INSIDE_OUT: a case study for enhancing public participation in mixed reality events

Marientina GOTSIS
Electronic Visualization Laboratory
University of Illinois at Chicago
gotsis@evl.uic.edu

Geoffrey BAUM
Electronic Visualization Laboratory
University of Illinois at Chicago
xiccarph@evl.uic.edu

Kang SUN
Electronic Visualization Laboratory
University of Illinois at Chicago
kbs@evl.uic.edu

Robynne H. GRAVENHORST
School of Kinesiology
University of Illinois at Chicago
anatomical@juno.com

Abstract. *INSIDE_OUT* took place on June 27th and 28th of 2003 at the University of Illinois at Chicago Electronic Visualization Laboratory, and on the AccessGrid. This case study examines a range of challenges and solutions to production and interaction design decisions for the mixed reality *INSIDE_OUT* performance event and the effectiveness of the solutions for transitioning passive audience members to active participation. The event featured live dancer improvisations by the Anatomical Theatre dance company based on the 'safety zone' choreographic exercise. *EnergyComposer,* a YG-based application provided the framework for a virtual performance space controlled through camera-based tracking of three different fluorescent color markers. The idiosyncrasies of the tracking system were successfully integrated into the aesthetic planning of the performance event. Two networked passive stereo displays were used to provide two different viewpoints of the virtual environment to the audience. In designing a multi-modal environment to accommodate a range of expert and inexperienced users, traditional wand navigation was disabled. Interactions that could fulfil both lengthy spatial exploration but required only a brief comprehension time were achieved by relying on full-body physical navigation for triggering events in a small, controlled environment.

1. Introduction

The *INSIDE_OUT* event was based upon the *EnergyComposer* application and served as the primary exhibition of the MFA thesis of Marientina Gotsis [1] at the Electronic Visualization Laboratory (EVL) of the University of Illinois at Chicago. The event was designed to be a significant departure from previous EVL-hosted MFA shows in that it would provide visitors the opportunity to both watch and actively participate in a mixed reality event. In addition, this event was the first to feature an optical tracking system developed at EVL. This system provided tracking for two networked passive stereo displays (C-Wall [2] and GeoWall [3]) using Ygdrasil (YG). Ygdrasil is built in C++ around SGI's OpenGL Performer™ visual simulation toolkit [4] and the CAVERNsoft G2 networking library [5]. YG extends Performer's hierarchical scene graph representation of the virtual world database by making the scene graph shared (**Figure 4**).

Figure 1. Lyndsae Rinio (left), LeAnn Vancil (middle) and Nadine Lollino (right)

Dance and performance artists have long used tracking systems and multimedia backdrops to enrich their performances. Groundbreaking performances, such as *Biped* (1999) by Merce Cunningham used a pre-rendered, abstracted dancer animation as a backdrop and as a virtual partner for the dance company. Other companies have developed real-time systems; one such system is the *MidiDancer* wireless sensory system by *TroikaRanch*. This system measures joint flexion to interactively affect video and audio used in the performances. *The Body Synth*™, a muscle-controlled MIDI controller created by Chris Van Raalte and Ed Severinghaus is often used by performance artist Pamela Z. Several of these systems were taken into account during the design of *EnergyComposer* and the planning of *INSIDE_OUT*. The *EnergyComposer* system was primarily influenced by David Rokeby's, *Very Nervous System* (1986-1990), a camera-based tracking system that translated gestures into sounds.

The major goals of the thesis show were defined as follows:

- To design a virtual performance application (*EnergyComposer*) that encourages full-body engagement for navigation of a virtual world instead of relying on hand control through a joystick, wand or spaceball®
- To ensure that the interaction system would adapt to varying user levels: complex enough for expert users (dancer) and straightforward enough to provide a satisfying experience to a first-time user
- To produce a mixed reality performance event that accommodates a medium-sized audience (15-20 people) and allows for their active participation (*INSIDE_OUT*)

Traditionally, EVL has used both tethered (Ascension pcBird®) and tetherless tracking (Intersense IS-900) for navigating virtual reality applications in the CAVE™ [6] and C-Wall systems. Neither tracking system provided a large enough tracking area, or the full mobility and ease of use required for the purpose of *INSIDE_OUT*. Use of the CAVE was excluded for two reasons: 1) EVL's four-walled CAVE provides limited space and freedom of movement when used by more than three people; 2) the enclosed space is not effective as a stage for a medium-sized audience. The solution was to convert one of the main EVL labs to a stage using a passive stereo screen (C-Wall) as a dynamic backdrop. Optical color tracking in development during the conception of the thesis was selected as the optimal system because it was tetherless and it enabled us to equip dancers and audience members with markers quickly and easily. The system prototype provided a tracking area close to CAVE dimensions (CAVE≈10Wx10Hx10D vs. stage≈10Wx8Hx10D, in feet) with the advantage of being functionally larger due to the lack of walls. Only two pieces of the corners closest to cameras were unable to be tracked due to field of view limitations. The entire physical stage area (tracked and non-tracked) amounted to more than 15x14 feet.

EnergyComposer served as the framework of virtual triggers and visuals that produced different animations and sounds in response to user movement. The application consisted of multiple layers of complexity to fulfil the choreographic theme of the performances and to accommodate differences in user mobility and physical body size.

The design of a stage in a computer lab environment presented additional challenges. The nature of the event and its intended audience necessitated many modifications of the lab environment to meet the needs of a performance event. The production process aimed to enhance the audience experience (especially if newcomers to virtual reality) or dance performance), and to entice audience members to actively participate in the show through their own improvisation. Visual design and aesthetic decisions will be minimally discussed in this paper. The following sections will detail the design of the *EnergyComposer*, the tracking system integration and several production solutions devised to enhance the 'learning-by-experience' process.

2. Designing a multi-modal performance space with *EnergyComposer*

The challenge of programming *EnergyComposer* lied in providing satisfying user feedback regardless of user skill and designing a goal-oriented virtual environment to encourage movement. The complexity of interaction had to satisfy long engagements by the dancers and brief explorations by audience members. A solution was to narrow down the movement scenario. We did this by setting the performance theme to a classic choreographic exercise known as the 'safety zone' improvisation [**7**]. This exercise is designed to augment the dancer's perception of empty space. Part of the process involves improvising short movement phrases that communicate feelings of comfort or discomfort in arbitrary zones. The exercise forbids becoming comfortable for too long in one particular space in order to allow for continual and fluid movement improvisation. The space has "then a very real effect on the dancer, forcing (them) to move differently, to behave in certain ways and not in others."[**8**]

EnergyComposer supported the safety zone exercise through a system of triggers that were 'fluid' in dynamics and forgiving in size. Three layers of triggers were programmed in the virtual application with an approximate 1:1 correspondence to the physical stage:

The first layer consisted of multiple static flat triggers placed at small intervals on the height axis (**Figure 2**). These triggers shifted the base soundtrack from A to G on the musical scale at the slightest detection of movement. This granted a sense of accomplishment to the user with very little effort to avoid the 'dead air' phenomenon and it also helped establish a meditative mood to benefit the improvisation theme.

Randomly dispersed clusters of box triggers of different sizes filled the entirety of the space and moved slowly in random trajectories (**Figure 2**). These triggers corresponded to visuals of the alphabet from A to Z. Upon intersection with the user's hand they output phoneme sounds and faded briefly in and out to mimic a 'highlighting' action. The randomness of the output audio established a sense of discomfort and frustration through the user's inability to spell a complete word.

The third layer of triggers consisted of seven large spherical volumes (~2 feet) that hovered in predetermined rectangular areas, strategically placed throughout the space. These objects were labeled A through G and corresponded to a unique color/shape animation and sound event. They carried distinct audio loops whose volume was amplified when the head tracker approached them, thus motivating the user to seek them out physically by listening to the audio feedback. This was made possible through software-based sound spatialization using *snerd*, YG's sound server. Intersection with the volume of the virtual objects resulted in lively and colorful events meant to provide the user with a greater level of achievement. Remaining true to the improvisational theme of the piece, these events were short-lived and the objects quickly returned to their more static behavior and smaller size and moved away from the user (**Figure 2** and **Figure 3**).

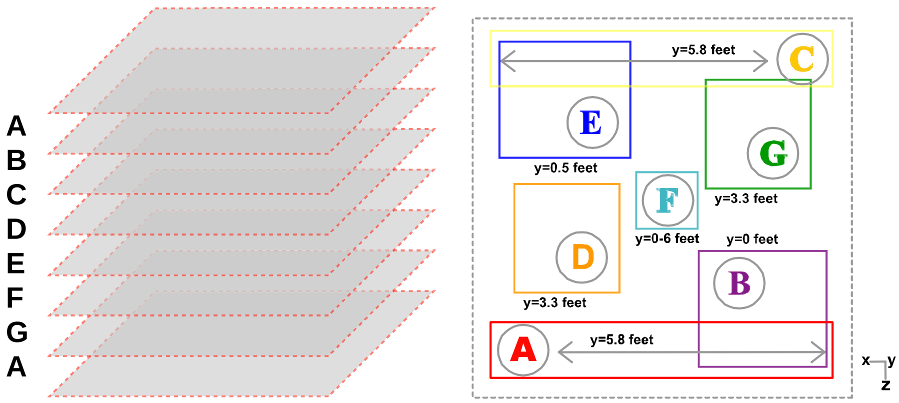

Figure 2. ***EnergyComposer's*** **first (left) and third (right) layer of triggers**

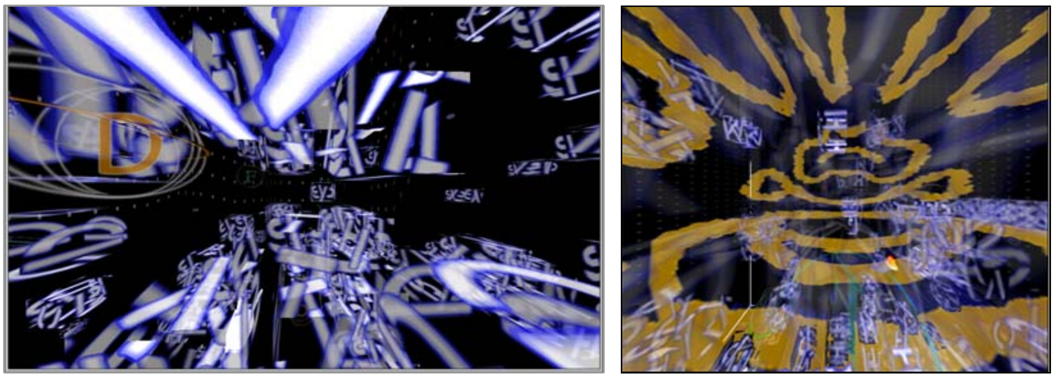

Figure 3. *EnergyComposer*'s second layer of visible triggers and 'D' and 'F' music/ animation objects (left) 'D' animation resulting from intersection with user hand (right)

Testing the experimental tracking system in conjuction with the virtual application and overall user experience was the most critical design step for the success of the application. A multitude of movement scenarios were accommodated for full-body engagement. The first two systems of triggers were alpha-tested in a CAVE by moving the wand with virtual navigation disabled. Adjustments were made later as we migrated to the single-wall display (C-Wall) and tetherless tracking system. Most of the behavioral programming for the larger and more eventful triggers was done during dance rehearsals. Triggers 'A' and 'C' could be triggered by reaching up with at least one hand (**Figure 1**, **left & Figure 6, right**). The corresponding soundtracks sounded optimal when standing tall near the object path. Triggers 'D' and 'G' could be triggered by bending knees and intersecting with the objects with at least one hand, or by kneeling on the floor and moving hands parallel to the floor (**Figure 1**, **right & Figure 6, middle**). The corresponding soundtracks sounded optimal when the user bent over into the orange and green zones, or was kneeling on the floor and tilting their head. 'E' and 'B' were optimally experienced when the user's head was near the floor with hands near their head, and in the case of purple, one had to come into full contact with the floor (**Figure 1**, **middle & Figure 6, left**). Trigger 'F' moved vertically in the center of the room and was optimally experienced when the user moved vertically with it.

3. Optical color tracking system integration

The tracking system utilized two Bumblebee® stereo cameras from Point Grey, Inc., to track a headband and two wristbands. Each Bumblebee camera contains two IEEE1394 cameras precalibrated to provide three-dimensional position data through stereo vision. A *camshift* algorithm from the OpenCV v0.9.5 (beta 3.1) library package was initialized to predetermined histograms for each of the three tracking markers, with one camera tracking the headband and the other camera tracking each wristband (**Figure 1**). The *x*-axis (horizontal) and *y*-axis (vertical) values were acceptably stable and generally noise-free for the purposes of the performance. However, the small disparity or baseline between the two cameras in each Bumblebee, compounded with noise from the *camshift* algorithm, led to more significant noise along the *z*-axis (depth), especially during *z*-axis movement (**Table 2 & 3**). To address this noise, a Kalman filter was introduced to smooth the *z*-axis data. Both statistical (see **6.Addendum**) and rehearsal observations were used to calibrate the Kalman filter to a degree which balanced a maximum of smoothing with a minimum of latency introduced by the algorithm. Tracking data was sent via the network from the tracking application, through System V shared memory, to a *trackd* device module (trackdSDK®) and subsequently sent to a *trackd* client and *EnergyComposer* (**Figure 4**).

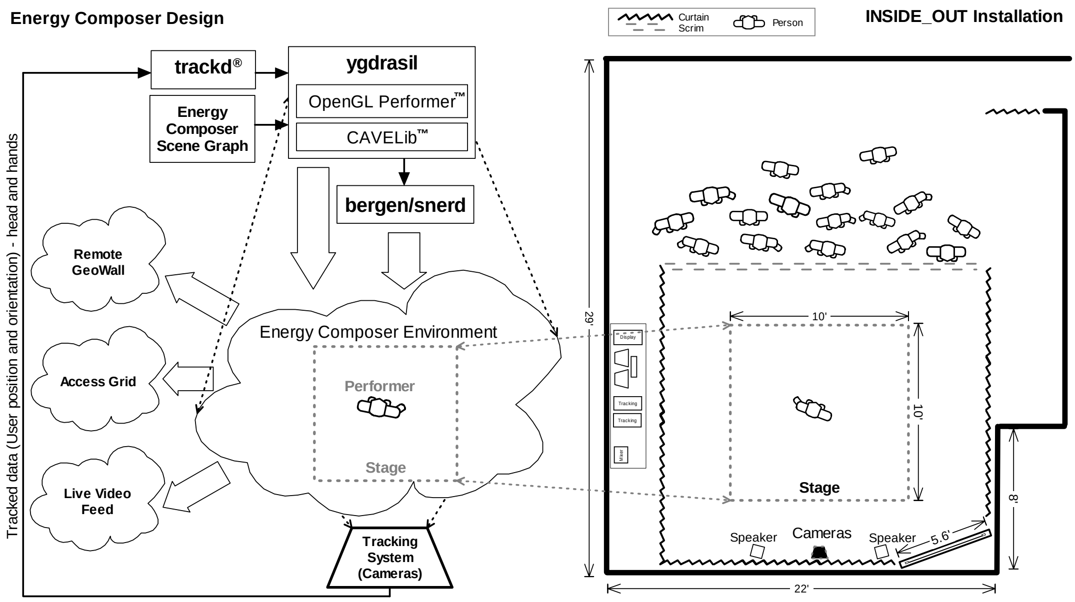

Figure 4. Component (left) and installation (right) diagrams

In anticipation of having the audience face the tracking cameras, a fluorescent color palette was assigned to the markers along with UV lighting to avoid clothing and background interference within the tracking system's field of view. A wide range of materials and paints were tested. The best possible results were achieved by custom fabrication of markers that were low-cost, easy to make and wear, and provided a consistently smooth and symmetric shape for a more stable centroid, regardless of the markers position.

In order to completely and evenly light the performance-space, the lab's institutional fluorescent ceiling lights were replaced by UV lights. The 'black light' look and feel was integrated into every aesthetic aspect of the show (*EnergyComposer* visuals, aural landscape, stage and costumes). The stage palette was isolated to black and white to minimize visual interference with the optical tracking. EVL's largest computer lab/conference room was transformed into a stage area by hanging black felt curtains, a translucent scrim and a custom designed and hand-built floor with stylistic elements that fluoresced under black light (**Figure 5**). The floor was necessary for injury prevention and cushioning and successfully withstood a week of rehearsals and two days of show traffic. The dancers wore full-body white costumes and UV-sensitive face makeup to be made visible to the broadcast cameras. Due to the sensitivity of the color-tracking, final calibrations proved necessary to compensate for the brightness and glow of the white scrim and the performers' white costumes under black light.

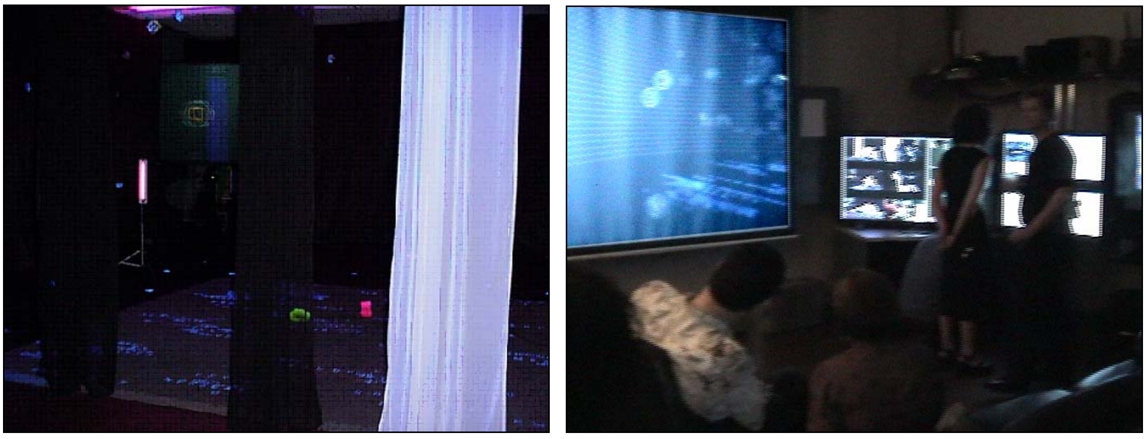

Figure 5. Stage (left) and waiting room (right)

Figure 6. Audience member interactions with *EnergyComposer* onstage after performances

4. Production solutions for creating a 'learning-by-experience' event

Audience expectations present a major challenge for producing an MFA show in a computer lab. Museums, galleries and theatres are designed with bare walls without visual clutter and with the provision of open space to allow observation from any angle or position. Therefore visitors come psychologically prepared with some preconceived notions of the physical space conditions. In the case of virtual reality events at EVL, many visitors come with no prior knowledge or experience of the technology, and some come with an understanding of the technology but no perception of a university research space as an exhibition or performance space.

When the technology is not portable enough, high-tech art often has to be shown in computer labs or spaces that were never designed for cultural events. Audiences can get easily distracted if they are overwhelmed by the technology of an art show. EVL's high-tech clutter posed the risk of confusing and overwhelming the incoming audience. We had to devise a strategy to 'psychologically' prepare different types of visitors and to educate them in advance of their experience-to-be-had and prepare them for active participation.

The event made use of two rooms: a waiting room and a performance space (**Figure 5, left & right**). The latter room featured the custom-designed stage and a C-Wall as a dynamic backdrop. The Anatomical Theatre dance company performed live ten-minute dance improvisations at one-hour intervals. After each improvisation, the public was invited to participate by improvising in-between each scheduled performance. The C-Wall displayed a narrow view of the performers' virtual perspective from within *EnergyComposer's* environment for the audience (*INSIDE* view). The waiting room consisted of a live video feed of the stage, a GeoWall providing the waiting audience with a view of the dancer's avatar performing in the virtual environment (*OUTSIDE view*), and various supporting screens (plasma panel with AccessGrid [**9**] video feeds, rehearsal clip archive and 4x4 tiled display with credits, event details and background information).

The waiting area provided the first step of psychological preparation (**Figure 5, right**). Since the performance room could only accommodate groups of 15-20 people, reservations were required and waiting visitors got a 'green room' preview of the performance's elements before they went in for the show. The second step was the brief introduction (~1 minute) made to the audience before each performance. This introduction by either a dancer or presenter briefly demonstrated what was being tracked, described the improvisational theme, and verified that the visuals and audio they had experienced in the waiting room were output in real-time and in response to the tracker's position. Following the introduction, the dancers' improvised performances served as an educational medium for the use of the space, and as the third step of psychological preparation. At the end of each dancer's improvisation, the performer or presenter invited willing audience members onstage to improvise the exploration of their own safety zone. They were guided and given movement suggestions as needed (**Figure 6**).

The intimacy of the stage's and the audience proximity to it were also effective in engaging audience members with the event. Very few chairs were made available to encourage the audience to move around freely during each performance. Additionally, the translucent scrim-like curtains between the audience and stage provided a feeling of individual isolation to the performer without compromising the visual appeal of the space. The requirement that everyone removes their shoes upon entry to the performance space (implemented to protect the delicate floors), unexpectedly served as a rite of passage for audience members.

5. Discussion/Conclusions

Testing for different scenarios was crucial from the beginning of designing *EnergyComposer*. Major adjustments were made to the virtual environment upon transition to the physical stage and during rehearsals with the performers. This resulted in being able to use the same application for both performers and members of the audience without requiring any intervention. This was desirable to avoid interrupting the event by any requirements of reinitializing, or reconfiguring the application.

With regular testing and rehearsals we were able to 'calibrate' and compensate for specialized tracking requirements and account for differences in movement style, height, mobility and expressiveness between performers. The two levels of complexity provided by *EnergyComposer*, created a dynamic space that was suitable for the choreographic theme of the show and was adaptable to the abilities of a range of potential users.

Besides the potential cost-effectiveness of the software-driven optical tracking, this alternative tracking system's lack of constraints provided a liberating experience to both author and user. Given the popularity of gaming-style navigation and the increasing incidence of repetitive stress injuries in the computing world, there is considerable room for the development of alternative interfaces.

Guiding VR audiences from active to passive participation requires more than devising motivation and ambience strategies. The use of dancers as teachers for movement motivation was a fruitful experience for the audience and for the dancers, because both were able to interact with each other and engage in an even exchange of imitation and extrapolation that is traditionally not implemented in performance shows.

6. Addendum

This section describes the noise and accuracy of the optical color tracking system used for the *INSIDE_OUT* event. The locations of our test sample points are shown on **Figure 7**. The tracked target was placed on a tripod and placed at the indicated points. The *y* position of the target was not varied, so any data describing the noise or accuracy of the *y* values should be considered with caution. Let p_i be the *i*th point and $p_{i,k}$ be the *k*th coordinate of point *i*.

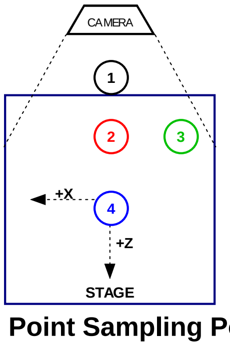

Figure 7. Point Sampling Positions

The camera height was 1.255m. The coordinate system used has the *z*-axis pointing out from the camera, the *x*-axis pointing to the right in the image, and the *y*-axis pointing down in the image. Noise was measured by reading the reported position of a stationary object over a period of time. See **Table 1** for the noise values. Two thousand data points were collected for each sample.

Measured point	x	y	z
1.	-0.1275m	-0.085m	1.07m
2.	-0.17m	-0.085m	2.21m
3.	-1.0125m	-0.085m	2.14m
4.	-0.16m	-0.085m	3.495m

Table 1: Measurements

Point and coordinate	μ	σ^2	Abs. Error of μ
$p_{1,x}$	-0.0747	1.4851e-08	5.2830e-02
$p_{1,y}$	-0.0894	2.0102e-08	4.3770e-03
$p_{1,z}$	1.0431	5.7483e-05	2.6853e-02
$p_{2,x}$	-0.0944	3.1458e-11	7.5550e-02
$p_{2,y}$	-0.0840	1.1178e-07	9.6101e-04
$p_{2,z}$	2.1744	3.3855e-04	3.5852e-02
$p_{3,x}$	-0.9583	2.8389e-05	5.4221e-02
$p_{3,y}$	-0.0886	3.1703e-08	3.5878e-03
$p_{3,z}$	2.1244	7.2861e-04	1.5571e-02
$p_{4,x}$	-0.0777	9.6994e-12	8.2315e-02
$p_{4,y}$	-0.0638	2.1331e-09	2.1205e-02
$p_{4,z}$	3.4142	1.0373e-03	8.0824e-02

Vector	Actual difference	Tracked difference	Error
$\|\|p_{1,x} - p_{3,x}\|\|$	0.8850	0.8836	0.0014
$\|\|p_{1,y} - p_{3,y}\|\|$	0.0	0.0008	0.0008
$\|\|p_{1,z} - p_{3,z}\|\|$	1.0700	1.0813	0.0113
$\|\|p_{1,x} - p_{4,x}\|\|$	0.0325	0.0030	0.0295
$\|\|p_{1,y} - p_{4,y}\|\|$	0.0	0.0256	0.0256
$\|\|p_{1,z} - p_{4,z}\|\|$	2.4250	2.3710	0.0540
$\|\|p_{2,x} - p_{3,x}\|\|$	0.8425	0.8638	0.0213
$\|\|p_{2,y} - p_{3,y}\|\|$	0.0	0.0045	0.0045
$\|\|p_{2,z} - p_{3,z}\|\|$	0.0700	0.0497	0.0203
$\|\|p_{3,x} - p_{4,x}\|\|$	0.8525	0.8806	0.0281
$\|\|p_{3,y} - p_{4,y}\|\|$	0.0	0.0248	0.0248
$\|\|p_{3,z} - p_{4,z}\|\|$	1.3550	1.2897	0.0653

Table 2 & 3. Noise Results (left) / Relative Point Differences (right)

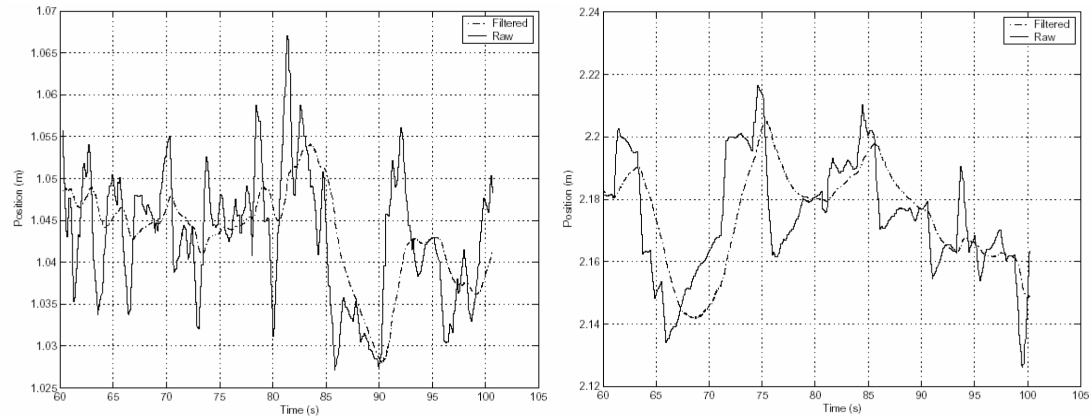

Figure 8. Data at p_1 (left) p_2 (right) on the *z*-axis

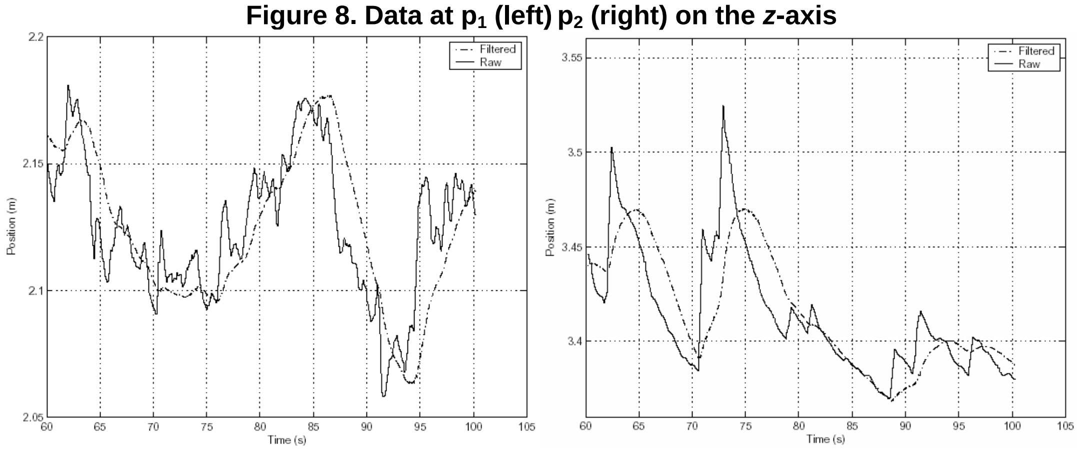

Figure 9. Data at p_3 (left) p_4 (right) on the *z*-axis

7. Acknowledgments

This project would have not been possible without Geoffrey Baum, Kang Sun and the generous volunteering of the Anatomical Theatre company, its artistic director, Robynne H. Gravenhorst and its members, Nadine Lollino, Lyndsae Rinio, LeAnn Vancil and Katie Wells. Special thanks to thesis advisors Dan Sandin, Drew Browning and Marta Huszar. Sound design and hardware was provided by Joe Reitzer. Video documentation was made possible by Seung Kang and Natt Mintrasak. Additional support was received by a great community of EVL student and staff volunteers.

The virtual reality and advanced networking research, collaborations, and outreach programs at the Electronic Visualization Laboratory (EVL) at the University of Illinois at Chicago are made possible by major funding from the National Science Foundation (NSF), awards EIA-9802090, EIA-9871058, EIA-0115809, ANI-9980480, ANI-9730202, ANI-0123399 and ANI-0129527, as well as the NSF Partnerships for Advanced Computational Infrastructure (PACI) cooperative agreement ACI-9619019 to the National Computational Science Alliance. EVL also receives funding from the US Department of Energy (DOE) Science Grid program and the DOE ASCI VIEWS program. In addition, EVL receives funding from the State of Illinois, Microsoft Research, General Motors Research, and Pacific Interface on behalf of NTT Optical Network Systems Laboratory in Japan.

8. References / Notes

[1] Currently employed by the University of Southern California, Interactive Media Division, mgotsis@cinema.usc.edu.

[2] D. Pape, J. Anstey, G. Dawe, Low-Cost Projection Based Virtual Reality Display, *The Engineering Reality of Virtual Reality 2002 SPIE Electronic Imaging: Science and Technology 2002,* San Jose, CA 01/24/2002 - 01/24/2002.

[3] J. Leigh, G. Dawe, J. Talandis, E.He, S.Venkataraman, J. Ge, D. Sandin, T.A. DeFanti, AGAVE : Access Grid Augmented Virtual Environment, *Proceedings of the AccessGrid Retreat*, Argonne, IL 01/16/2001 - 01/16/2001.

[4] J. Rohlf, J. Helman, IRIS Performer: A High Performance Multiprocessing Toolkit for Real-Time 3D Graphics,' *Proceedings of SIGGRAPH 94 Computer Graphics Conference*, pp. 381-395, 1994.

[5] K. S. Park, Y. J. Cho, N. K. Krishnaprasad, C. Scharver, M. J. Lewis, J. Leigh, A. E. Johnson, ``CAVERNsoft G2: A Toolkit for High Performance Tele-Immersive Collaboration,'' *Proceedings of the ACM Symposium on Virtual Reality Software and Technology 2000,* pp. 8-15, 2000.

[6] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, `Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE,' *Proceedings of SIGGRAPH 93 Computer Graphics Conference*, pp. 135-142, 1993.

[7] L. A. Blom, L.T. Chaplin, The Intimate Act of Choreography. pp. 46-47, ISBN: 0-8229-5342-0, University of Pittsburg Press, 1989.

[8] Ibid.

[9] Argonne National Laboratory / NCSA Alliance, ``Access grid web page,''
(URL: http://www-fp.mcs.anl.gov/fl/accessgrid/)