

Research Report

Room With A View (RWAV): A Metaphor For Interactive Computing

Larry Koved
Ted Selker

IBM Research Division
T. J. Watson Research Center
Yorktown Heights, NY 10598

LIMITED DISTRIBUTION NOTICE

This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. It has been issued as a Research Report for early dissemination of its contents and will be distributed outside of IBM up to one year after the date indicated at the top of this page. In view of the transfer of copyright to the outside publisher, its distribution outside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g., payment of royalties).

Room With A View (RWAV): A Metaphor For Interactive Computing

Larry Koved
Ted Selker

IBM Research
T. J. Watson Research Center
Yorktown Heights, NY 10598

Abstract: The desktop metaphor demonstrates the utility of spatial separation of tasks. Room With a View (RWAV) extends this metaphor to make greater use of three dimensional spatial layouts and gestures. In this multi-surface user interface metaphor, the work surfaces, such as walls, shelves and tables, are projected with icons of books, files, etc. These icons represent information and other interactive media. Documents or drawings are displayed on a high resolution active blackboard, or on hand-held clipboard-sized interaction surfaces called ViewBoards. The user moves around the information space by gesturing with ViewBoards or blackboard chalk. RWAV's goal is to improve interaction with large amounts of spatially organized information.

Room With A View

Background

The desktop metaphor, introduced on the Xerox Alto computer [1], has since been embodied in a variety of window management systems. A basic premise of the metaphor is that we engage in a variety of tasks during the normal course of our work. The messy desk helps us to start work on one task, suspend it, and begin another task. Elements of the computer-based desktop are represented spatially through a set of icons, such as filing cabinets, folders, paint palettes, etc. These icons are indicative of the objects that we may have in a real office. It is fairly common for people to spatially organize their work, both on a work surface (two dimensional), and making piles of papers and books on the work surface (three dimensional). The RWAV metaphor extends the desktop by organizing information on multiple display surfaces.

The desktop metaphor has been used extensively as a user interface. However, many tasks are too complicated to be easily organized on a single screen. When there is a need to present substantial amounts of information, the desktop becomes cluttered and restrictive. The *Rooms* metaphor was created to mitigate this problem [2]. Rooms uses "doors" to move from "room" to "room". This represents moving from one desktop to another. Each room is used to contain separate projects and retain spatial layout on each desktop. While the Rooms metaphor extends the number of workspaces available for use, the window into the information space is still limited to a single desktop (computer display) at a time. The user must then remember which room contains the relevant information. Navigating the maze of rooms to locate the room with the desired information is problematic. The limited amount of display space available on the screen imposes a cognitive load that is perhaps unnecessary.

The desktop and Rooms metaphor are accepted as improving users' orientation in their tasks. Techniques for orienting people requires display real estate. A conflict arises between devoting real estate to the active workspace and presentation of the overall logical structure of the system or project. If more display space is devoted to detailed drawing or text, structural information must be condensed or obscured.

Another issue with current workstations involves problems in immersing the user in a project with a single screen. The rest of the office intrudes, the keyboard is shared and the display surface is limited. Improving immersion is an exciting research area people are exploring with interesting new hardware solutions. The "Put That There" project uses hand gestures, speech recognition and a wall sized display for human-computer interaction [3, 4]. A user points at parts of a display and asks verbal questions. The computer then responds to these commands. The project demonstrates how multiple I/O technologies can be combined to create a new interaction paradigm. The user is literally surrounded by their terminal. Unfortunately, as a workstation, the room gives the user no more spatial data than a terminal. Head mounted display projects have also attempted to give users a sense of immersion, such as the ability to walk through a building [5]. These have been successful at creating spatial immersion, but not at presenting detailed information. These projects employ new display hardware to improve users sense of orientation in information they are working with but have not increased a person's ability to structure and work with detailed information.

To get an idea of how to be immersed in a work situation and have structural information available, let us look at a normal office situation. I may leave a slip of paper on my desk. During the normal course of work, this slip of paper will be moved and end up under some

books and folders. When I need the information on the slip, I use a number of cognitive and motor skills to get access to the information. First, I recall the general location of where I put the slip of paper. I also remember the color and the general form of the writing patterns on the paper -- scrawling at the top is the axes for a graph. I was interrupted by a phone call and moved some books on top of the slip, and then put my new mail on top of the books. By lifting the books and mail, I am able to retrieve the slip of paper.

In contrast to the use of spatial and visual information, a logical search could be performed. Searching for a text string using Boolean operators may be fast, but generating the appropriate search criteria may be difficult and require several attempts. People are very good at remembering things they have seen in the past [6]. A search based on context, location and visual appearance may be more efficient, time wise, than a textual or graphical search.

In the above paragraph, I use the office surrounding me as an aid in organizing my work. Below we develop the office, its shelves, walls and counters as work surfaces to store structural or orienting information. Multiple displays can capture the process of starting, suspending and reorganizing work tasks. We describe a metaphor for using low resolution computer displays for organizing structural information and high resolution displays for manipulating detailed information.

Scenario

RWAV extends the messy desk metaphor. RWAV combines multiple display surfaces -- hand held displays, and walls or counter-sized displays -- for presenting visual output together with several different techniques for generating computer input. These use principles of direct manipulation [7]. The following is a scenario using the RWAV environment.

Early Monday morning, my office is dark as I approach the open door. Inside is a table and a few chairs. The walls and tables are bare. The exception is a set of objects that look like pads of paper or clipboards.

As I turn on the computer, the artifacts of my office are projected onto the walls and desktop. On the wall are some diagrams and notes for the project I was working on late Friday afternoon. Several stacks of "papers" and "books" are on the desk. These were the result of browsing through the library for articles relevant to a current project. On the wall is a calendar, clock, several personal photographs, a game and an icon of a hi-fi stereo. Pointing at the stereo turns on the morning news.

I grab one of the *ViewBoard* tablets. The ViewBoard is a computer display with a touch sensitive surface, a primary means for interacting with the computer. It acts as a panning, zooming and selecting device when it is aimed at a wall or counter top. The movement of the ViewBoard over the desk is representative of the panning function supported by the ViewBoard. To look at the schedule for the day, I aim the ViewBoard at the calendar on the wall. A high resolution image of the data represented by the icon appears on the ViewBoard. After panning the ViewBoard over to today's date, pressing the ViewBoard will make the selection. Aiming the ViewBoard at the counter and pressing again puts a copy of today's agenda on the counter. I notice an appointment to see Ted at 9:30. I aim the ViewBoard at an icon on the wall representing documents related to our project. To retrieve them, I press the ViewBoard and aim it at a spot on the table. Releasing the ViewBoard places the stack icon of documents related to our project on the table.

Along the wall above the desk is a set of icons. Among them is a "in-basket" indicating new mail has been received. As was the case with looking at the calendar, a ViewBoard is

used to get more detail about the icons on the wall. Taking a ViewBoard, I aim it at the in-basket and press. The mail headers are listed on the ViewBoard. To read the mail, I point my finger at the item of interest. For scrolling through the text and graphics, I run a finger along the edge of the display surface. To scroll down, I run my finger down index tabs that are displayed along the edge of the screen, and vice versa to scroll up. Within the mail are references to previous pieces of correspondence and related documents. To retrieve one, I point to the item of interest and it appears on the ViewBoard. If more display space is needed for reading, I can grab another ViewBoard and place it on top of the original ViewBoard, then remove it, to make a copy.

Ted walks into the office for our meeting. By pointing to an icon representing our project, the state of my work environment is saved and the desk is cleared. All of the material previously spread out is now neatly organized as an *office icon* off on the corner of my desk. The materials for the project Ted and I are working on are now displayed on the table and the wall. The various documents, sketches and notes are in the same locations on the wall and table as they had been during our previous meeting. The spatial organization was preserved when it was turned into an office icon after our previous meeting. When collaborating on the project, we can look at two separate ViewBoards or use a high resolution projected black-board.

Ted remembers some diagrams and new articles that are in his office. Since they are relevant to our discussions, he opens a window to his office. By pointing to an icon of his office on the ViewBoard, a "window" looking into his office is presented on the wall. With the ViewBoard, he is able to zoom into an area of his office for closer inspection. Pointing his ViewBoard at a stack of materials and pressing the ViewBoard, he is able to grab them and bring them into our meeting and close the window to his office.

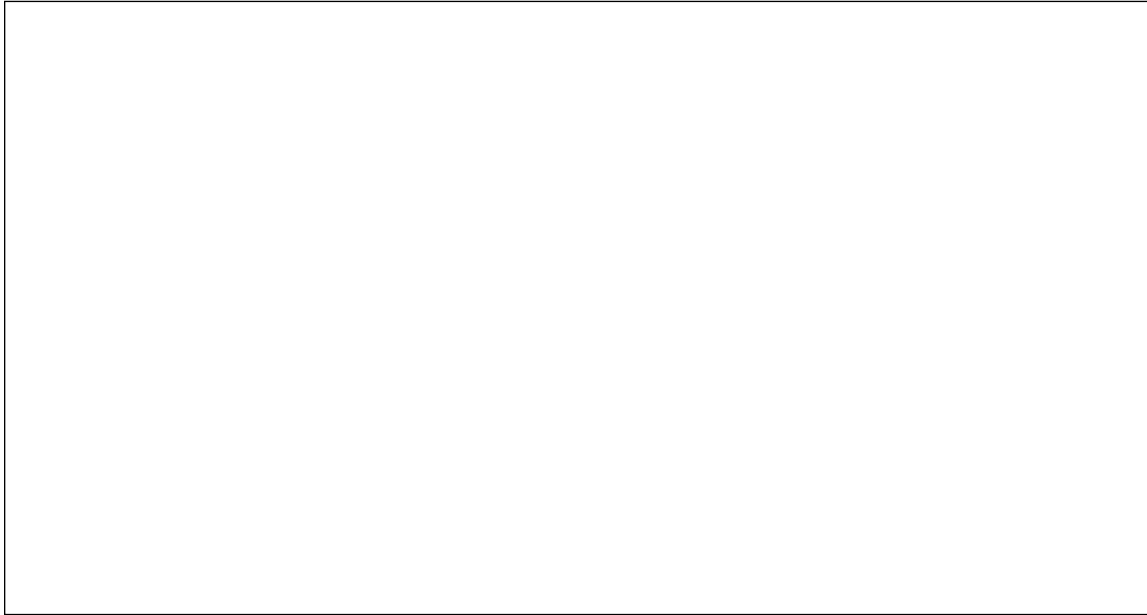
The meeting is enhanced by the RWAV technology. We use the ViewBoards to draw diagrams, search the library and our notes for articles and information. When we run out of room on a ViewBoard, we can grab another ViewBoard, transfer the information to an icon on the wall or file it into a folder. Cross references to our notes and library references are made by using gestures on the ViewBoard surface.

After Ted leaves, I continue to work on the diagrams until my manager arrives. The materials for the project with Ted are filed away as an office icon. Aiming a ViewBoard at a telephone directory, I scan for the phone number for the lab. Selecting the directory entry will place the phone call. Among the icons projected on the wall is a video insert showing a picture of Bryan when he answers the phone. Conversely, Bryan also sees a video image of us working in my office. My manager would like to see the figures on performance and latency in the system we are building. I grab a ViewBoard, aim it at the iconic image of the stack of documents that contain the numbers and press to retrieve the data...

Enabling the Scenario

Off the shelf technologies are integrated into a single system to enable the Room With A View scenario. These include flat panel displays, touch sensitive surfaces (i.e., touch screen technology), three dimensional space trackers, video or laser projectors, window managers, hypertext, and gesture recognition.

The ViewBoard is an LCD flat panel display with a touch sensitive surface. Some capabilities afforded by a touch sensitive surface include editing of documents, pointing and



A RWAV consisting of a table with two viewboards on top, and a rear projection video screen with icons.

selecting, and the creation of simulated buttons. To simplify the implementation, the ViewBoards are attached to the computer via an umbilical cord for power and data lines.

ViewBoards are panning and zooming devices when used in conjunction with images on the walls and counter tops. The computer stores the geometry of the room and objects within it, such as the location of the walls and counter tops. A three dimensional tracking device (Polhemus 3Space tracker) senses the position of the ViewBoard [8, 9]. The user holds a ViewBoard and moves it around the room, aiming it at a wall or counter top. The images that appear on the wall or counter are magnified and presented on the ViewBoard, creating the panning function. By pressing and moving the ViewBoard closer to the wall or counter, the image is magnified or zoomed, giving greater detail. Moving away from the wall or counter reduces the level of detail. Pressing a different part of the ViewBoard surface selects and locks an image onto the ViewBoard so that the user can move the ViewBoard without the panning/zooming functions enabled. The user can then work on the displayed objects, such as documents, spreadsheets and correspondence, directly on the ViewBoard.

The ViewBoard takes on many of the functions previously performed by a mouse, yet has greater flexibility. The 3D tracking capabilities provide 2D cursor tracking on the wall and counter displays. In addition, the extra degree of freedom enables panning and zooming functions that would not be as easy to accomplish with a mouse. Virtual buttons replace the mouse buttons. Because the buttons are on the display surface, the number of and definition of the buttons are dynamic and reconfigurable. Their function can be made readily apparent by changing how they are labeled.

When additional work space is needed, one ViewBoard can momentarily be placed on top of another to make a copy. With several ViewBoards, multiple documents can be spread out as though each were a piece of paper on a table. This is a key advantage over the traditional computer with only a single display -- as more display space is needed, another ViewBoard can be put to use. We see this as a key advantage over the Rooms scenario because multiple display surfaces are available. These displays can be spread out on a table or

floor to provide the necessary working space frequently needed when working on writing a paper, analyzing data or doing design work.

A single ViewBoard can be treated like a piece of paper, but takes on additional significance in the RWAV scenario when several are used together. ViewBoards can be logically linked to form larger documents such as books. For example, two side by side ViewBoards may represent an open book. When a scrolling or page turning gesture is performed on one ViewBoard, the image on both ViewBoards is updated.

Since ViewBoards are electronic, they provide shared views for multiple people in a group setting. If each person in a group has their own ViewBoard, each of the ViewBoards can have a shared image. As one person makes changes to the image on their ViewBoard, the image is updated on the other ViewBoards [10-13]. Similarly, changes on the ViewBoard can also result in changes to the image on the walls or counters.

The images on the walls and counters are projections of computer generated images. One of the simplest ways to accomplish this is to take the video output from a computer display adapter and feed it into a video projector. Using a rear projection screen will allow the user to stand in front of the wall without occluding the projected images. If necessary, the images can be reflected off of mirrors to reduce the projection space requirements.

Because the projected display is used for presenting structural information, resolution is not the critical factor. The surfaces appear to be bulletin boards with items tacked to them, books and objects on shelves, or objects hanging on the wall that can be pulled down for closer inspection. Glancing around a room the user finds and organizes projects in various stages of completion. These large surfaces give the user a frame of reference or context in which work is being done -- to display icons of books, papers, notes, etc. These create the immersive, three dimensional environment in which work is performed. Fine details are not critical. The use of the large display space is oriented more towards organizing information. Detailed work is accomplished through the ViewBoards.

The interaction between the ViewBoards and the large projection surfaces is a key component of RWAV. The three dimensional organization of the images spread across the room create a work environment. This provides the means for organizing information for accomplishing tasks. The ViewBoards are a tool for interacting with the information in this three dimensional space.

Status

A prototype of RWAV exists. A video projection system has been installed to provide the iconic images that would be projected onto a wall or counter top. These images are displayed onto a rear projection screen. Processes using the X Window SystemTM[14] provide the icons and text appearing on the projection screen and ViewBoard.

Some of the ViewBoard's rudimentary functions are provided, such as the selection function. When an icon is selected, the appropriate window appears on another display unit. Instead of using simulated buttons on the ViewBoard, a mouse is used. Additional functions will be added as the prototype progresses. By the time of the conference, the flat panel

¹ X Window System is a trademark of the Massachusetts Institute of Technology

display with a touch surface and the 3-D position tracking should be integrated into the system.

Summary

Traditional uses of interactive computing have been to replace typewriters, perform calculations, make drawings and organize lists of information. The desktop metaphor is a departure from these uses. It allows people to use the computer to organize items in their computer spatially. Due to space constraints, people have resorted to making lists in windows (i.e., contents of filing cabinets or folders). By allowing additional work surfaces, RWAV separates detailed work surfaces from spatial organization elements of the work area.

Room With A View extends the interactive computing metaphor beyond the desktop. It is a technique for organizing information in a three dimensional spatial layout more indicative of the ways in which we do our daily work. Information is presented and organized on the working surfaces around us -- the walls and counter tops, in addition to active work areas that we call ViewBoards. Through hand gestures, speech and other I/O technologies, we are able to communicate our intentions to the computing environment that surrounds us.

RWAV is a framework for working in large information spaces on a daily basis. It enables users to structure the spatial organization of the work environment to meet their needs. RWAV preserves the spatial organization of tasks between work sessions. When tasks are suspended, their spatial organization is preserved for later sessions.

Part of the motivation for RWAV is to provide a test bed for experimenting with new techniques for interacting with computing systems. In its current incarnation, RWAV allows us to look at the basic premise that three dimensional organization and presentation of information is useful.

References

1. C. P. Thacker, E. M. McCreight, B. W. Lampson, R. F. Sproull and D. R. Boggs. Alto: A Personal Computer. In D. Siewiorek, C. G. Bell, and A. Newell, editors, *Computer Structures: Principles and Examples*, McGraw-Hill, New York, 1982.
2. Stuart K. Card and Austin Henderson, Jr. A Multiple, Virtual-Workspace Interface to Support User Task Switching. *CHI + GI 1987 Conference Proceedings*, 198-192, ACM, April 1987.
3. Richard A. Bolt. *The Human Interface: Where People and Computers Meet*. Lifetime Learning Publications, Belmont, California, 1984.
4. Richard A. Bolt. 'Put-That-There': Voice and Gesture at the Graphics Interface. *Computer Graphics*, 14(3):262-270, 1980.
5. Frederick P. Brooks, Jr. Grasping Reality Through Illusion -- Interactive Graphics Serving Science. *CHI '88 Proceedings*, 1-11, ACM, May 1988.
6. R. N. Shepard. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, (6):156-163, 1967.
7. Ben Shneiderman. Direct Manipulation: A Step Beyond Programming Languages. *IEEE Computer*, (August):57-69, 1983.
8. Frederick H. Rabb, Ernest Blood, Terry O. Steiner and Herbert R. Jones. Magnetic Position and Orientation Tracking System. *IEEE Transactions on Aerospace and Electronic Systems*, AES-15(5):709-718, IEEE, September 1979.
9. Polhemus Navigation Sciences Division of McDonnell Douglass Electronics Company, 3SPACE® USER'S MANUAL, Colchester, Vermont, 1987.
10. International Business Machines Corporation, Cooperative Viewing Facility: Version 2 General Information (GC34-4005), Cary, N.C., 1985.
11. Gregg Foster. CoLab, Tools for Computer-Based Cooperation: A Proposed Research Program. University of California, Berkeley, UCB/CSD 84/215. December 1984.
12. Sunil Sarin and Irene Greif. Computer-Based Real-Time Conferencing Systems. *Computer*, 33-45, October 1985.
13. W. K. English and Doug C. Engelbart. A Research Center for Augmenting Human Intellect. *Proceedings of the National Computer Conference*, 395-410, IFIPS, 1968.
14. Jim Gettys, Ron Newman and Robert Scheifler. Xlib-C Language X Interface. Massachusetts Institute of Technology, Cambridge, Massachusetts. 1987.

Copies may be requested from:

IBM Thomas J. Watson Research Center
Distribution Services F-11 Stormytown
Post Office Box 218
Yorktown Heights, New York 10598