Kevin Ponto*[†], Falko Kuester[†], Robert Nideffer[¥], Simon Penny*

Virtual Bounds: A Teleoperated Mixed Reality

*Arts Computation Engineering

[†]Calit2 Center of GRAVITY

[¥]Laboratory for Game Culture and Technology

The University of California, Irvine

Irvine, CA 92697

kponto@uci.edu, Phone: 949-824-2109, FAX: 949-824-7780

ABSTRACT

This paper introduces a mixed reality workspace that allows users to combine physical and computer-generated artifacts, and to control and simulate them within one fused world. All interactions are captured, monitored, modeled and represented with pseudo-real world physics. The objective of the presented research is to create a novel system in which the virtual and physical world would have a symbiotic relationship. In this type of system, virtual objects can impose forces on the physical world and physical world objects can impose forces on the virtual world. *Virtual Bounds* is an exploratory system allowing a physical probe to navigate a virtual world while observing constraints, forces, and interactions from both worlds. This scenario provides the user with the ability to create a virtual environment and to learn to operate real-life probes through its virtual terrain.

KEYWORDS

Augmented Reality, Mixed Reality, Tangible Bits, Analog Gaming, Remote Control

ACKNOWLEDGEMENTS

INTRODUCTION

Over the past 20 years, the lines between the virtual and physical worlds have become increasingly blurred. Computer graphics techniques can now generate images that are nearly photo-realistic, and artificial intelligence has taken great strides in conquering the Turing test. Yet, physical and virtual realities stay juxtaposed when it comes to interaction paradigms. Human-computer-interaction (HCI) is generally a one-way communication in which the physical world tells the virtual world what to do. The virtual world, on the other hand, has no control

over the physical world, and may, at best, provide limited force-feedback. Haptics and forcefeedback devices have been developed as a means to address this. Consumer-based versions of these devices usually provide the user with a tactile feedback, most commonly in the form of a rumbling controller. While these systems are evolving, most of the existing feedback devices are intended to simply augment the physical world systems, but are not powerful enough to modify them.

Figure 1
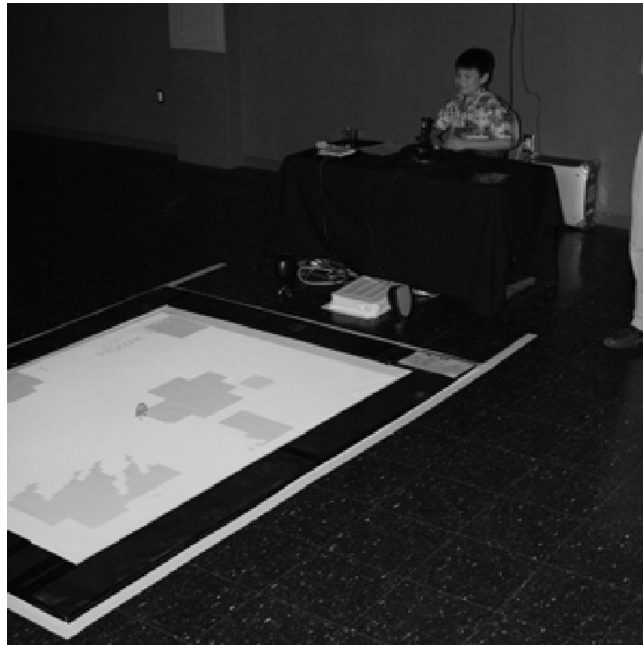
One field in which haptic devices have proven to be effective is in teleoperations. Teleoperation refers to actions that are performed remotely with the operator removed from the actions either due to differences in location or scale. Teleoperated vehicles (probes) are often used in situations in which safety, working environment, accuracy or practicality mandate the use of machines.

For example, teleoperated planes can be built smaller and faster than planes that require a pilot and can operate at higher g-force levels. Teleoperated robots can be sent into dangerous situations to disarm explosives, perform deep sea rescues, and to do reconnaissance when a human life would be subjected to extreme danger. In the medical field, teleoperated robots are used to perform surgeries remotely when a physician is unable to be present. Similarly, miniature teleoperated robots may used to perform surgeries that would be too invasive if

performed with traditional surgery techniques. Another field in which haptics has proven useful is in mixed reality. Mixed reality systems are composed of elements that are partially virtual and are partially physical. The advantage of a mixed reality system is that it can combine the tangible aspects of physical objects with the freedom and control available in the virtual world. [1]

This paper introduces a mixed reality workspace that allows users to combine physical and computer-generated artifacts, and to control and simulate them in one fused world. All interactions are captured, monitored, modeled and represented with quasi-real world physics (as shown in Figure 1). The objective of the presented research is to create a multimodal interface in which the virtual world and physical world have a symbiotic relationship. In this type of system, virtual and physical objects can impose forces, constraints and actions on each other. This type of system is used for a variety of applications requiring the virtual and physical world to exist on a more balanced playing field. In this way, *Virtual Bounds* is an enactive interface that makes use of non-symbolic forms of knowledge (ie actions) to create systems intuitively accessible to average users. A proof of concept system for a teleoperated system is presented that is amenable for teleoperations training.

Related Work

Mixed Reality research draws heavily from fields such as teleoperation and augmented reality with particular focus on machine vision and digital image processing, tangible interfaces and design, physics modeling and simulation to name a few.

In the early 1990's, Taylor et al. [2] developed a virtual reality interface for a scanning tunneling microscope allowing users to "touch and feel" a surface at a microscopic level. In addition, users were allowed to make controlled modifications, such as sculpting, at an atomic scale. In another example, researchers use haptics to emulate the feel of piloting a teleoperated vehicle directly. Brady et al. [3] developed a system to relay forces and torques experienced by a teleoperated vehicle back on the users controller. This allowed those piloting the vehicle to experience the same sensations that they would inside the vehicle .

Remote controlled devices are used as an interface for a wide range of applications. Kaplan et al. [4] present internet accessible telepresence systems for a remote controlled vehicle. A camera was mounted on the front of the RC car, giving the user a first-person perspective of the surroundings. Lawson et al. [5] combined the small RC car perspective with a virtual environment, creating an augmented teleoperated system. These systems combine the advantages of viewpoint and remote control, while superimposing additional information, such as dimensions, to enhance the users' situation awareness.

Image processing and object recognition are other vital components of teleoperations systems. Krueger [6] was an early pioneer in the field of virtual reality and human computer interaction systems in the 1960's and 1970's. His work focused on systems in which user actions could be interpreted directly without the user communicating through an physical interface. His most recognized work, *Videoplace*, used innovated techniques to gather user information. The user was placed in front of a screen with backlighting, allowing a computer vision system to acquire user actions. From this interface, Krueger defined over 50 different types of interaction paradigms for the user to explore.

Since the 1970's, machine vision has progressed steadily towards realtime processing as the result of increased computer performance. Crowley, de Verdi et al. [7] tracked individual colors in order to follow the movement of objects on a deskspace.

Objects tracked in the real world can also be used as an interface to the virtual world. Ishii and Ulmer [8] of the MIT Media Laboratory coined the term "tangible bits", to describe these types of systems. As stated by Ishii and Ulmer, "Tangible Bits is an attempt to bridge the gap between cyberspace and the physical environment by making digital information (bits) tangible. We are developing ways to make bits accessible through the physical environment." [8]

Using this concept, Chun et al. [9] created interface for athletic-tangible computer supported cooperative play, termed *PingPongPlus* . They utilized an overhead camera to track a ping-pong ball during a normal ping-pong match. In this system, an overhead projector projects images onto the table to augment the real world game play. Several different modes were created, in which different goals and images were displayed to enable different types of gameplay. Chun et

al. found that their system did not only augment reality, but translated reality by creating new game interaction paradigms for participants to interact with.

Other researchers, such as Mark Billinghurst, have used machine vision to create augmented reality interfaces [10][11][12]. By determining the placement and orientation of real world objects (feducials), virtual objects can be overlaid corresponding to the users perspective. Users of these systems are commonly required to wear a clear head-mounted-display in order to achieve this hybrid viewpoint. The main advantage of this augmented reality technique is that a 3D view of both the virtual and physical world is available.

Other mixed reality projects have used remote control vehicles for the purposes of gaming [13][14]. These systems have a control structure similar to that of *Virtual Bounds*, consisting of overhead projectors, tracking devices, and user control units. However, these systems aspired to augment shooter-style video games, while the goal of *Virtual Bounds* was to create an enactive teleoperation environment.

Technical Approach

Figure 2
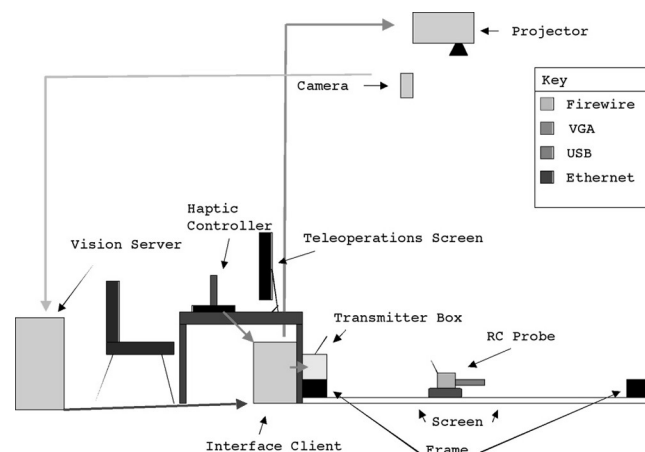
The mixed reality system described in this paper consists of five main components (as shown in Figures 2 and 3). The first component is the physics probe controller, a haptic joystick that allows the user to interact with the system. This joystick sends messages through an interface computer, which processes user input, converts it into control sequences for the remotely operated probe and then

6

forwards these sequences to the transmitter box. The transmitter box subsequently sends the appropriate radio signals to control a small RC vehicle, which acts, in turn, as the probe into the virtual world.

Figure 3
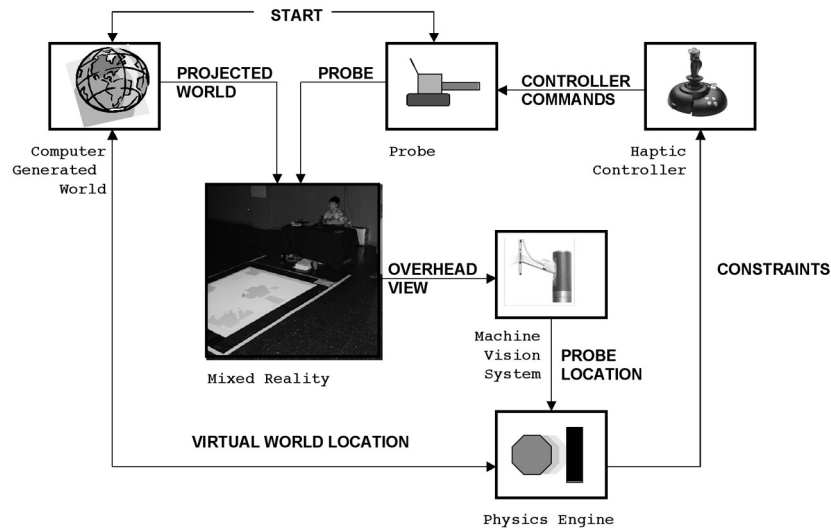
The machine vision system, consists of a camera connected to a computer node tasked with determining position and orientation of the probe. This computer acts a tracking server and continuously sends the reference coordinates target object as an UDP stream.

The physics engine acts as a UDP client, and receives these packets. The physics engine uses this data to translate the physical world position into a corresponding virtual world location used to determine if any virtual world forces should be enacted back on the probe. Simultaneously, the physics engine provides the forces that the probe exerts inside the virtual world.

Once this process is finished, the virtual computer-generated world can be implemented to reflect the current state of the mixed reality world. The fourth component of the system allows all virtual objects to be constructed, animated and displayed. Virtual events are also sonified and outputted through series of speakers.

The display system is the fifth component in the system and consists of an overhead projector projecting onto a thin custom build, floor mounted screen. A frame was built around the screen for aesthetic, as well as practical reasons. The

frame acted as a failsafe to the probe moving beyond its restricted physical world bounds and provided stability to the screen material.

To begin interaction with this system, the user is required to first charge the probe. Once charged, the probe can be placed on top of the computer-generated environment. At this point, the vision system recognizes the presence of the probe and begins to track its movement. The user is now able interact with the system through a haptic controller. The physical world movement of the probe is constrained by virtual objects, their corresponding physical properties and the physics engine.

Physical Probe Controller

The physical control of the system is divided into three separate modules, the joystick, the interface client and the transmitter box. The user interface into the system is through a haptic joystick, which communicates with the interface client. A Microsoft Sidewinder 2 joystick is used for the proof of concept system. The joystick input is monitored by the interface client for position changes (up, down, left, right). Intuitively, pushing the joystick up, moves the probe forward. Pulling it back moves the probe in reverse. Moving the joystick to the left, turns the probe left, while moving the joystick right turns the probe right. For the presented case study, the trigger button causes the probe to emit virtual projectiles and the throttle is used to control whether the probe fires projectiles that will build or destroy virtual terrain.

The monitoring interface client sends commands to the transmitter box based on the state of the haptic controller and the virtual world. The transmitter box is built using two external inputs for power and data. A 9V DC input is necessary to run the internal electronics, and a USB input is needed to communicate with the interface client. The box internally uses a Keyspan USB to Serial Converter, to send serial signals to a PIC microprocessor. The serial signal runs at 19,200 baud with 8 data bits and 1 stop bit. The PIC microcontroller then interfaces with a modified 27 MHz RF transmitter to send the appropriate commands to the RC Probe. The transmitter box subsequently sends the appropriate RF signal to the probe. The probe receives these transmitter signals and moves accordingly.

The RC Probe is made from an ECOMAN Remote Control Mini Tank. This product operates on the 27 MHz frequency and has a 96:1 gear ratio allowing it to move slowly, with enhanced torque. These features are especially useful since many small RC products move at very high rates of speed. The ECOMAN Remote Control Mini Tank utilizes two treads. Each tread can either move forward or backward allowing the probe to either move forward, backward, or spin. The tank itself was modified with an overlay mask, colored yellow on the body and orange on the turret. This was done to remove the militaristic association with the RC probe and to aid in the color tracking routine. Florescent variations of the colors yellow and orange were chosen to be minimally variant under varying lighting conditions.

Machine Vision System

The machine vision system is tasked with obtaining the physical world position and orientation of the RC probe. In order to achieve this, a camera (Apple iSight ®) is mounted over the projection area and image data passed to the machine vision system for processing. This small compact Firewire® camera captures 24-bit images at 640x480 pixel resolution at 30 frames per second. The camera was mounted 10 feet off the surface and covers an area of 6 feet by 4.5 feet area, resulting in an image resolution of 4.6x4.1 mm per pixel.

Figure 4
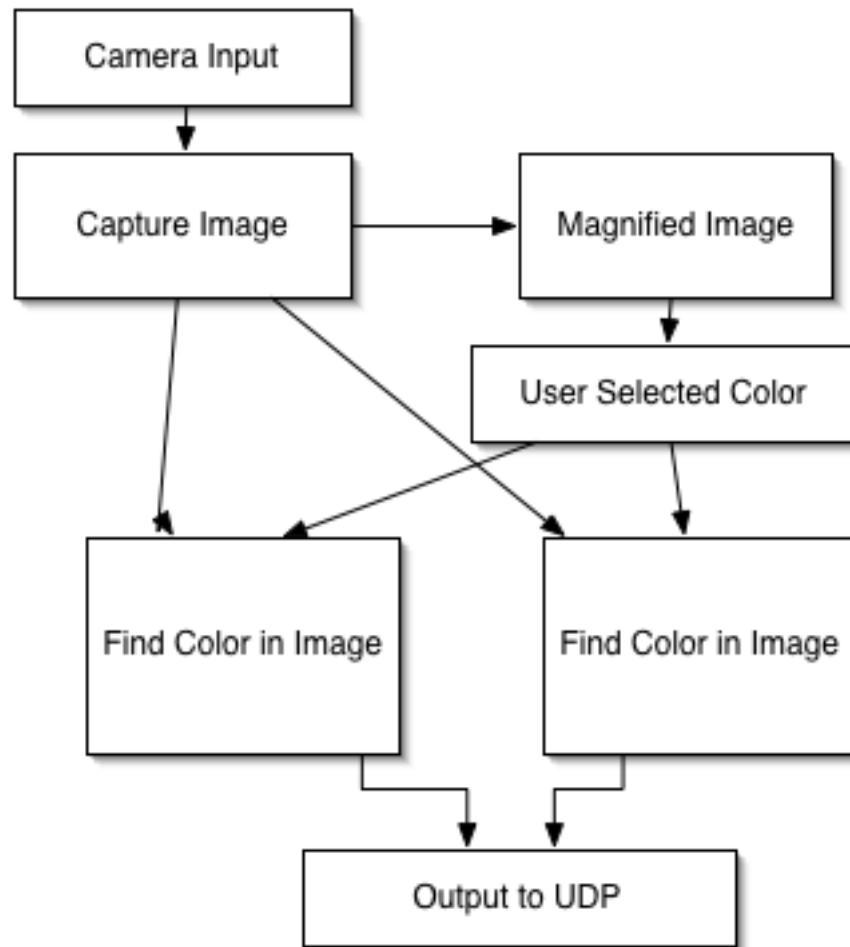
The data is transmitted through a Firewire® cable to the vision server. The vision server selected is a Macintosh Dual 1.25 GHz G4 and the image processing pipeline developed uses cv.jit (cvjit) in combination with MAX/MSP Jitter (cycling74). This configuration was selected for its ease of use and rapid prototyping support. Intel's OpenCV library is currently being used for more complicated or time critical tracking procedures.

The system reliably captures the probe using its color markings and provides the probe position and orientation within the physical world coordinates. Figure 4 shows a system diagram. First camera data is acquired, and presented to the operator. The operator chooses a magnified image around the probe to analyze further. From this, the operator chooses the hue and saturation values that will correspond to the front and rear of the probe. The image-processing pipeline analyzes the camera data to determine the centroids of the color values. These

values are then output via the UDP network protocol to the physics engine and the virtual computer-generated world.

If the probe is not in the camera space, a coordinate value of −1 is passed to the physics engine to alert the system that the probe is no longer on-screen. The machine vision system also thresholds the necessary amount of change of the probe seen by the camera before passing the values on to the physics engine. This prevents hysteresis of the probe's position, which is extremely noticeable in the first person perspective.

Physics Engine

The Physics Engine is given the task to enact virtual forces back on the RC probe. To determine what forces should be put upon the RC probe, its physical position needs to be correlated with its virtual world position. To do this, the physics engine collects the position data of the RC probes front and rear from the UDP stream generated by the machine vision system. Since these coordinates are in camera space, transformation into the virtual world reference coordinate system is necessary.

A simple calibration step is provided, allowing the user to identify a set of reference points (the corners of the physical screen space) to correlate the coordinates to the virtual space. The same calibration points also establish the reference frame for the machine vision system. The calibration step is only required during the initial system setup.

Once the physics engine has knowledge of the RC probe boundaries, it can determine the location of the probe from a fixed point. This is accomplished using inverse bilinear interpolation. The advantage of this technique is that the camera can be oriented in any direction, and does not need to be aligned with the projection system.

Once the referenced virtual world location of the probe is determined, a simple lookup can be performed to check if the RC Probe has collided with an obstacle. If so, a virtual force can be applied back on the probe. The physics engine is also required to determine if the actions of the RC probe will enact any forces onto the virtual world. Once all of these forces are reconciled, the final virtual world is created.

Virtual World

The virtual world was designed to provide an alternate reality for the RC probe to traverse and created in C++. OpenGL was used to represent the 2D and 3D virtual world. In addition, the OpenAL framework was used for audio output. Open AL provides a simple method for importing and playing sound files with sound characteristics based on virtual world locations. This is useful for generating sound effects that provide audio feedback for virtual/physical world interactions.

A mixed reality study in the form of a computer game was developed to allow the exploration and analysis of the muti-modal interaction techniques and interfaces. The game allows users to build obstacles for the RC probe as well as to destroy them. Users are able to visualize the symbiosis between the virtual and physical world and to test the capabilities of the system.

Virtual characters in the system are confined to the same boundaries as the RC probe. Some virtual characters, specifically, the assets, move according to very simple flocking behavior with their goal to reach the collection box. The collection box bounces around the virtual world like a ping-pong ball, reflecting off virtual obstacles as well as the RC probe. The game itself is described in more detail below.

Once the virtual world is constructed, it can be visually represented in different forms. The most powerful representation is the projection onto the mixed reality surface. However, since the virtual model is readily available, a remote operator can select between a birds-eye-view of the scene, and a first person view of the probe, or a combination of the two.

Visualization System

The visualization system is responsible for creating representations of the virtual world in the physical world. The system is composed of three components, the projector, the frame, and the screen. The projector is mounted 15 feet above of the floor and generates an image that is 4 feet by 3 feet. A Dell 2100MP projector is used to produce an 800x600 image at 110 MHz refresh rate.

The screen is placed on the floor below the projector. The screen is made out of a specialized 9 foot by 9 foot sheet of white, reflective plastic. The sheet

provides excellent reflective properties at a cost of approximately two dollars per square foot.

A frame was built around the projection area out of painted foam-core. The frame was built to designate where the projection area ended as well as to provide a failsafe boundary for the RC probe. Speakers were also placed around the user to provide specialized sound.

Mixed Reality Case Study

The pervasive nature of computer games and interface technology is well suited as a mixed reality test bed. Exposure (previous experience) of users allows for a natural and intuitive transition of skills and their application in teleoperations. In this case, a simple game scenario was selected, in which users are asked to guide assets (small blue key-like shapes) into the collection box (bouncing square). To do this, the RC probe can shoot virtual projectiles that either build or destroy virtual barriers. Effectively, the RC probe acts like a shepherd herding sheep. The game is very simple, giving points for the number of assets collected in a given period of time. High scores are collected and listed on the upper corner. Figure 5 shows the system in operation.

Figure 5
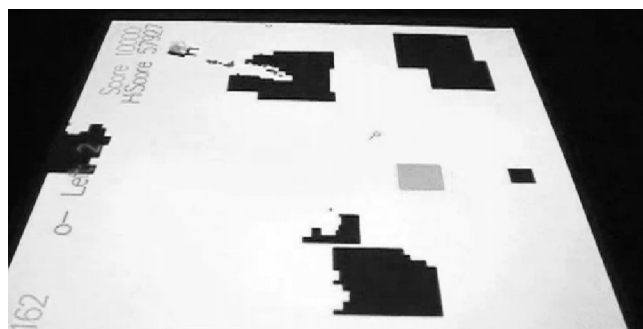
The barriers that the RC probe creates/destroys, constrain the movement of the virtual characters and at the same time constrain the movement of the physical RC probe. This creates a symbiosis between the virtual and the physical world, as the virtual world creates virtual boundaries on the physical world. For many users, making and destroying these boundaries proved to be as amusing as the game itself.

Results

The mixed reality system achieved an effective fusion between the virtual and physical worlds. The virtual world forces seem to be intuitive and "make sense" in the minds of the tested users.

As for individual components, the vision system can to track the RC probe at approximately 20 frames per second. System delay between the acquisition of position from the vision server, to the response on the client server is negligible. At this acquisition speed, the system accurately monitors the RC probe while moving forward and backward, but latency is apparent when virtual projectiles are emitted and the user spins the RC probe for extended periods of time. While this delay is noticeable, it was not considered to be disorienting for those testing the system.

Figure 6

The response time between movement of the haptic joystick and movement of the probe is less than 20 ms. Once instructed as to how the controls worked, users had few complaints about the accuracy of their intended movements and the resulting probe response. Overall the system preformed well during extended testing. The most significant problems during the a one week exhibition were RC probe reliability and battery life. Higher-grade robotic probes can eliminate these problems in the future. Figure 6 shows the removal of a probe from the operating surface during the exhibition.

User Performance Appraisal

Qualitative and quantitative measures were employed to assess the performance of this mixed reality system. Qualitative user assessments commented positively on the intuitive nature of the interface, believability of the mixed reality scenario, and the enjoyability of the operation. Negative comments focused on the need for fine-tuning the interface with respect to subtle movements and joystick responsivity. Quantitative performance measures assessed the response times when operating the system from the real-world perspective and solely from the teleoperations perspective. Five subjects participated in nine measures (trials) of response time for each mode. Data were analyzed using a repeated-measures ANOVA (Statview 5.0.1). All subjects were able to successfully manipulate the probe for all modes and in all trials. Response times did not vary significantly by trial ($p = 0.14$, i.e., no apparent learning effect) but did vary by mode ($p = 0.02$). The teleoperation mode required approximately twice as long to manipulate the probe in the specified sequence of tasks (3.47 ± 2.65 versus 5.55 ± 3.84 seconds, difference $= 2.08$ seconds). Although statistically significantly longer, this manipulation time was not unreasonable for the required task.

Conclusion

This paper presents a new mixed reality approach that combines research in computer graphics, visualization, and virtual reality with machine vision, haptics and teleoperations. The resulting system provides a multimodal interface allowing for manipulation of teleoperated probes in virtual environments.

Virtual Bounds provides a technique to train users of teleoperated systems. By allowing users to create virtual environments that produce physical world feedback, users can quickly explore and operate in a mixed reality space. This allows for greater flexibility compared to creating a full real-life model of the simulation environment.

This system also has potential utility in the field of teleoperated surgery. Robotic arms could be tracked in the same manner as RC probes. This would allow surgeons to train on virtual patients. Currently, teleoperated surgical systems train on physical models, whereas, the current mixed reality system would allow surgeons to easily swap one virtual patient with another, recreating

difficult surgical situations and providing easy analysis of the surgeon's teleoperation performance.

REFERENCES

[1] Ohta Y. and Tamura H (1999) Mixed Reality Merging Physical and Virtual Worlds. Ohm-sha & Springer-Verlag

[2] Taylor, R., Robinett, W., Chi, V., Brooks, F., Wright, W., Williams, R., Snyder, E., "The Nanomanipulator: A Virtual-Reality Interface for a Scanning Tunneling Microscope," In Proceedings of SIGGRAPH'93, pp. 127-134.

[3] Brady, A., MacDonald, B., Oakley, I., Hughes, S. and O'Modhrain, S. (2002) "RELAY: A futuristic interface for remote driving," In Proceedings of EuroHaptics 2002, Edinburgh, UK, pp. 8-10

[4] Kaplan A, Keshav S, Schryer N & Venutolo J (1997) An Internet accessible telepresence. Multimedia Systems 5(1): pp. 140-144.

[5] Lawson S, Pretlove J, Wheeler A (2002) "Augmented Reality as a Tool to Aid the Telerobotic Exploration and Characterization of Remote Environments", Presence: Teleoperators and Virtual Environments, Vol 11(4), pp. 352-367.

[6] Krueger M (1983) Artificial Reality. Addison-Wesley

[7] Crowley J, Hall D, de Verdi V (2000) Object recognition using coloured receptive fields, Proceedings of the Sixth European Conference on Computer Vision, Lecture Notes in Computer Science, vol. 1842, Springer, Berlin, 2000, pp. 164–177.

[8] Hiroshi I, Ulmer B (1997). "Tangible bits: Towards seamless interfaces between people, bits and atoms." In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '97, Atlanta, GA, Mar. 22–27), S. Pemberton, Ed. ACM Press, New York, NY, pp. 234–241.

[9] Chun B, Ishii H, Orbanes J, Pardiso J, Wisneski C (1999) "PingPongPlus: Design of an Athletic-Tangible Interface for Computer-Supported Cooperative Play," In Proceedings of Computer Human Interaction (CHI'99), pp. 394–401.

[10] Billinghurst M, Kato H, Poupyrev I, Imamoto K, Tachibana K (2000) "Virtual Object Manipulation on a Table-Top AR Environment." In Proceedings of 1rst International Symposium on Augmented Reality (ISAR'00), Munich, Oct. 2000, pp. 111-119

[11] Billinghurst M (2002) Augmented reality in education, new horizons for learning. Internet as of March 22, 2006, www.newhorzons.org/strategies/technology/billinghurst.htm

[12] Billinghurst M, Kato H, Poupyrev I, Imamoto K, Tachibana K (2001) "The MagicBook: A Transitional AR Interface." Computers and Graphics, November 2001, pp. 745-753.

[13] Sugimoto M, Kojima M, Nakamura A, Kagotani G, Nii H, Inami M (2005) Augmented Coliseum: Display-Based Computing for Augmented Reality Inspiration Computing Robot In Proceedings of Horizontal Interactive Human-Computer Systems. pp. 3 – 8

[14] Metaxas G, Metin B, Schneider J, Shapiro G, Zhou W, Markopoulos P (2005) SCORPIODROME: An Exploration in Mixed Reality Social Gaming for Children. In Proceedings of ACM conference on Advances in Computer Entertainment.