SMALLab: a mediated platform for education

David Birchfield
Arts, Media and Engineering
Arizona State University
+1.480.965.3155
dbirchfield@asu.edu

Thomas Ciufo
Arts, Media and Engineering
Arizona State University
+1.480.965.9438
thomas.ciufo@asu.edu

Gary Minyard
Theatre for Youth
Arizona State University
+1.480.965.9438
gary.minyard@asu.edu

ABSTRACT
In this article we describe the design and realization of SMALLab, a new Student Centered Learning Environment [SCLE] that utilizes interactive digital media in a multimodal sensing framework. We draw on methodologies from interactive art, computer music composition, animation, and educational technology to facilitate learning with digital media that emphasizes collaboration and human-to-human interaction. We describe the realization of this learning platform, and discuss the design of a complementary curriculum that helps students develop a deeper understanding of movement and sound.

Categories and Subject Descriptors
K.3.1 [**Computers and Education**]: Computer Uses in Education – *collaborative learning, computer-assisted instruction*

General Terms
Algorithms, Documentation, Performance, Design, Experimentation.

Keywords
K-12 Education, Situated Multimedia, Learning, Experiential Media, Student Centered Learning Environments, Interactivity, Constructivism.

1. INTRODUCTION
For K-12 education to keep pace with the rapid technological advances in other sectors of our society, we must develop new approaches to education that harness emerging technologies, enable collaborative learning, bridge the physical/digital realms, and prepare all students for the dynamic world they are entering. We must devise innovative strategies that engage the creativity and innate curiosity of our students, and we must design educational activities that engage minority and underserved students and those with diverse learning styles.

Our approach to education is informed by related research in open –ended educational philosophies that support exploratory and active learning approaches. Constructivism posits that learners construct much of what they learn [1], and emphasizes the necessity of play and exploration in self-guided learning, particularly in mediated environments [2]. Instructional teaching can play a complementary role in these learning environments where advanced students and teachers providing a “zone of proximal development” that speeds learning [3]. Howard Gardner’s thesis in the *Theory of Multiple Intelligences* [4] posits that there are multiple forms of knowledge, including musical and kinesthetic, that support and reinforce one another. This research supports evidence that students have diverse learning styles that are influenced by their life experiences, cultural background, and genetic predisposition. Active and visual learners are examples of students who thrive in environments that differ from those provided by conventional book and lecture teaching methods.

Figure 1. Students and Instructor interact in SMALLab

Dynamic, interactive media offers a great number of opportunities to engage learners, but they have been underutilized in classrooms to date. Fixed media, such as films, can serve to vary instructional methods, but they do not engage active learners. There has been extensive prior work in the development of distance learning frameworks that can deliver educational content to remote learners [5]. Many distance learning courses do offer opportunities for mediated interaction through email, chat rooms, and posting sites. However, students’ learning does not benefit from the direct human-to-human interaction that can motivate and speed their learning as in classroom settings. Educational video games have been introduced in a number of classrooms as a complement to conventional teaching methods [6]. Games can be a powerful motivational force. However, because of their reliance on mouse/keyboard/controller interfaces, most games offer only limited modes of engagement that restrict the naturally expressive and social capabilities of students.

One of the most exciting strategies for engaging learners that has emerged in recent years is the development of technology-based learning systems that are highly inquiry-based. The most effective such learning systems are those that are learner-centered, knowledge-centered, and assessment centered [7]. Student Centered Learning Environments (SCLE) encompass a great variety of such learning systems, including open-ended learning environments, microworlds, goal-based scenarios, cognitive apprenticeships, and constructivist learning environments [8]. SCLEs have been extensively used for math and science education but have not as commonly been developed for arts and media education [7].

In this article we describe our work in the development of an innovative SCLE that engages diverse learners in an extensible framework. In Section 2 we describe SMALLab, a situated multimedia environment for learning. We then describe portions of an arts/science curriculum that is delivered with the media system. Finally, we summarize our preliminary results and describe our future work.

2. SMALLab

Central to our work is the development of the Situated Multimedia Arts Learning Lab [SMALLab]. SMALLab is a physically situated, interactive media environment developed by a team of artists, educators, engineers, media designers and psychologists at the Arts, Media and Engineering program at Arizona State University. We conceive of the system as a platform for experiential education that engages multi-sensory capabilities and natural creativity of students. Here we describe the hardware and software architecture of the system.

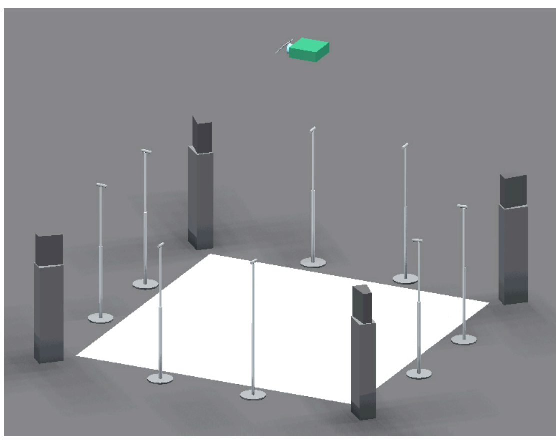

Figure 2. Schematic of SMALLab hardware arrangement

Figure 2 shows the hardware arrangement of SMALLab. The physical space fills a 15' X 15' footprint and extends up to a 12 foot ceiling.

SMALLab's open framework facilitates social interaction and collaborative learning. It is designed with an emphasis on constructive as opposed to instructive modes of learning, but supports a diversity of teaching methods. The system is low-cost, reconfigurable, and can be transported for easy installation in a classroom or community center. Our work has focused on the design and realization of an environment that emphasizes human-to-human interaction and supports collaborative learning approaches. We are mindful of the real world logistical challenges that face schools today. Thus, we have sought to design a system that addresses the needs of students and is suitable for classroom use.

A computing cluster drives the interactive system with custom software for fused multimodal sensing, context modeling, and dynamic visual and sonic feedback. The lab allows teachers and groups of students to interact with one another and their composed media worlds through free play, structured movement, and vocalization. The use of a multimodal sensing apparatus engages the naturally expressive capabilities of K-12 students. Multimodal real time feedback provides a platform for curriculum design that addresses the needs of students with diverse learning styles.

We have adopted a modular and extensible approach to interaction design that progresses concurrently with empirical testing with students and teachers, and is guided by learning objectives and outcomes. A local network supports communication between modules via multicast UDP. Here we present an overview of the system architecture.

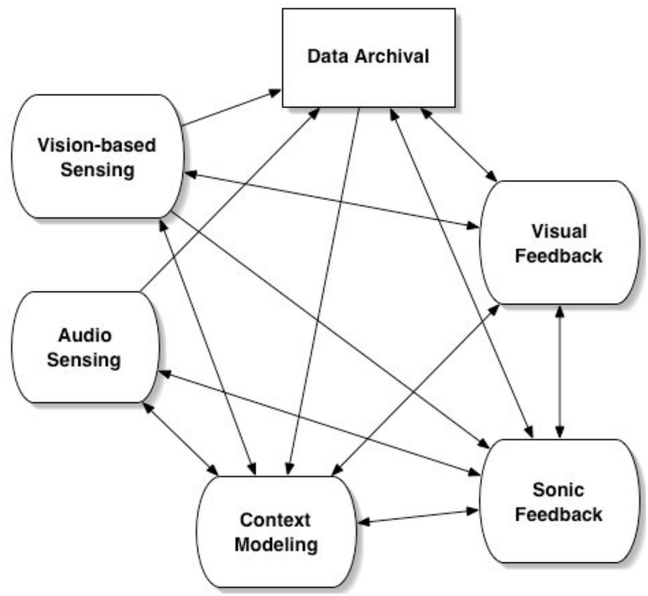

Figure 3. Schematic of interconnected software modules

2.1.1 Vision-based Sensing

Three *Point Grey Dragonfly2* (VGA resolution at 60FPS) video cameras are mounted around the active space. These are connected serially on the same IEEE1394 bus to the PC through a single port. The videos from these three cameras are automatically synchronized. The tracking system uses color information to locate the 3D positions of multiple objects in the space. 2D candidate object locations in each camera view are first obtained using the Continuously Adaptive Mean Shift (CAMSHIFT) algorithm [9, 10] provided in OpenCV [11]. Triangulation is then used to find out the 3D object locations. Before each session, the external camera calibration parameters are computed using the multi-camera calibration toolbox [12].

The current system is capable of tracking up to five objects at forty frames per second. This three-dimensional position data, along with velocity and magnitude, is sensed and broadcast to other modules. We are experimenting with various scenarios that combine student and object tracking in order to detect movement patterns that arise from collaboration between students. For example, the sensing system is capable of extracting patterned movements such as primitive shapes that emerge over short time spans.

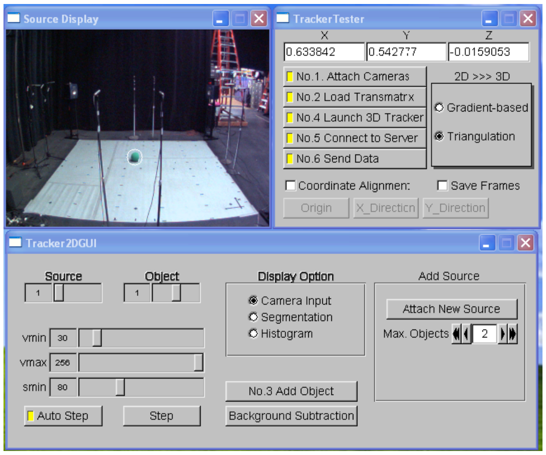

Figure 4. Multi-camera, color based vision tracking system

2.1.2 Audio Sensing

In addition to this visual sensing framework, we have implemented a module that allows students to interact through vocalizations. Through experimentation and exploration, students can use their own voices to develop a deeper understanding of sound and movement.

An eight element microphone array is positioned around the space, and is utilized in two ways. First, raw sounds can be recorded in real time and inserted into a database of soundfiles for later use in interactive exercises. Second, sound analysis modules extract features from the incoming audio that can be used to drive interactive media in the space. For example, we are currently developing a learning module where students explore and describe aspects of physical movement by shaping related contours with the pitch of their voices. In this exercise, a rising and falling pitch contour in the voice might be coupled with the trajectory of a thrown ball.

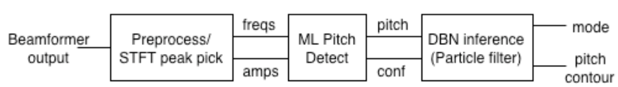

Figure 5. Audio sensing processing diagram

Our engine for tracking pitch contours is sketched in Figure 5. This approach is extensible to other features and other types of sonic gestures that might involve amplitude shapes and timbral attributes. Here, the signal is picked up by the microphone array using a simple delay-sum beamformer focused at a given location in the active space. The beamformer signal is preprocessed by lowpass-filtering and downsampling at a rate of 8 kHz since in most applications, pitches above the Nyquist rate of 4 kHz are rarely found.

Sinusoidal peaks are estimated from a hopped short-time Fourier transform using the parabolic interpolation method of Smith and Serra [13]. Each peak consists of an amplitude and frequency value. Next, for each STFT frame, a maximum-likelihood pitch detector [14] is applied to compensate for missing and spurious peaks. Finally, a simple dynamic Bayesian network (DBN) modeling piecewise up/down/constant pitch motions is used in post-processing. The DBN identifies modes (time intervals) where the pitch is going up, down, or staying at a constant level, as well as intervals where no pitched sounds are being made. Additionally, the DBN smoothes over occlusion artifacts caused by transitory interference (e.g., audience noise, someone tapping the microphone). The mode information and the smoothed output pitch contour shape are broadcast to other processing modules in SMALLab.

2.1.3 Sound Feedback

Sound plays a critical role in SMALLab and many of our learning exercises depend on the immersive, three-dimensional nature of sound. Four raised speakers and one subwoofer surround the space. We have developed software to project spatialized, reactive sound into the space. An extensible database of soundfiles supports this module, and through in-classroom and web based interfaces, both students and teachers can contribute sound content.

Our current work in this area extends prior research in the development of interactive installations [15], and borrows techniques from *musique concréte* [16] and concatenative music composition [17]. We have designed specialized learning modules that allow students to record their own sounds from the environment and then discuss, share, and interact with those sounds in SMALLab. During classroom activities and via the Edulink website (described in Section 2.1.7), these collected sounds can be annotated and auditioned by students and teachers. These annotations inform our models for interaction and allow for the delivery of sonic feedback that is adaptable to individuals and groups of students.

2.1.4 Visual Feedback

A top-mounted video projector displays interactive visual content on the floor of SMALLab. In contrast to related work CAVE environments [18], we have sought to develop an architecture that promotes social interaction and collaboration among groups of students. The absence of projection screens surrounding the space subverts many biases of screen-based media, and creates an open physical environment.

Our feedback frameworks utilize still images and video clips that are collected and annotated by the students as described in Section 3.2. In addition, we have developed a three-dimensional graphics engine using OpenGL within the Max/MSP/Jitter programming environment [19]. Interactive graphics modules are coupled with specialized learning exercises to assist in the development of students' understanding of spatial relationships, movement dynamics, and activity patterns.

2.1.5 Continuous Archival

We have developed a software module with a MySQL database backend to archive all sensing and feedback data in real time. This stored data can be accessed for a number of purposes. First, during a given learning session, students can recall and replay movement passages to reflect on their activities. Second, this data is used to update real time context models that can inform our feedback mechanisms. Finally, this data can be used for evaluation and assessment by providing a detailed view of students' activities over multiple time scales. For example, we are currently examining the relationships between student movement patterns and sonic feedback to better understand how sound can be used to influence movement in service of more efficient learning.

2.1.6 *Context Models*

Informed by domain knowledge from the arts, sciences, education, and psychology, we are developing computational models of context. Currently this work is focused on the understanding of students' sound and movement interactions. The model first extracts the *elements* and *principles* underlying the artifacts of each student's interaction. Second, it tracks how this context develops over time and is individuated amongst different students. The context consists of *lexica* representing the students' *modes of interaction* within the space, plus a *syntax* describing how lexica relate across different times or modalities.

For example, in one of our developing learning modules, students are asked to express an abstract idea, such as "joy", using their voices and interactive sounds in SMALLab. Corresponding sound-oriented lexica include "swooping" (up-down) pitch contours in the students' vocalizations, identical or contrasting contours in the sounds projected by the feedback apparatus, and an idea that the student is expressing through sound and movement. A corresponding syntactic rule is that, given a set of sounds tagged by a particular student as related to "joy", the pitch contour of the student's vocalization is more likely to match that of the feedback immediately preceding the vocalization provided that both share "up-down" patterns. Conversely, they might be more likely to contrast if the feedback contains strictly "up" *or* "down" gestures. Such complex syntactic relations may differ from student to student, but they nevertheless form an integral part of how each student constructs the concept, "joy," through interaction. In a broader sense, syntax may encompass the patterns in sound that emerge from the relationship of three-dimensional position and the dynamics of movement during various learning modules.

2.1.7 *Edulink Online Resource*

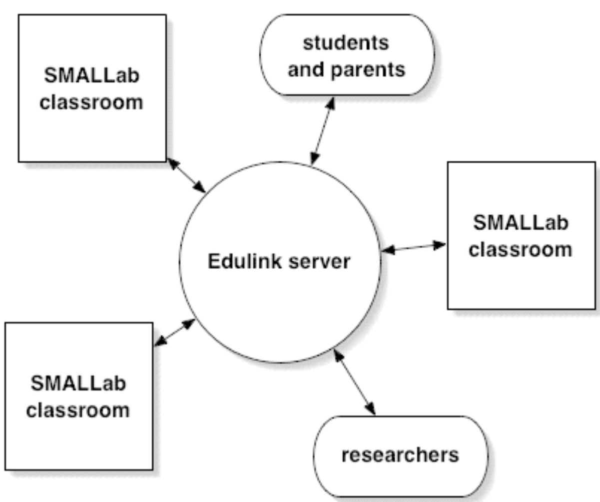

Figure 6. Edulink extends beyond and between classrooms

Our work in this project is focused on the development of a SCLE that facilitates direct human-to-human interaction between groups of students and teachers. Nonetheless, the Internet plays a powerful role in extending beyond and between classrooms. We have developed the Edulink website [20], an interactive online portal that links between SMALLab installations, extends classroom activities to the homes of students and parents, and serves as a resource for teachers and researchers. Edulink provides a navigation interface for archived movement data that can be downloaded and manipulated in the classroom. It facilitates uploading audio and visual media content. That content can be annotated by students and teachers, then downloaded for use in the classroom. Both of these mechanisms act as direct links between students and teachers. This link allows interaction between diverse groups of learners who bring varied perspectives to the classroom. In this fashion, the classroom can extend beyond the four walls of one site to encompass a large and varied student population.

Finally, our researcher team periodically shoots video documentation of learning sessions. This video documentation is then uploaded to the Edulink site. There, researchers can view the documentation and submit their annotations that assess the tools, techniques, and progress of students in the sessions. This annotated observational data is a foundation of our hybrid assessment methods. Furthermore, because this documentation is searchable it facilitates teacher training, as ideal teaching practices can be observed and modeled in the classroom.

3. CURRICULUM DESIGN

Our development of the SMALLab and related curricula embraces constructivist and experiential learning philosophies. Schank and Cleary [21] describe key characteristics of interactive learning environments that inform our work. In particular, we strive for diverse environments where students can learn by doing, learn through reflection, and learn through exploration.

3.1 Learning Objectives

Two overarching and complementary learning objectives guide our work: students should develop a deeper understanding of movement and sound. These broad objectives are subdivided into demonstrable skills that are taught and evaluated through specialized curricula. We are using a modular approach to curriculum design wherein sets of specialized learning exercises can be grouped to form complete sessions. Multiple sessions can be strung together to form learning paths spanning across weeks or months in a classroom.

We have subdivided our overarching learning objectives into a set of skills that we are currently building into our curriculum. Specifically, we have designed learning modules and assessment strategies that guide students to:

- deconstruct properties of sound and movement
- reproduce properties of sound and movement through vocalization and movement in the space
- articulate properties of sound and movement with increasingly sophisticated language
- recognize abstracted sound and movement features from composed and real world environments
- navigate complex sonic environments
- use interactive sound and visual feedback to communicate an idea

We will now describe several inter-related learning modules that allow students to develop and demonstrate skills in these areas through a number of instructional methods.

3.2 Learning Modules

The ***Video Harvesting*** learning module allows students to record material from environments outside the classroom that can then be brought into the SMALLab database of media content. The module is designed to help students abstract basic properties of movement from complex environments, and to develop an increasingly sophisticated vocabulary to describe movement. In this exercise, a facilitator leads students on a guided walk with a video camera in an outdoor environment. As they walk, students are asked to discuss movements they see and to articulate the diversity of patterns, dynamics, and degrees of complexity that are at work. When students discover compelling movements, they are asked to shoot video footage of the events. Examples of harvested video includes ant traces in sand, traffic movement at various distances, clouds slowly moving across the sky, and close-up footage of a leaf blowing in the wind.

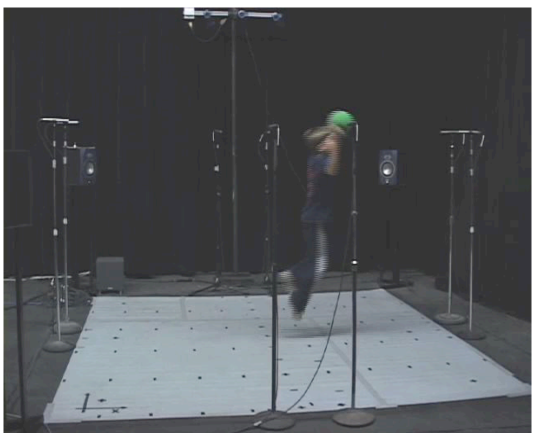

Figure 7. A student creates a *Sound Poem* in SMALLab

The ***Video Movement Mimicry*** module explores students' ability to reproduce properties of movement through physical interaction in the space. During the exercise, students' harvested video is projected onto the floor of SMALLab. The facilitator leads a discussion of the properties of the video and asks student to move the color object in a way that mimics the observed movements. As the student moves, their movements are archived and stored. The student movement pattern is replayed by a virtual ball that is projected onto the floor of the space and mixed against the original video. Students can compare their movements against the original video and engage in active discussions regarding the attributes of movement. This exercise provides an opportunity to experience movement, and to engage complex movement patterns through experimentation and exploration. Their learning is motivated by the use of their own video clips and movement patterns that they selected. The exercise serves as a catalyst for discussion of the forces that drive complex movement including gravity, inertia, and friction. This embodied interaction with dynamic media provides the opportunity for students to make their learning visible to teachers and themselves.

The ***Sound Poem*** is the most sophisticated module in our curriculum. It supports free exploration and creation in an interactive sonic environment. It is designed to improve students' ability to navigate complex sonic environments and use interactive sound/movement to communicate an idea. For this exercise, we have created an interactive environment that fully utilizes the multimodal sensing, and real time feedback apparatus. The primary goal is to provide students with an opportunity to express themselves through movement and sound, and learn about themselves in the process. Most traditional learning environments are physically passive, and don't allow the student to use their bodies. As pictured in Figure 4, this exercise encourages the students to construct and shape their sonic environment using full body movement and gesture.

This sound feedback engine, built in Max/MSP, contains five sound file players that can be spatialized into two-dimensional locations in the space. The volume of the five players is controlled by the x/y position of the ball in the space, so as the ball gets closer to a given location, the sound assigned to that location also gets louder. These locations are reconfigurable, and the proximity to gain control curves for each sound location can be changed according to a student's request or curriculum needs. We have developed an extensible database of sound files for this playback system, consisting of both synthetic and recorded sounds. Through interaction with the sounds, students can build individual sets that express a particular idea as they move in the space. In addition to the location-based mixing, the velocity of the ball modulates the playback speed and pitch of a given sound. This allows students to drastically manipulate the quality of a given sound by staying in the same general location, but quickly moving the ball. The z-plane location of the ball controls the influence of a band pass filter on each sound. This allows students to sweep the filter range from low to high as they move the ball from the floor to up over their head. Filter resonance varies directly with the accumulated ball velocity.

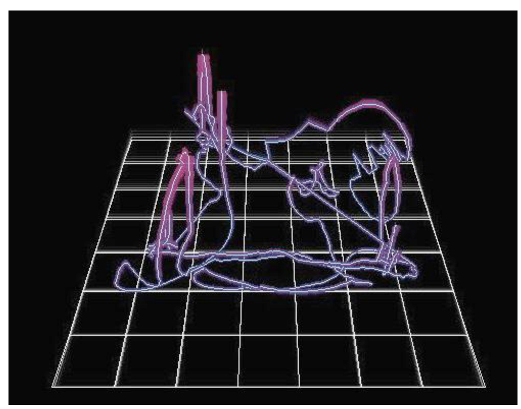

Figure 8. 3D Archive traces of a student's movements

The ***Sound Poem Reflection*** exercise combines the exploratory model of the Sound Poem with guided discussions from the ***Video Harvesting*** and ***Movement Mimicry*** exercise. This learning module allows students to develop a deeper understanding of the relationships between sound and movement embedded in their Sound Poems. Again, as students create their Sound Poems, all movement data is archived, and the facilitator can replay the archived data as a virtual ball replays their movement and sound creations in real time. Students can now reflect on the movements and sounds from a passive perspective to discuss the patterns that emerge. To aid in understanding of a longer history of their movements, an accumulated composite trace of their three-dimensional movements (eg. Figure 8) is also projected on the floor and rotated for multiple viewing angles. This visualization is enhanced by dynamically increasing the line thickness, color warmth, and intensity as the ball moves upward on the z-plane. Consequently, students' movements generate a perceptual color model that clearly reflects their activities. The facilitator can lead the students in discussions of their movement and sound interactions, reflecting on similarities and differences across the poems, and articulating various salient features. The orientation of such discussions around the students' actual sound interactions motivates their learning and facilitates richer understanding.

4. PRELIMINARY RESULTS

We are currently undertaking pilot studies with middle school students who participate in after school programs on the campus of Arizona State University under the auspices of the Herberger College for Kids. This initial study has focused on a small set of learners, but has yielded encouraging results that have already lead to revision and improvements in our methods.

Firstly, we observe that students are highly motivated to participate in these learning exercises. Often, they are initially shy and reserved in their interactions, but through the introduction of increasingly active movement and sound exercises, students lose their inhibitions and grow noticeably more engaged and expressive in their use of dynamic sounds and movements.

Secondly, students report that they enjoy the social and collaborative aspects of SMALLab. An important design goal has been to develop a situated learning environment that supports human-to-human interaction and promotes collaborative learning exercises. While many screen/mouse/keyboard interfaces can isolate students and limit their creative capabilities, we are encouraged that here, students are able to engage with their peers in a meaningful way.

Thirdly, we can see that our use of interactive media is having an impact on students. For example, we expected that the interactive sound design utilized in the Sound Poem engine should bias students to move in particular ways. Specifically, the topology of virtual sound locations should reinforce 2D spatial movements toward the four corners of the space. The interactive relationships between z-plane movements and object velocity should encourage shaking and swooping physical gestures. Indeed, analysis of archived movement data reveals that these similarities emerge for individuals and across groups of students.

5. CONCLUSIONS AND FUTURE WORK

We have described the design, realization and preliminary testing of SMALLab. We have designed an innovative curriculum for structured learning that helps students to develop a deeper understanding of movement and sound. Our pilot studies with middle school students have been promising, and we are encouraged to extend and enrich these methodologies.

In our future work with the SMALLab platform will focus on all aspects of sensing and feedback. In particular we will develop more sophisticated models for data fusion that will yield more robust tracking for multiple students and objects in the environment. Our models for interactive feedback will grow more malleable and adaptive as students and teachers gain confidence and experience with the system. We will extend the current curriculum by designing more learning modules and experimenting with varied combinations in a given learning session. Finally, we will embed SMALLab systems in numerous regional and national classrooms in order to undertake longitudinal study of the impact of our work across extended timeframes and across multiple sites.

6. ACKNOWLEDGMENTS

We gratefully acknowledge that these materials document work supported by the National Science Foundation CISE Infrastructure grant under Grant No. 0403428 and IGERT Grant No. 0504647.

7. ADDITIONAL AUTHORS

Gang Qian, Wilhelmina Savenye, Hari Sundaram, Harvey Thornburg, Christopher Todd. Arizona State University. [gang.qian, willi.savenye, hari.sundaram, harvey.thornburg, christopher.g.todd] @ asu.edu.

8. REFERENCES

[1] Schunk, D.H., Learning Theories (2nd Ed.). 1996, Englewood Cliffs, NJ: Merrill, Prentice Hall.

[2] Jonassen, D.H., Computers and Mindtools for Schools: Engaging Critical Thinking. 1999, Englewood Cliffs, NJ: Prentice-Hall.

[3] Vygotsky, L.S., Mind in Society: The Development of Higher Psychological Processes. 1978, Cambridge, MA: The Harvard University Press.

[4] Gardner, H., Frames of Mind: The Theory of Multiple Intelligences. 1993, New York: Basic Books.

[5] Simonson, M., et al., Teaching and Learning at a Distance: Foundations of Distance Education (2nd Ed.). 2002, Upper Saddle River, NJ: Merrill, an imprint of Prentice Hall.

[6] Kafai, Y.B., Minds in Play: Computer Game Design as a Context for Children's Learning. 1995, Hillsdale, NJ: Lawrence Erlbaum Associates.

[7] Bransford, J.D., A.L. Brown, and R.R. Cocking, eds. How People Learn: Brain, Mind, Experience, and School. 2000, National Academy Press: Washington, DC.

[8] Land, S.M. and M.J. Hannafin, Student Centered Learning Environments, in Theoretical Foundations of Learning Environments, D.H. Jonassen and S.M. Land, Editors. 2000, Lawrence Erlbaum Associates: Mahwah, NJ.

[9] Bradski, G.R., Computer Vision Face Tracking for Use in a Perceptual User Interface. Intel Technology Journal, 1998. Q2.

[10] Comapiciu, D. and P. Meer. Robust Analysis of Feature Spaces: Color Image Segmentation. in Computer Vision and Pattern Recognition. 1997. San Juan, Puerto Rico.

[11] Open Source Computer Vision Library, http://www.intel.com/technology/computing/opencv.

[12] Svoboda, T., The Multi-Camera Self-Calibration Toolbox. http://cmp.felk.cvut.cz/~svoboda/SelfCal.

[13] Smith, J.O. and X. Serra. PARSHL: An Analysis/Synthesis Program for non-Harmonic Sounds Based on Sinusoidal

Representation. in International Computer Music Conference. 1987.

[14] Thornburg, H. and R.J. Leistikow, A New Probabilistic Spectral Pitch Estimator: Exact and MCMC-approximate Strategies, in Lecture Notes in Computer Science #3310, U.K. Wiil, Editor. 2005, Springer Verlag.

[15] Paine, G., MAP 1: An Interactive Virtual Environment Installation. 1998.

[16] Chadabe, J., Electric Sound. 1997, New York: Prentice Hall.

[17] Schwarz, D. A System for Data-Driven Concatenative Sound Synthesis. in Digital Audio Effects. 2000. Verona, Italy.

[18] Cruz-Neira, C., D. Sandin, and T. DeFanti. Virtual Reality: The Design and Implementation of the CAVE. in ACM SIGGRAPH. 1993.

[19] Cycling'74, Max/MSP/Jittter: Graphical Audio and Video Programming Environment, http://cycling74.com, 2006.

[20] Birchfield, D., W. Savenye, and H. Thornburg, AMEEd Edulink. http://ame4.hc.asu.edu/edulink, 2006.

[21] Schank, R.C. and C. Cleary, Engines for Learning. 1995, Hillsdale, NJ: Lawrence Erlbaum Associates.