implementation the user points in the direction where the shadow is to appear and presses a button. The light source is then positioned behind the user in the direction opposite to the direction pointed. The distance d_l between the light source and the user is a function of the distance d_h of the user's hand to the user's body. Because the user is restricted by arm length, the distance is exaggerated by the system. For example: $d_l = d_h^2 + c$. This approach allows the user to control both the location of the shadow and its size, and as a result the C/D ratio of the input.

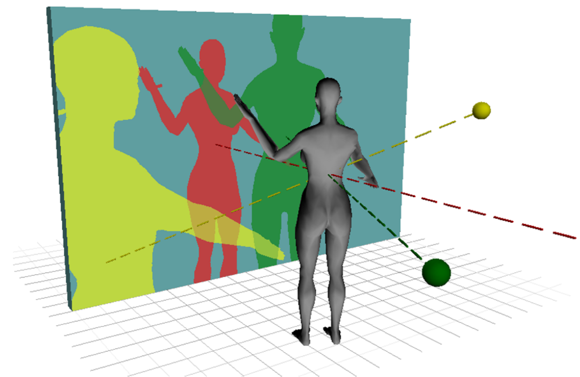

Figure 8. A user's virtual light source can be positioned based on different behaviors. Three such behaviors are shown, colour coded. Green: *user following*, Red: *orthographic*, and Yellow: *manually positioned*. Note that the colours are for illustrative purposes only.

Behavior Transitioning
This is a means of managing transitions between other behaviors. When switching from one behavior to another it is undesirable for the light source to jump instantly from one position to another. This can cause confusion for the user and collaborators. Instead, the system transitions from the position calculated by the old behavior function $p = f_o$ to the position calculated by the new behavior $p = f_n$ over a short period of time T by calculating a linear blend of the two functions $p = (1 - t/T) * f_o + (t/T)f_n$. This produces continuity of the shadow projection.

COLLABORATIVE INTERACTION TECHNIQUES
Large display systems are frequently used to support co-located collaboration, and ideally they should seamlessly support natural collaborative interactions. Although our current sensing and modelling approach focusses mostly on the *geometric* properties of users and environments, it is possible to extract an indication of collaborative intentions based solely on user geometry, and to further leverage this through specific techniques.

Synchronized Shadow Projections
When users are collaborating, inter-user coordination is a concern equal in importance to raw interaction performance. However, the importance of collaboration depends on how closely users are collaborating. Users positioned at opposite ends of a large display are likely working independently, whereas users positioned directly beside each other are likely collaborating closely. The synchronized shadows technique uses inter-user proximity, following design guideline **D3**, as an indicator of the degree of collaboration, and alters the shadow behaviour to change in a manner that supports each user's current collaborative state.

In the technique, when users are not collaborating closely, each of their shadows follows its own behaviour independently (e.g. user following). As two users approach and enter each other's private space, however, the shadows synchronize (Fig. 9). Synchronization means that the shadows alter their projection in order to be *consistent* and to *mini-*

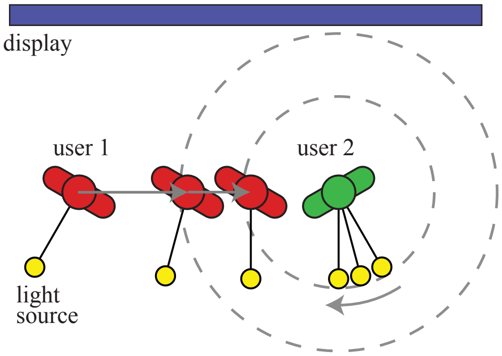

Figure 9. As user 1 enters user 2's outer and inner collaboration threshold light sources for the users transition from *user following* to *orthographic* behavior.

Figure 10. Private data is shared by the literal action of passing it to the other user's hand.

mize conflict. Consistency means that the shadows reflect a believable real-world lighting situation. For example, if user 1 is to the left of user 2, then user 1's shadow should also be to the left of user 2's. To minimize conflict, we enforce the condition that shadows not overlap. The more shadows overlap, the more likely it is that users will be confused.

Once the users are judged to be within collaboration range the system transitions to a lighting model consistent with the set of requirements. The orthographic lighting model fills these requirements: as users approach one another each of their lights transitions to the new model. Collaborative range can be defined as desired, but a good value is in the range of 45cm-120cm, identified by Hall [11] as being a typical radius for private space.

Access Control and Conflict Management

Management of private data is a concern in collaborative systems. Users must have a means not only of moving data between different privacy states, but the privacy state of all information artifacts must also be clear to users. We have built our access control protocols to center around the theme of *social awareness & skills*. We make use of standard social conventions to govern the handling of private data.

We enforce privacy by requiring all access to private data to take place in the literal body frame of reference (personal space), whereas access to public data takes place in the display's frame of reference. For example, in order for a user to move private data from body storage to the display, the user must first directly access that storage through their torso. Once the file has been moved to the shared display, however, it can be accessed in the display's frame of reference by any user. This follows design guideline **D3**. In another scenario, if user 1 wants to grant user 2 permanent access to a personal file, the user must physically and literally pass the file to the other user's hand (Fig. 10). Their hands must come in close proximity in order for the file to be passed. This protocol of forcing private information access to occur in personal space builds on a person's sense of their own private space, and also allows users to observe each other directly, making use of often subtle human cues to aid in the coordination of the sharing task. This follows design guideline **D4**.

PRELIMINARY EVALUATION

We described a design framework and a large set of novel interaction techniques. Fully evaluating all the techniques would require several controlled experiments, and is beyond the scope of this paper. We instead gathered preliminary user feedback from six users, with the goal of guiding future development. Each user was introduced to the different application features and interaction techniques, and was then given an opportunity to explore the system. To simulate a collaborative environment the experimenter served as a colleague. Notes were taken about user behaviour, and feedback was gathered both during and following the session. Each session lasted approximately half an hour.

All users seemed able to understand the concepts behind the interaction techniques. After one or two tries users were able to use the body-centric metaphor for tool selection, and similarly were able to navigate personal file space. Commenting on the body-centric approach in general, one user observed "You can't mess up!" The different lighting behaviors were also easily understood, as were the collaboration protocols. This suggests that basing interactions on real-world metaphors was a good decision. Nevertheless, there were several lessons learned that can guide improvements.

First, several participants commented that performance and realism are important in supporting the power of the shadow metaphor for interaction. The system exhibited occasional "hiccups," where there was an observable delay before rendering refresh. These delays broke the users' mental models of the reality of the shadow representation. There appears to be a threshold of accuracy that the shadow must achieve in order for the user to benefit from the embodiment and the binding of personal and extrapersonal space.

An interesting comment relates to tool placement. A participant asked if it was better to place commonly used tools on the left side of the body for a right-handed user, in order to make selection with the dominant hand easier. The answer is unclear, as it has been shown that a person is able to reach more accurately using proprioception with their left hand, if they are right-handed [6]. The difference between dominant and non-dominant sides in proprioception is something that should be further investigated.

Another issue that arose is that it was sometimes difficult for participants to remember the state of the two different hands. Each hand can be in a different mode, which is more complex than normal desktop systems where only a single cursor mode has to be remembered. It was suggested that the visualization could be improved to help users understand which hand is doing what. This is likely the best path to take, unless haptic feedback can be integrated to give direct information to each hand to help differentiate between modes.

Yet another comment centered on the physical device that was used. The Wiimote is designed to be held in a manner that suggests it is a pointing device, similar to a laser pointer. Unfortunately this is inconsistent with our approach, and caused at least one participant to attempt to activate a tool by *pointing* at a body part, instead of by *placing* the device at the body part. It is worth considering other input devices that do not present the affordances of a pointing device. An even better solution would be to improve the body model to a degree where an input device is not needed. Researchers have investigated input using hands in mid-air [30], and these approaches could be integrated into our system.

CONCLUSIONS AND FUTURE WORK

We have taken a body-centric approach to supporting interaction for very large wall displays. Our approach is inspired by the reality-based and whole body philosophies of interaction technique design. This allowed us to leverage themes such as *body awareness & skills* and *social awareness & skills*. Our goal in using this approach is to foster techniques that are, among other things, easy to learn to do, easy to interpret, and expressive.

We began by describing design principles that helped guide our work. This included the description of various interaction spaces, including personal, peripersonal, and extrapersonal. These different spaces segment the various frames of reference relevant to interaction with, and display of, different information artifacts. Careful design serves to bind these different spaces, and support interaction. We also examined some social rules of interaction that can guide the design of interactive systems.

We then described an implementation of a three-module *sensing*, *modelling*, and *interaction* architecture that enabled our interaction technique development. This implementation was a significant advancement beyond the architecture described in previous work, allowing for real-time calculation of a geometric scene model describing users and displays in the context of a shared interactive environment.

Based on the implemented architecture and our design principles we were able to develop a number of body-centric interaction techniques appropriate for use with various large wall displays. These include single user techniques for storing virtual tools directly on a user's own body, a technique for accessing personal information based on the metaphor of a user's torso as a container, body-based control surfaces, and a number of behavioural models for controlling the motion of a user's personal light source. We also developed several collaboration techniques, including a technique for synchronizing users' shadows to ease collaborative work, and a number of protocols for enforcing access control and managing conflict.

An important next step in our work is to support the development of a more fine-grained body model. While our model is holistic, in the sense that it represents the user's entire body in the context of the environment, it is not yet a very detailed model. Of particular importance is a more accurate model of the user's hands and fingers. Many existing interaction techniques rely on manipulation using individual fingers. We would like to integrate these techniques with our whole-body techniques. This would unify previous hand-specific work with our whole-body approach in a beneficial manner. Future work could also involve integrating new sensing techniques, such as improved vision algorithms for body tracking and the soon to be released Microsoft Kinect (formerly Natal) gaming system.

Another important next step is to extend the model to capture more than just the geometric properties of the scene. Models of mental process and intent could be very useful in guiding interaction techniques. We have made initial steps in this direction by modelling some collaborative protocols, but there is much work left to be done. Our immediate effort will center on developing new modelling modules.

In addition, we plan to continue developing new body-centric interaction techniques. This will involve the design of both new means of manipulation and corresponding feedback mechanisms. We will adapt our existing and future techniques with special consideration for multiple display environments, including handheld devices and tabletop displays.

ACKNOWLEDGEMENTS

Funding was provided by NSERC through operating and equipment grants, scholarships, and the NECTAR strategic research network, by the GRAND NCE, and by Defence Research and Development Canada. We thank Christine MacKenzie and Jocelyn Keillor for their invaluable input.

REFERENCES

1. J. K. Burgoon, D. B. Buller, J. L. Hale, and M. A. deTurck. Relational messages associated with nonverbal behaviors. *Human Communication Research*, 10(3):351–378, 1984.

2. B. Buxton. Surface and tangible computing, and the “small” matter of people and design. In *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pages 24–29, 2008.

3. M. Cherubini, G. Venolia, R. DeLine, and A. J. Ko. Let’s go to the whiteboard: How and why software developers use drawings. In *Proceedings of CHI ’07*, pages 557–566, 2007.

4. G. Cocchini, N. Beschin, and M. Jehkonen. The fluff test: A simple task to assess body representation neglect. *Neuropsychological Rehabilitation*, 11(1):17–31, 2001.

5. C. L. Colby. Action-oriented spatial reference frames in cortex. *Neuron*, 20:15–24, 1998.

6. S. H. B. Daniel J. Goble. Task-dependent asymmetries in the utilization of proprioceptive feedback for goal-directed movement. *Exp Brain Research*, 180:693–704, 2007.

7. W. G. Darling and G. F. Miller. Transformations between visual and kinesthetic coordinate systems in reaches to remembered object locations and orientations. *Exp Brain Res*, 93, 534–547.

8. N. J. Felipe and R. Sommer. Invasions of personal space. *Social Problems*, 14(2):206–214, 1966.

9. S. C. Gandevia, K. M. Refshauge, and D. F. Collins. Proprioception: peripheral inputs and perceptual interactions. *Adv Exp Med Biol*, 508:61–68, 2002.

10. C. Gutwin and S. Greenberg. A descriptive framework of workspace awareness for real-time groupware. In *Proceedings of CSCW ’02*, pages 411–446, 2002.

11. E. T. Hall. *The Hidden Dimension*. Peter Smith Publisher Inc, 1992.

12. C. Harrison, D. Tan, and D. Morris. Skinput: appropriating the body as an input surface. In *Proceedings of CHI ’10*, pages 453–462, 2010.

13. N. P. Holmes and C. Spence. The body schema and multisensory representation(s) of peripersonal space. *Cognitive Processing*, 5(2):94–105, 2004.

14. R. J. Jacob, A. Girouard, L. M. Hirshfield, M. S. Horn, O. Shaer, E. T. Solovey, and J. Zigelbaum. Reality-based interaction. In *Proceedings of CHI ’08*, pages 201–210, 2008.

15. S. Klemmer, B. Hartmann, and L. Takayama. How bodies matter: Five themes for interaction design. In *Proceedings of DIS ’06*, pages 140–149, 2006.

16. M. W. Krueger, T. Gionfriddo, and K. Hinrichsen. Videoplace – an artificial reality. In *Proceedings of CHI ’85*, pages 35–40, 1985.

17. R. Lozano-Hemmer. *Under Scan*. emda & Antimodular, 2007.

18. A. Maravita, C. Spence, C. Sergent, and J. Driver. Seeing your own touched hands in a mirror modulates cross-modal interactions. *Psychological Science*, 13(4):350–355, 2002.

19. Y. Miwa and C. Ishibiki. Shadow communication: system for embodied interaction with remote partners. In *Proceedings of CSCW ’04*, pages 467–476, 2004.

20. F. Pavani and U. Castiello. Binding personal and extrapersonal space through body shadows. *Nature Neuroscience*, 7:14–16, 2003.

21. Y. Rogers and S. Lindley. Collaborating around vertical and horizontal large interactive displays: which is best? *Interacting With Computers*, 16:1133–1152, 2004.

22. Y. Rogers and T. Rodden. *Configuring spaces and surfaces to support collaborative interactions*, pages 45–79. Kluwer Publishers, 2003.

23. G. Shoemaker, A. Tang, and K. S. Booth. Shadow reaching: a new perspective on interaction for large displays. In *Proceedings of UIST ’07*, pages 53–56, 2007.

24. G. Shoemaker, T. Tsukitani, Y. Kitamura, and K. S. Booth. Whole body large wall display interfaces. In *Extended Abstracts of CHI ’10*, pages 4809–4812, 2010.

25. H. Strömberg, A. Väätänen, and V.-P. Räty. A group game played in interactive virtual space: design and evaluation. In *Proceedings of DIS ’02*, pages 56–63, 2002.

26. E. Sundstrom and I. Altman. Interpersonal relationships and personal space: research review and theoretical model. *Human Ecology*, 4(1):47–67, 1976.

27. A. Tang, M. Finke, M. Blackstock, R. Leung, M. Deutscher, and R. Lea. Designing for bystanders: Reflections on building a public digital forum. In *Proceedings of CHI ’08*, pages 879–882, 2008.

28. A. Tang, J. Lanir, S. Greenberg, and S. Fels. Supporting transitions in work: Informing large display application design by understanding whiteboard use. In *Proceedings of GROUP ’09*, pages 149–158, 2009.

29. S. Vaishnavi, J. Calhoun, and A. Chatterjee. Binding personal and peripersonal space: evidence from tactile extinction. *Journal of Cognitive Neuroscience*, 13(2):181–189, 2001.

30. D. Vogel and R. Balakrishnan. Distant freehand pointing and clicking on very large, high resolution displays. In *Proceedings of UIST ’05*, pages 33–42, 2005.