Implementation and capabilities of a virtual interaction system

Arthur W. Joyce, III and Anna C. Phalangas

Alfred I. duPont Hospital for Children / University of Delaware
1600 Rockland Road, Wilmington, DE 19803, USA

joyce@asel.udel.edu, phalanga@asel.udel.edu

ABSTRACT

Virtual Interaction refers to a technique for interacting with computer generated graphics. Graphical objects are overlaid on live video of the user. A chromakey separates the user (foreground) from the background resulting in a silhouette of user. The computer causes the graphical objects to move in relation to the silhouette so that they appear to interact with the user. This paper presents the implementations of the system, some techniques for interaction and discusses using the system as a tool for physical therapy.

1. INTRODUCTION

Virtual Interaction is the name of a low immersion, virtual reality system developed at the Alfred I duPont Hospital for Children and the University of Delaware. The users view themselves participating in virtual worlds, consisting of computer generated graphics, with real-time interactions. The purpose of this project is to create interactive environments, which can enhance therapy by providing motivating activity tailored to each user's interests and physical abilities. Virtual interaction provides physical activity for individuals with disabilities without requiring direct interaction with real physical objects. It also offers the potential for tracking clinically relevant performance information.

The system is moving into its third implementation. This paper presents these implementations and discusses the basic requirements of any implementation. It also discusses several interaction modes and their computational requirements.

2. BACKGROUND

In the early 1970s, Myron Krueger developed a system called VideoPlace, which provided a number of interactions between users and computer generated graphics (Krueger 1990). The users were back-lit to provide a strong contrast in the video signal. This almost binary signal was easily separated into the foreground/background regions needed for the interactions. Users saw their silhouettes interacting with the graphics.

Researchers at the Massachusetts Institute of Technology (MIT) have developed the IVE (Interactive Video Environment) and the ALIVE (Artificial Life IVE) systems (Maes 1995). These systems use color segmentation to separate the user's silhouette from the background without any special backdrops. The MIT research has focused primarily on intelligent, artificial creatures that react to the user's gestures. The users see themselves in their physical surroundings along with the computer-generated characters. The graphical characters are carefully generated so that they may move both in front of or behind the user. The gesture recognition and character responses occur over several seconds.

The Vivid Group (Toronto, Canada) currently markets the Mandala VR System that uses a blue backdrop and a chromakey to separate the user's image from the background. The blue background is replaced with a computer generated scene and interactive graphics. The interactions are immediate, video game like responses. The Vivid Group uses proportional control based on areas of interest to control the action in many of their games (MacDougall 1996).

Research being performed at the Alfred I. duPont Hospital for Children and the University of Delaware explores the use of this technology as a means of providing therapeutic and exercise activities for individuals with disabilities (Joyce 1995). The project, called Virtual Interaction, has the goal of further exploring the

Proc. 2nd Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Skövde, Sweden, 1998

benefits noted by Krueger and developing an inexpensive implementation. The focus of the development is to provide interactivity with low-latency feedback. Although the interactions do not have to take the form of a game, this paper often refers to the motivational activity as playing within the context of a game. Real therapy still requires real work, but the goal of this research is to add fun, game-like elements to it.

An example of how Virtual Interaction might be used is the implementation of the video game, *Breakout*. (The images in Figure 3 are based on this game.) Atari introduced Breakout as an arcade game in 1976. A ball bounces between a paddle and a wall of bricks. Each time the ball hits a brick, that brick is removed. (In order to create a more playable game, the bricks have been moved from their original vertical orientation to the brick ceiling shown in the pictures.) With Virtual Interaction, the user becomes the paddle. The player steers the ball toward the remaining bricks while trying to keep the ball from hitting the bottom of the screen.

3. VIRTUAL INTERACTION IN PHYSICAL THERAPY

In most instances, the ability to understand the physical world is learned from direct experience. The ability to move, understanding relationships with one's environment and other physical concepts are learned through practice and refinement. For individuals with disabilities, the opportunities to interact and learn may be limited. Virtual reality may be able to offer individuals with disabilities the opportunity to manipulate objects and experience situations not otherwise available to them.

Krueger found that people are motivated to respond when they see their images. Their sense of curiosity led them to explore when the graphical objects in the image responded to their actions. Krueger hypothesized that the engrossing nature of the activity could be used in physical therapy to provide distraction from repetitious and tedious exercises. He saw from first hand evidence that individuals with disabilities could be motivated to push their capabilities in order to interact with a virtual environment. (Krueger, 1990, p198)

Although no formal studies have been done, anecdotal evidence agrees with Krueger's hypothesis. A Virtual Interaction system was demonstrated at RESNA 97, held in Pittsburg, Pennsylvania, and attendees were invited to try the system. A young man with Cerebral Palsy played a remarkable game of Breakout using his head to control the ball. He played the game with great intent and, based on his respiration rate, got an exceptional upper trunk workout. Others have tried the system in the laboratory and have responded favorably to its potential as a therapy tool.

Virtual reality has potential in cognitive or educational applications. (Helsel, 1992) Although the Virtual Interaction project has focus primarily on physical therapy, an example of a simple cognitive exercise is its implementation of the Milton Bradley game *Simon*. This memory game presents a sequence that must be repeated back by the player. The sequence increases in length each round. A number of possible adaptations are envisioned, such as adjusting the number of choices to the cognitive level of the user. A physical challenge might be added by programming the buttons to move a little further away each round.

As part of the Virtual Interaction project, time has been invested observing therapy sessions in a variety of settings and discussing the system with therapists. This provided ideas for potential applications and has helped to begin to establish a design methodology. In general, therapists have been enthusiastic over the possibilities of Virtual Interaction. A common theme among the discussions was that therapists are always looking for ways to keep therapy and exercise interesting. The discussions confirmed the hypothesis that motivation is an important element in physical therapy.

Virtual Interaction is a low-immersion virtual reality environment that does not burden the user with goggles or sensors. This avoids problems associated with high-immersion virtual reality systems, such as motion sickness and cumbersome equipment. Like all virtual reality, the Virtual Interaction system allows the environment to be defined in any manner. Gravity, mass and inertia can all be ignored or redefined to tailor the interaction to the abilities of a user. While the system is intended to require the users to physically participate in the environment, there is no real physical interaction. The lack of physical contact, for example, makes it an appropriate tool for providing exercise for burn patients with painfully sensitive skin. The behavior of the virtual objects can be customized for each particular activity. For instance a virtual breeze may be added in order to gently push objects in a direction and emphasize activity on that side of the user.

A diverse population with different levels of physical and cognitive abilities must be considered when creating activities for therapy. Programs should be designed so that the therapist has a user-friendly interface to adapt the activity quickly and easily. There may be other factors that require consideration when designing activities. It has been noted that some young children (and cognitively young individuals) passively listen to audio stimulation and do not resume physical activity until the sound has stopped. (Fell 1994, Contaldi 1993) What was intended to be motivational becomes their sole focus and inhibits their participatory activity. It is likely that a similar effect will be seen with video interactions. As long as there is movement on the screen, some individuals may simply sit and watch. The graphical objects may have to cease activity completely in order to solicit participation from these users.

The system can log the user's activity and potentially track their performance. For instance, Breakout logs each point of contact between the ball and the silhouette. The aggregate provides some insight into the ability of an individual to reach different areas of the playing field. When distracted and motivated by the game, individuals will perform at their maximum capability. Data-logging can be used to show growth in accomplishments over time. It might also demonstrate differences between initial expectations and actual achievements. One of the most powerful lessons that individuals can learn is that there might be a difference between their self-perceived capabilities and their actual abilities.

4. THE VIRTUAL INTERACTION SYSTEM

The basic Virtual Interaction configuration is shown in Figure 1. The user sitting in front of a blue backdrop, facing a camera and video display (TV). The camera and the TV are connected to the Virtual Interaction system's electronics, which reside in a personal computer. (Note that there are two video displays, the TV and the computer monitor.) An alternative configuration uses an overhead camera and a blue table to provide a desktop interaction field. Other modes of use might include multiple users, or focus on isolated body parts, for instance focusing on the leg for kicking exercises.

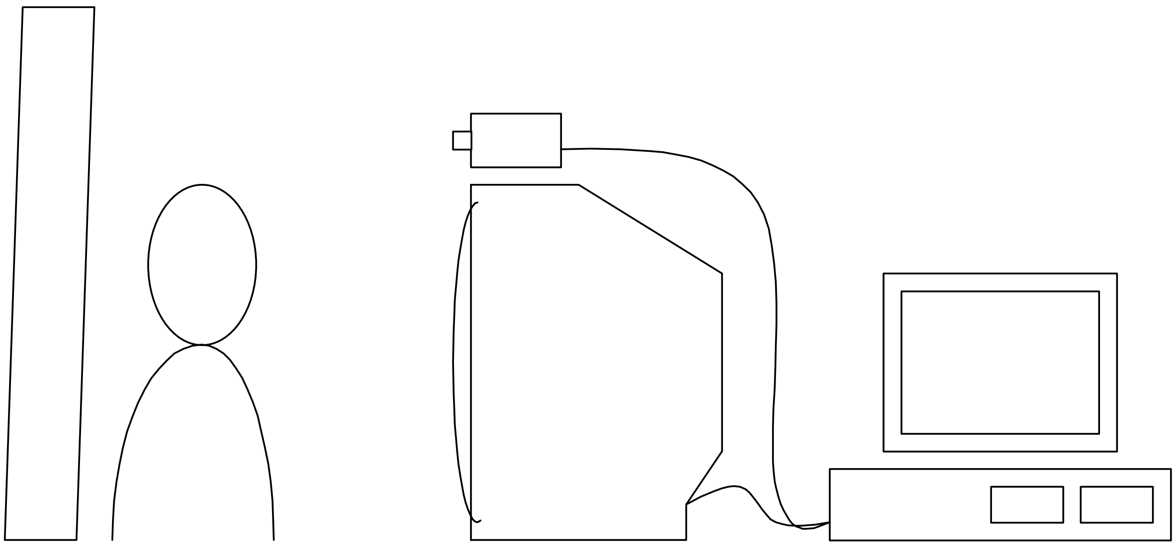

Figure 1. *The user sits in front of a blue backdrop. The camera on top of the TV sends the video image to the computer. The computer mirrors the video, adds graphical objects to it and sends it out to the TV, where the user sees him or herself interacting with the virtual objects.*

There are several considerations that must be reviewed when setting up a system. The camera must be far enough away from the user so that the camera's field of view is wide enough to provide a reasonably large playing field. Video conferencing cameras generally have wide-angle lenses with 45-degree fields of view. The full span of the arms can be seen if the user is five feet or so from the camera. This is also a good distance for users to be from the TV. They can comfortably see the entire screen but are still close enough to see detailed interactions. The backdrop needs to be as large as the camera's field of view. It should be as flat as possible and lit evenly from the front. Typical office lighting is generally good enough as long as the backdrop does not create a shadow over itself or the user, and the user doesn't create a shadow on the backdrop.

5. DISCUSSION OF SYSTEM REQUIREMENTS

The basic requirements of the system are an overlay system and foreground/background separation. The overlay system merges the live video with computer generated graphics to produce the video that the users see. The foreground/background separation produces the silhouette used for controlling the interactions. In addition, a large display is desired for output. Incorporating audio output can enhance the illusion of interaction as well as providing other game playing elements.

Figure 2 shows how the video flows through the electronics. The video from the camera is fed into the live video frame buffer and a chromakey. The chromakey output is also stored in a frame buffer, which the computer queries for its foreground/background information. The computer-generated graphics are drawn in the graphics frame buffer. These two video streams are merged and converted back to an analog video signal, which is sent to the TV display. It is possible to eliminate the chromakey and its frame buffer by adding computer access to the live video frame buffer. The computer makes the foreground/background decisions based on what pixel values are read from the live video frame buffer. This also makes alternative tracking schemes, such as flesh detection, possible.

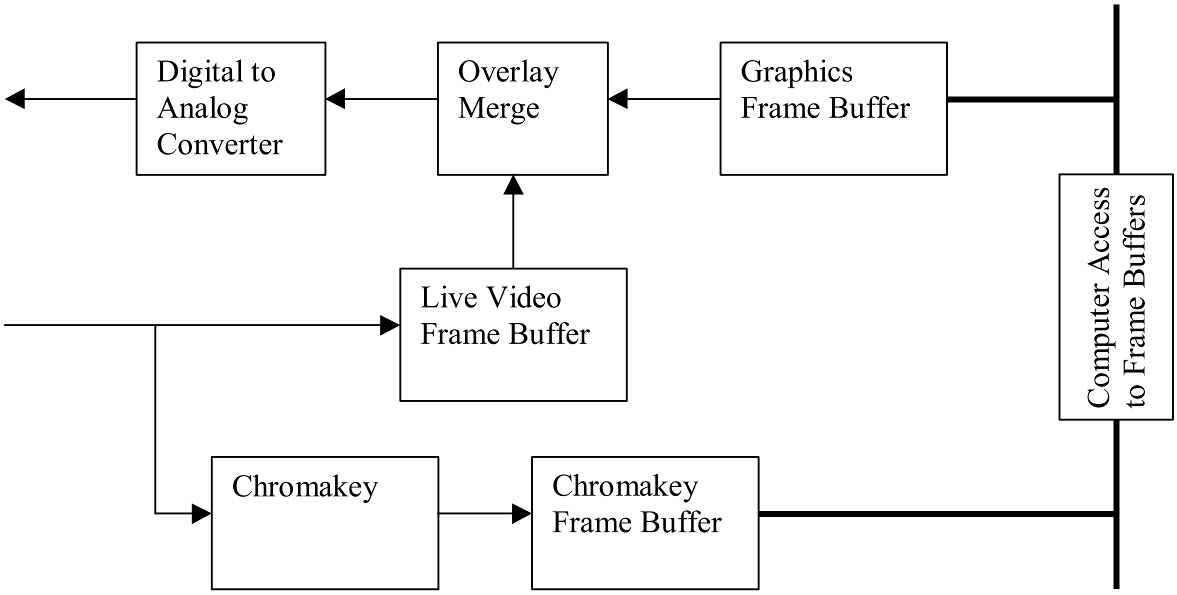

Figure 2. *The video flows from the camera into the Live Video Buffer and the Chromakey. The computer reads the Chromakey buffer for its foreground/background information. The computer generates the graphical objects in the Graphics Buffer. This is merged with the live video and sent to the TV display.*

Figure 3 shows examples of images at various points in the system. The camera produces an image of the user, which is digitized and sent to the Live Video Frame Buffer and the Chromakey. The Chromakey produces a binary foreground/background image. The computer samples this image and causes the graphical objects to react to the user's silhouette. The graphical objects are drawn in the Graphics Buffer. (The image in the graphics buffer is not shown in Figure 3.) The live video is mirrored by reversing the horizontal scan and is merged with the graphics video stream. The resulting video is shown to the user.

Figure 3. *Example images at various points in the Virtual Interaction system. The left image is typical of the camera output. The center image is the user's silhouette, as produced by the chromakey. This is the image that the system reacts to. The image on the right is the resulting image shown to the user.*

Most users are familiar with their images in a mirror and are able to use the visual feedback to position themselves. A camera and a monitor provide the view of an outside observer, which is not a model of interaction that most users are comfortable with. Horizontally flipping the video image produces the mirror-like feedback that users are accustomed to. This is such a natural feedback mechanism that most users do not recognize the difference until it has been pointed out to them.A life sized, or near life sized video image provides the user with some sense of immersion. A large screen output helps to create a one-to-one correspondence that allows the user to treat the system as a magic mirror. Large screens also help users to see detailed graphics and intricate interactions. Television sets are currently the most cost-effective means of providing this large format. It is therefore important that the output video be available in NTSC or PAL format.

Sound can enhance the illusion of interaction. The sound seems to define the moment of impact as well as indicate that some amount of energy has been expended. The system requirements are an ordinary sound card and a few wave files. Synthesized musical instruments can also be used to produce other auditory elements or provide the auditory output for a virtual musical keyboard.

6. IMPLEMENTATIONS

The project is in the early stages of its third implementation. The goal has been to develop a system that would be widely available to therapists and clinicians. If the cost can be held down, it might be economical enough to be used in private homes to continue physical therapy outside of clinical settings. Off-the-shelf technologies have been used wherever possible and all of the implementations have been based on IBM compatible PCs.

The first implementation integrated commercially available products with some custom electronics. A Creative Labs Video Blaster was used to merge live video with graphics generated under Microsoft Windows 3.1. The resulting video was seen on the computer's VGA display. A Video Blaster version FS 200 was used to vertically flip the video in real time. Turning the camera over on its back and vertically flipping the video produced the requisite mirrored video. Since accessing the frame buffer on a Video Blaster caused the video to freeze, an alternative method for sampling the color content of the image was required. A Micro Search ChromaKey+ (an analog chromakey) and a custom 1-bit deep digitizer were used for the foreground/background separation.

An AverMedia Averkey converted the VGA signal into NTSC format. The computer's VGA display mode had to be 640 by 480 pixels with the entire screen occupied by the output video. Keyboard input was used to activate pop-up dialogs to adjust the game parameters. Since the NTSC output simply echoed the computer's display, these dialogs were also seen on the user's display.

All of the electronics have been integrated onto a single, fully custom printed circuit card for the second implementation. This has been done to overcome some limitations of the first system, push the technology and add some features. This card takes NTSC video in and puts the resulting video out in NTSC format. The Live Video and Graphics buffers are double buffered for real-time, transparent access to the digitized video and smooth animation. Because the video and graphics buffers are completely separated from the computer's VGA display system, the VGA display no longer shows the output video. Therefore, the VGA display can be used to present other data, for instance to show the therapist the user's performance or allow adjustments to the game while it is being played. It might also be used to allow the therapist to have participatory input.

In addition to the basic game requirements, the custom card contains extra resources that were included to facilitate research toward more sophisticated systems. There is an FPGA-based processor connected to the capture bus and an additional graphics buffer referred to as the Background Buffer. The FPGA-based processor is intended to process the video stream on a per-pixel basis. The envisioned uses are to implement a hardware chromakey, hardware based region finding, or simple image processing algorithms such as edge detection. The Background buffer's output is multiplexed to either go to the FPGA-based processor or go to the output video multiplexer. The foreground/background separation can be accomplished without the blue backdrop if the Background buffer is filled with a reference image, which is then compared to the live video by the FPGA-based processor. Alternatively, the FPGA-based processor could be used to implement a chromakey, which could control the video multiplexer to allow a background image to replace the blue backdrop.

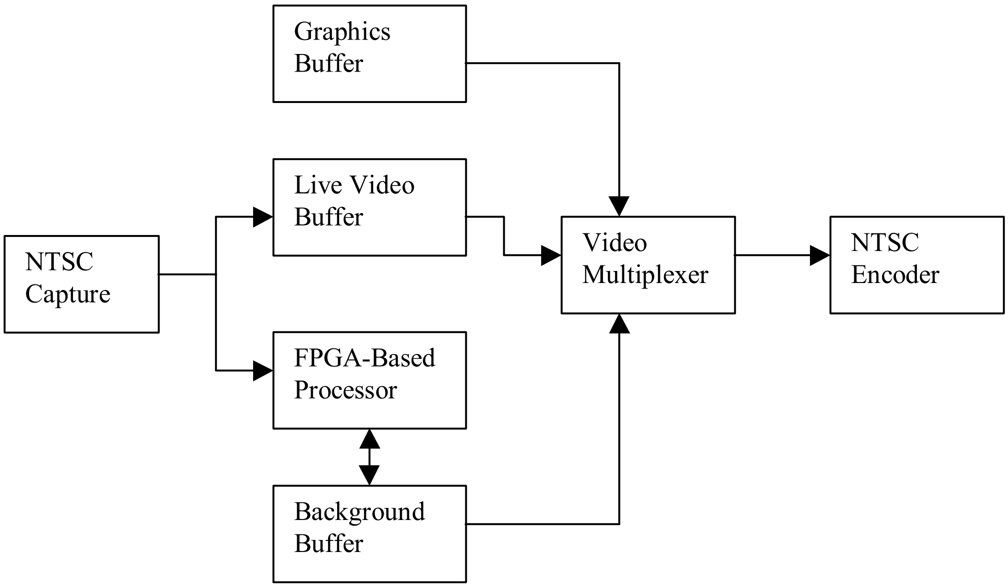

Figure 4. *The architecture of the custom Virtual Interaction card contains additional resources. An FPGA-based processor and a background frame buffer have been included. These are intended to support research into more sophisticated scene generation and more sophisticated foreground/background separation with feature tracking.*

The custom card contains four Motorola XC56166 digital signal processor chips, one for each of the three frame buffers and one for the FPGA-based processor. Each frame buffer consists of four Texas Instruments TMS55161 VRAM chips, thereby giving each frame buffer two pages of 1024 by 512 memory. The memory controllers are implemented in Xilinx XC3142 FPGAs, which allows the memory models and addressing modes to be changed to suit an application. The NTSC decode is performed by a Brooktree Bt819 and the NTSC encoder is a Bt856.

The card has a 16-bit ISA bus interface and uses a single 16-address slot in the IO address space. The PC does not have direct access to the frame buffer memory but rather issues commands to the DSP to read from or write to the memory. Although this provides mediocre performance when operations are handled on a pixel by pixel basis, the DSPs are a significant resource and can be programmed for higher level functions. For instance, if run-length encoding is used for the graphical objects, the fill rate improves from about 60 thousand pixels per second to roughly 2 million pixels per second.

The third implementation uses the newly available VigraVision card by VisiCom (San Diego, California, USA). This card provides video capture, NTSC/PAL output, fast access to the video buffer, and can be set to mirror the video. This card, introduced in the spring of 1998, is the first commercially available product that meets all of the system requirements at reasonable cost. Although this system uses an overlay capture system similar to the Video Blaster, the display adapter does not have to use the 640x480 mode. Therefore, over half of the area on the computer monitor is still available for control functions and performance monitoring.

In addition to the video capture and overlay requirements, any implementation must meet certain rendering requirements, which are application dependent. Basically a system must be able to create smooth animation, which requires both processing power and double buffering. None of the demonstration programs are particularly demanding, so the rendering power has not been an issue. The VideoBlaster system uses Windows 3.1, which does not support double buffering. To prevent the erase and redraw operations from creating visible artifacts, both operations are combined into a single BitBLT operation. This operation copies a bitmap that has been expanded, so that when it is copied it restores the area previously occupied by the graphical object to the background color of the overlay. The custom implementation has a dedicated frame buffer for the graphical overlays. The buffer is driven by a DSP, which performs many of the pixel level operations. The VigraVision system uses Microsoft Windows 95 with DirectX. It incorporates the features of a modern video card with double buffering and a graphics accelerator, as well as the requisite live video overlay.

All three implementations use chromakeying with blue backgrounds to determine the user's silhouette. Any chromakey is subject to the fundamental bandwidth limitations of the incoming NTSC signal. Color changes are smeared over four to eight pixels, which makes the vertical edges of the silhouette ragged. Interactions must filter out this noise and recover gracefully from the inevitable errors. Controlled lighting can produce superior definition of the silhouette but the goal of this project has always been to provide a system that can be placed in a clinical setting with as little physical disruption to the environment as possible. Instead of controlling the lighting, the focus has been on working around the problems created by a less than perfect chromakey separation.

7. INTERACTIONS

The basic implementation of *Breakout* uses a bounce interaction, which tests for the silhouette in the path of the ball (or any other object using the bounce interaction). The program tests at several points around the periphery of the ball in order to detect a glancing blow as well as frontal contact. For a medium speed game, the ball moves about 6 pixels during a 33ms frame period. (That equates to moving across the screen in about 3-4 seconds.) Therefore the bounce interaction tests about 30 pixels per frame period. When a contact is made, further checks are performed to eliminate spurious noise and then the normal of the silhouette is calculated. The dot product of the normal and the motion vector produces the magnitude of the impact. Adding twice that amount of motion in the direction of the normal gives a mirror-like deflection.

Interestingly, the addition of a collision sound as the ball bounces off of the user's silhouette enhances the illusion that the purely graphical ball actually made physical contact with the user. Sound is associated with the expenditure of energy and hearing a sound seems to provide the user with the sense that a physical, energetic collision occurred.

Simple switches can be used to activate anything for which the application calls. The Simon game mentioned above uses simple switches. An application that demonstrates the use of switches as well as audio interactions is the virtual musical keyboard. Each switch is tied to a different note, which is played though the MIDI interface by a Sound Blaster FM Synthesizer. In order to debounce a switch, the program tests a small region for activation and tests a larger, overlapping region once activated. The silhouette must therefore be well outside of the activation region before the switch is reset and a new switch contact can be initiated.

An application that uses a drag and drop interaction is an anagram unscrambling game. The computer presents a set of letters and the user drags the letters into their proper sequence. The drag and drop interaction is activated with touch and the graphical object then moves with the user's silhouette. The interaction continues until the program tells it to drop, either because the user held still, or the object reached a goal. The drag expects to track an appendage which is longer than it is wide, for instance, an arm, or on a smaller scale, a finger. The tracking continually calculates the major axis of the appendage, using the normals along the side to estimate the direction of the axis. The most extreme point along the major axis is used as the tip and is the location of the drag point.

An interaction mode that allows users to swat objects is being developed. The swat interface must not only determine that the ball has been hit, but where the silhouette was before it hit the ball. The program being developed uses an anticipatory proximity measure, which measures how close, and in what direction the silhouette is to where the ball is going to be. When contact is made, changes in the proximity measurement determine the magnitude and direction of the deflection. Rapid hand motions by the user may be on the order of 30 pixels in a single frame period (assuming a typical upper body and arm span camera view). This is about the same as the diameter of the ball or the thickness of the hand viewed edgewise. Because the silhouette can come from any direction, looking ahead by two frame periods requires the program to examine regions that are about 150 pixels on a side. A sparse search is used to find the silhouette and then the estimate of the closest point is progressively refined. Because of the high-speed access to the video buffer, the VigraVision system is able to store the raw video and delay the search until contact has been made. The custom card system must perform the search during each frame period.

8. ERRORS AND OTHER CONSIDERATIONS

The system is not perfect and it makes mistakes. Given the particular application and diversity of abilities expected among the users, graceful error recovery is important. For instance, in Breakout, the ball can become imbedded in the silhouette. Once inside, it is surrounded by the silhouette and bounces continuously. The net effect is a ball that does little more than wiggle until the user moves enough to free it. Requiring the users to make significant movements is not a viable option for error recovery. Instead heuristics are employed. For this error, the heuristic prevents the ball from bouncing continuously by insuring that the ball moves some minimum specified distance without striking the silhouette. If the silhouette continues to be found, the ball is considered trapped and it simply continues in its present direction until it emerges out of the silhouette. It then reverts to its normal bounce mode.

The swat interface is being developed because experience with the Breakout demo has shown that users want to swat and push the ball, which is how they would interact with a real one. The bounce interface tests for contact along the leading edge of the ball under the assumption that the motion of the ball causes all contacts. Therefore, it assumes that the contact point is along the edge of the silhouette. However, when the silhouette has moved since the last test, that contact point may be inside of the silhouette and there is no edge on which to base the normal. Another error that occurs is push through. The user hits the bottom of the ball intending to push it up. The program correctly sees the silhouette and bounces up. But the hand continues to move and pushes through the ball. The ball then sees the bottom of the hand in front of it and bounces downward. These experience points out that a user's expectations and perception of the interface must be considered when designing the interactions.

There are other sources of errors, not the least of which is the misclassification of the foreground and background. Interactions must be designed to work with less than perfect silhouettes. The fundamental NTSC bandwidth issue and motion blur both help to create ragged edges. Controlled lighting and a camera with a high-speed shutter can provide superior separation, but the resulting system would result in a significantly higher impact on the clinical environment.

9. FUTURE RESEARCH

Future system development can explore more sophisticated interactions. It is not possible to track a feature inside of the silhouette. For instance, the hand cannot be used to drag an object across the body. Once the silhouette of the hand is inside of the silhouette of the body, it disappears as a trackable feature. An application may be able to compensate for this limitation with heuristics. The ALIVE System at MIT estimates the locations of the hands based on the body outline. If the arms are away from the body, they are easily tracked. If the arms do not protrude from the basic silhouette of the body, then the contour of the body is used to estimate the orientation of the arms. A change in width of the body at the hips indicates that the hand is at the user's side. If there is a change in the mid-torso region, the arm is bent at the elbow. Ideally, the system might provide more sophisticated feature tracking perhaps by using flesh detection.

The system will be studied as a tool to aid physical therapy. One of the reasons behind the migration to the VigraVision card is the need for stable, commercially available hardware for use in clinical settings. The technology demonstration programs need to be modified and extended in order to create activities that can be used with a variety of individuals.

10. CONCLUSION

Virtual Interaction can provide interactions that are engaging and believable. Experience with the demonstration applications has shown that the interactions are somehow tangible but the environment is a harmless cartoon world. An indication of the tangibility of the objects can be seen in the amusement users get when balls bounce off of their heads. They eagerly explore what effect their actions elicit. The mirrored self-image provides easily understood feedback.

Virtual Interaction should provide a unique, customizable, motivational tool for therapy. The applications should include good exercise and learning environments for individuals with disabilities. By making these activities more fun, the system has the potential to motivate a person to do more in therapy, thereby improving their well being.

Virtual Interaction can be implemented with reasonably inexpensive, off-the-shelf technology. Current modes of interaction should be extended and further modes, such as the simple gesture recognition demonstrated at MIT, developed. Improved technologies, such as flesh detection and more sophisticated feature tracking, should allow more capabilities and open new modes of interactivity.

Acknowledgements: This research has been funded by the U.S. Department of Education, Grant #H129E20006 from the Rehabilitation Services Administration and Grant #H133P30003 from the National Institute on Disability and Rehabilitation Research, The National Science Foundation, Grant #HRD-9450019, and the Nemours Research Programs.

11. REFERENCES

E. Contaldi, P. Taney, (1993) Introduction to Computer Based Music for the Consumer, *Presentation at DATI Annual Conference*, Oct 4-5, 1993, Newark, Delaware

H. J. Fell, H. Delta, R. Peterson, L. J. Ferrier, Z. Mooraj, M. Valleau, (1994) Using the Baby-Babble-Blanket for Infants with Motor Problems: An Empirical Study, *ASSETS '94*, p77-84, Marina del Rey, California (Author's note: The statement concerning the cessation of activity was part of the presentation but was not noted in the paper.)

S. Helsel (1992) Virtual Reality and Education, *Educational Technology*, May 1992, p 38-42

A. W. Joyce, III, A. Phalangas, (1995) Virtual Interaction; An Interface for Individuals with Disabilities, *Proceedings of RESNA '95*, Vancouver, Canada, p 425-427

Myron Krueger, (1990) *Artificial Reality II*, Addison-Wesley Publishing Co., Reading, Massachusetts

P. Maes, T. Darrell, B. Blumberg, A. Pentland, (1995) The ALIVE System: Wireless, Full-body Interaction with Autonomous Agents, *M.I.T. Media Laboratory Perceptual Computing Technical Report No. 257*

F. MacDougall, (1996) Video Image Based Control System, *U.S. Patent Number 5534917*, July 1996