

An Immersive, Multi-User, Musical Stage Environment

Matthew Reynolds

Bernd Schoner*

Joey Richards†

Kelly Dobson

Neil Gershenfeld

Physics and Media Group

MIT Media Laboratory

20 Ames Street, Cambridge, MA 02139

{matt,schoner,bigjoe,kdobson,gersh}@media.mit.edu



Figure 1: The Flying Karamazov Brothers. Stage equipped with camera, sonar receivers, and screen.

Abstract

A multi-user, polyphonic sensor stage environment that maps position and gestures of up to four performers to the pitch and articulation of distinct notes is presented. The design seeks to provide multiple players on a stage with the feeling of a traditional acoustic instrument by giving them complete control over the instrument's expressive parameters and a clear causal connection between their actions and the resulting sound. The positions of the performers are determined by a custom ultrasonic tracking system, while hand motions are measured by custom-made gloves containing accelerometer units. Furthermore, juggling clubs are illuminated dynamically to make complex juggling patterns more apparent. The system is currently on tour with the Flying Karamazov Brothers juggling troupe.

*Present address: ThingMagic LLC, Cambridge, MA 02142

†Present address: Trimble Navigation Ltd, Sunnyvale, CA 94088

CR Categories: B.4.2 [Input/output and data communications]: input/output devices—Channels and controllers; J.5 [Arts and Humanities]: Performing Arts

Keywords: Applications, HCI (Human-Computer Interface), Object Tracking, Spatialized Sound, User Interface Design

1 Introduction

The Flying Karamazov Brothers, a group of four internationally acclaimed performers, often seek to communicate and explain high-level concepts and phenomena through their juggling. For their show *L'Universe*, which opened in Seattle in January 2000, they chose to explain (and playfully distort) the history and concepts of physics: planets are represented by juggling balls, large pendula explain Newton's law, and complex juggling patterns (fig. 5) are used as a metaphor for chaos. Technology is integrated with the show to enhance the artistic experience, but also to ease the explanation of scientific concepts and to make performance patterns more transparent to the audience. For example, a large back-projected screen is used to emphasize events on stage, play prerecorded video sequences, create interesting interactions with objects on stage, and provide visual feedback for the performers and audience.

One special portion of the show involves a musical instrument in the form of a stage environment which transforms the motion of the performers into musical notes. The Flying Karamazov Brothers have historically been very interested in music as well as in juggling. For example, they have in the past used the clapping sounds of juggling props to create interesting audible rhythmic patterns,

and they often juggle to the accompaniment of a symphony orchestra. One of their dreams has been to play a giant instrument that uses their entire body as its interface. In our system, the stage position of each performer is mapped to notes of different pitches, while the performer's hand motion is used to trigger, control, and release notes. The necessary position-tracking function is accomplished with sonar receivers located at the outer vertices of a large hexagon (fig. 3 and 6). Performers wear sonar-transmitter hats as well as small sensor units on each hand that measure acceleration in three axes. A wearable computing device worn on the performer's belt digitizes the sensor data, coordinates the sonar pings, and communicates to the base station through a 900MHz wireless link. A system controller coordinates the sonar pings from each of the performers, takes timing measurements from the six sonar receivers, and handles data transfer from the hand sensors. Both the wireless link and the sonar system use a sonar-synchronous Time Division Multiple Access protocol for supporting multiple performers. The system controller communicates through an Ethernet link to software running on a PC. The PC's software maps the raw sonar data onto the XY-positions of the performers and into MIDI notenumbers for reproduction by a sampling synthesizer. Graphics depicting player position and action are projected onto the screen behind the floor to provide feedback for the players as well as the audience. The four performers take part in a quartet and choreographed dance using the stage floor as their instrument and performance space. In a darkened theater, they rely on the visual feedback from the graphics system to locate themselves on the stage while performing music.

In a second application, the spatial information is used to dynamically illuminate juggling clubs depending on the functional position of each juggler in the Karamazovs' complicated and dynamic juggling patterns. A pattern recognition algorithm running on the PC detects each performer's role in the juggling pattern in real-time and assigns a color to each of the jugglers, which is transmitted through the controller's RF link to the performer's wearable computer. Each performer's wrist unit then transmits a low frequency signal modulated with the color data to the juggling clubs. Different colored LEDs inside the clubs are illuminated according to the requested color. The resulting system is a visually impressive dynamic display of the complexity of the juggling pattern, as assembled in real time by the stage computing systems.

2 Related prior work

2.1 Large-scale interactive musical systems

Acoustical instruments, especially in the Western tradition, are causal and predictable—in fact the predominant goal of the process of learning a traditional musical instrument is to have the musical instrument behave in the most predictable way possible. Certain physical actions are assumed to cause a certain musical event to happen. Complexity is created by careful superposition of notes and multiple performers. Mathematically speaking, most traditional musical instruments are predictable, stationary, driven systems.

Many large-scale interactive electro-acoustic or sensor-based digital music systems have a very different character. A specific gesture doesn't result in a specific sound or note, but typically causes a whole sequence of sounds, so that a single performer is able to create an arbitrarily complex sequence of sounds. Rich vertical and horizontal musical structure and material are often triggered by simple gestures or a prerecorded sequence is slightly altered in one or a few ways through user input. The price for this musical complexity is a lack of control of all degrees of freedom of the music. Because of the limitations of either the sensor systems or the musical software it is often impossible to create a single note of a specific pitch, loudness, and duration. This limits the flexibil-

ity of such digital music systems in a stage environment, where it is important to have audience-visible correlation between the performer's actions and the sound output.

The list of examples of large electronic music systems is long. Only a few examples exemplify the desire to produce a fully expressive musical system. Michel Waisvisz's *HANDS* is an early electronic music interface designed to go beyond the expressive possibilities of a keyboard [12]. Waisvisz assembled a variety of sensors on a glove and the performer's body, among them a sonar system to detect the vertical hand movement. Sonar transmitters are mounted on the palms of the performer's hands. Matching sonar receivers on the performer's pants receive these signals and detect Doppler frequency shifts due to hand motion.

In *SOUND=SPACE* at Centre Pompidou in Paris (France, 1991), another such system, Ralf Gehlhaar developed an ultrasonic echolocation [5]. The approach was inspired by Polaroid's ultrasonic ranging device for auto-focusing cameras. Gehlhaar's system tracks over an area of up to 100m. Up to 48 transmitters send sonar impulses, which are reflected by objects in the space at a maximal distance of 11m. Unlike our system, Gehlhaar's approach does not uniquely identify objects and performers.

Brain Opera, an interactive opera by Tod Machover, is composed of many sub-installations which map a few input degrees of freedom into complex pre-sampled or MIDI-generated sounds [11]. Its design philosophy is therefore very different from the authors', because it is composer-centric in the sense that one user gesture can expand into a complex musical output that is preselected and under the composer's, not the performer's, control. For example, hundreds of rhythm pads map user touch to different percussion sounds, where, apart from some impact on loudness, the user has no control over the generated MIDI sequence. The *Future Music Blender*, a chair that reads the user's arm gestures in three dimensions, provides a fine grid of active zones from which the user may select dynamically. It also gives visual feedback indicating the played zone. Fundamentally, however, it allows the user to navigate among preselected streams of music rather than enabling the performer to produce arbitrary sounds.

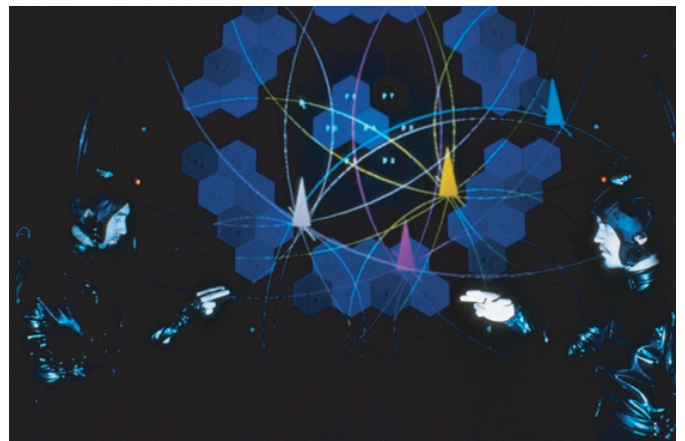


Figure 2: Ivan and Pavel Karamazov performing *Floor Waltz* by Doug Wieselmann. The back projected graphics on the up-stage screen illustrate the position of the performers on the floorboard and the sonar receivers.

At Siggraph '98 David Bianciardi and Ryan Ulyate introduced *Interactive Dance Club*, a large installation using input from the dancing audience to trigger disco-style music [1]. The installation was guided by ten key principles, e.g. under all circumstances the

music had to sound “good” and the effect of the user input had to be clearly perceivable. Indeed, the disco style of the music is preserved convincingly and the dancers do interact with the system in a causal way. However, the degree of interaction and influence available to each user is limited. In this particular system the trade-off between user control and expressivity and sound output favors the musical result over the amount of user control.

Paradiso’s *Dance Shoe* integrates 13 different sensors in a dance shoe [8]. Accelerometers in three axes, gyro sensors, pressure sensors, and a sonar system to track the dancer are collocated on a board which is integrated with the shoe and transmits the sensing information through a wireless link. Unlike the system described in this paper, the sonar chirps are sent from a base station and received by the shoe. The generated sound is a function of the spatial location of the dancer and the other sensor values, but again the mappings are indirect, resulting in a continuous musical floor. Reference [9] describes Paradiso and Sparacino’s earlier efforts of tracking gestures and motion on-stage: a) a multimodal conducting baton using an infrared LED on the tip of the baton whose light is tracked by a camera using a position sensitive detector (PSD), b) a scanning laser range-finder to track hands in front of a screen, and c) a vision system to locate people in front of a screen.

Paradiso provides a comprehensive summary of his and other researchers’ contributions to enabling technology for interactive musical systems in [10]. The overview includes musical interfaces of all kinds, sensor spaces, and novel controllers. Wanderly and Battier edited a collection of articles on the more fundamental aspects of Gestural Control of Music [13], mostly abstracting from the technical challenges.

In this paper we present a digital musical system that tries to stay as close to the tradition of acoustical instruments as possible. The objective is to recreate an ensemble experience, where every note is important for the music and where every note has expressive potential.

2.2 Stage-scale People and Gesture Tracking Systems

During the design phase for this project the authors identified the following major design requirements: a) identification and tracking of four performers simultaneously at approximately 30Hz aggregate rate and 5cm precision b) completely wireless operation with long battery life c) provision of redundant receiver units so that multipath or occlusion problems do not prevent successful tracking d) integration with and synchrony to a gesture data collection system capable of using hand gestures to trigger musical notes, e) modularity so that its components can be easily swapped in or out in case of malfunction and f) “road-ready” and quick to set up and calibrate in a wide variety of different performance venues. Given these design constraints most technologies and commercially available tracking systems were ruled out easily. For example, magnetic field or vision-based techniques could not be used because of the uncertain metallic object distortion and lighting conditions that are routinely encountered when on the road with a traveling Vaudeville show. Sophisticated compensation of venue-to-venue differences in software was ruled out, because such calibration could not be adequately handled by non technical theater personnel. In the end we decided against a commercial system and settled on a custom solution integrating ultrasonic position tracking technology with an accelerometer-based gesture sensing technology.

There is a great deal of prior work relating to the underlying ultrasonic position sensor technology that is used in this system. For example, ultrasonic tracking of users of augmented reality systems is presented in the ConstellationTM system [3], while the use of a person as a passive sonar reflector is described in the *SOUND=SPACE* system presented in [5]. These systems, while

they fail to meet the ease of deployment and unique identification requirements respectively, provide a good introduction to the challenges surrounding the use of ultrasonic devices for position tracking. Another approach, using modulated infrared light, is described by Welch and Bishop in [14] and [4]. This system offers the advantages of very high resolution and high update rate (2000 updates per second with 0.5mm precision are claimed) but is very complex and requires a surveyed infrared emitting ceiling. Our approach, as described below, yields significantly lower resolution and update rates but at low cost and low complexity with few setup or calibration hassles. Prior work in gesture sensing technology is described in [10]; there are many examples of other systems using micro-machined accelerometers as gesture sensors but none (to our knowledge) that are as tightly integrated with the position tracking system.

3 Hardware design

3.1 Sonar tracking system

The Sonar Tracking System is one of the key subsystems of the stage instrument. It is a complex system which is intimately tied to the Time Division Multiple Access (TDMA) communication system described in Section 3.4, as the two systems are synchronous and controlled by the same real-time operating-system software.

3.1.1 Basic principles of the sonar tracking system

The Sonar Tracking System uses the measured time of flight of short, Gaussian-shaped pulses of 40KHz ultrasound, emitted by a transducer array mounted on the performer’s hat, to provide 3D position solutions. This ultrasound frequency was chosen due to the availability of small, inexpensive ceramic transducers for this frequency. The performer’s hat contains eight of these ceramic transducers to cover a full 360 degree radiation pattern. Each ranging pulse consists of a Gaussian-envelope pulse of 1msec duration resulting in 40 cycles of ultrasound per ranging pulse. These pulses are produced at an aggregate repetition rate of 30Hz. The system is limited to 30Hz update rates because the velocity of sound in air at room temperature is approximately 340m/sec. For the hexagonal stage area of approximately 10m major dimension a higher pulse repetition rate would cause range ambiguities, since in each 33.3msec update period, a sonar pulse can only travel about 11.3m.

3.1.2 Sonar receiver units and the RS-485 stage network

Each Gaussian ultrasound pulse transmitted by the performer’s transducer hat is received by six sonar receiver units mounted at the six vertices of the stage hexagon. These six sonar receiver units consist of a tuned homodyne receiver which discriminates the desired ultrasound pulse from the received background noise. Each receiver unit carries two separate receiver chains to reduce the effect of multipath interference on the reception of a sonar pulse—in radio terms this is called a diversity receiver. In the event of multipath sound propagation (echoes) or obstruction of a particular pulse, the receiver reports “no result” and the software Kalman filter ignores that receiver. Each receiver unit is microprocessor controlled and networked to all others via a wired RS-485 stage network. The microprocessor digitally measures the time between a synchronization signal sent over the RS-485 network from the stage’s TDMA System Controller and the center of the received pulse to 10 μ sec precision. This time is then reported back to the TDMA System Controller for data collection and subsequent processing. The theoretical accuracy of such a system would then be 340m/sec divided by

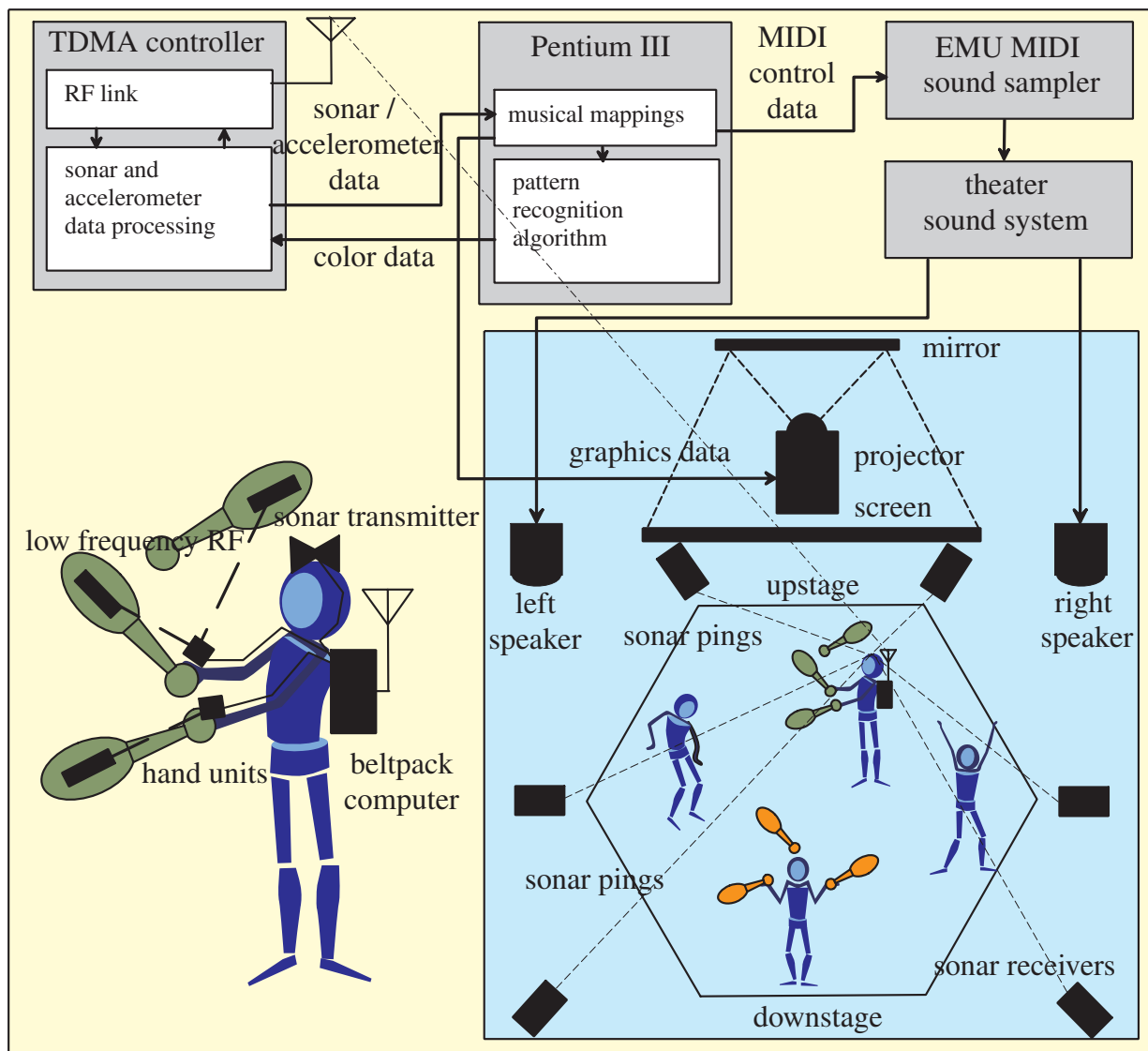


Figure 3: System overview: computing, smart costumes, and stage arrangement.

$10\mu\text{sec}$ yielding 3.4 mm precision. This of course is unachievable—in practice we generally achieve typical measured accuracies of approximately 1 centimeter in X, Y, or Z position, which could be met with $34\mu\text{sec}$ measurement precision.

We compensate for sensor misalignment or sound velocity changes (the velocity of sound in air depends on frequency, air temperature, pressure, and humidity) by overdetermination. By taking six measurements we can determine the velocity of each sound pulse individually and thus compensate for sound velocity changes, clock skews, etc.

3.2 Wearable Beltpack computers

Each performer carries a special purpose wearable computer in a small satchel worn on the belt. These computers, which we call *Beltpacks*, are ruggedized to withstand the rigors of the stage environment. Size, reliability, and power consumption requirements precluded the use of commercial embedded computer modules, so we designed our own. They are enclosed in a reinforced cast alu-

minum boxes the size of a normal builder's brick, approximately 7cm by 5cm by 20cm, and weighting about 1kg each. They are powered by two C-size alkaline batteries and run approximately 10 hours on each set of batteries.

3.2.1 Wearable Beltpack computer internals

Each Beltpack computer consists of a Microchip PIC microcontroller running a tiny real-time operating system of our own design, along with a custom 900MHz data radio unit, made by Micrilor Corp and modified by the authors for fixed latency. This radio communicates with the stage TDMA controller system. Additionally, each Beltpack manages eight channels of 12-bit A/D conversion connected to accelerometers worn on the performer's wrists (the Wrist Accelerometers), and two low-frequency *Short Range Radio* transmitters (so called because they have a transmission range of approximately 1m) operating at 1.28MHz which communicate with juggling clubs when they are handled by the performer. Each Beltpack also provides a connection to the performer's hat, which car-

ries an ultrasonic transducer module for the position-tracking function as described above.

All functions performed by the Beltpack unit are synchronous to time stamps sent over the 900MHz radio channel by the stage TDMA System Controller. The Beltpack listens to these time stamps and performs its measurement duties, responding only in its assigned time slot to the TDMA controller system. In this way the four performers may be tracked and their measurements gathered apparently simultaneously. Measurements of the Wrist Accelerometer states are obtained at a 60Hz rate from each performer.

3.2.2 Wrist accelerometer units

Each performer wears two Wrist Accelerometer Units, one on each wrist. The Wrist Accelerometers are three-axis measurement units based on the Analog Devices ADXL202 micro-machined accelerometer unit. Each ADXL202 contains two micro-machined accelerometers capable of measuring -2g to 2g with 5mg precision. The performer's two Wrist Accelerometers each contain two ADXL202 devices mounted at right angles, along with operational amplifiers for signal filtering and level conversion to 0-4.096V outputs. Additionally each Wrist Accelerometer Unit contains a solenoidal antenna for the 1.28MHz Short Range Radio transmitter which is used to send messages from the Beltpack to the Juggling Club Units.

Because they are worn in a specially made glove on the performer's wrist and are subject to a great deal of mechanical stress, the Wrist Accelerometer printed circuit boards are encapsulated in a very strong urethane rubber potting compound, made by the Smooth-On Corp. and poured in molds made by the authors on a numerically-controlled milling machine.

3.3 Juggling club units

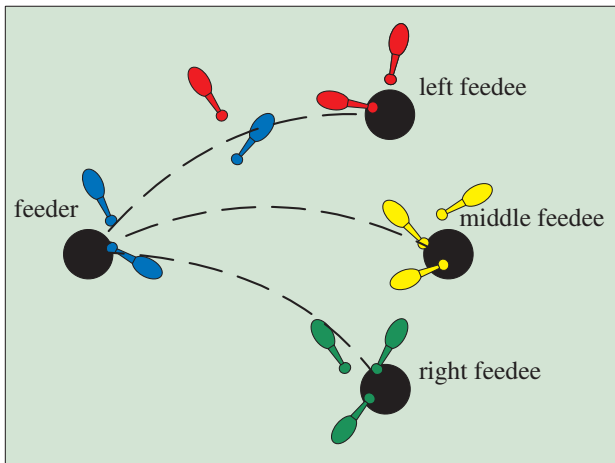


Figure 4: Jazz juggling with color clubs. Feeder (on the left): blue, left feedee: yellow, middle feedee: red, right feedee: green.

Each of the 12 juggling clubs in use during this portion of the L'Universe show contains an internal custom-printed circuit board, called the *Club Board*. This device consists of a receiver for the 1.28MHz Short Range Radio signals emitted by the Beltpacks via antennas on both of the performer's wrists, a microcontroller to manage and decode these messages, sixteen high brightness In-GaAs LEDs, and a switching power supply to supply these circuits with power from two rechargeable NiMH batteries in the neck of the juggling club. The Short Range Radio signal consists of a short

message directing the Club Board to illuminate the club in one of four colors: Red, Green, Yellow, or Blue. As each of the colors is represented by four LEDs and each LED has a brightness of approximately 5600 millicandelas (mCd), the aggregate illumination of each club is approximately 22.4 candelas.

3.4 TDMA system controller

The heart of the Stage Instrument system is the TDMA System Controller. This too is a custom special-purpose computer unit. It is the source of all timing information for all the sonar, juggling club, and Wrist Accelerometer data. It consists of three ports: a 900MHz data radio port which talks to the Beltpack units, an RS-485 Stage Network transceiver, and an Ethernet network port. It additionally contains a status panel indicating the status of each of the four Beltpack units and the six sonar receivers on a set of red or green LEDs for easy problem diagnosis. Sonar ping times are received from the Sonar Receivers via the RS-485 Stage Network, and are fused with data from each of the eight Wrist Accelerometers in real time. This data is then packetized and sent as UDP broadcast packets over an Ethernet network port to the Pentium III Stage PC which runs the music and visualization algorithms. Data flowing back from the Stage PC arrives and is transferred back to the Beltpacks to cause the color of each juggling club to change.

The TDMA controller's interface to the PC is via Ethernet [7]. Instead of serial port hassles, we send IP packets via the host PC's Ethernet network interface. The Microsoft Windows NT operating system's support for TCP/IP networks is far superior to its support for the normal PC serial port causing the software to run much more smoothly with this method.

4 Applications

4.1 Recursive tracking algorithm

The TDMA sonar system delivers up to six timing measurements (six receivers) per performer. The time it takes a sonar chirp to get from the transmitter to the receiver is proportional to the speed of sound and the distance from the performer. This constrains the location of a particular performer to a circle centered in the receiver positions (fig. 6). In a noise-free environment with a perfectly tuned system, the radii of the six circles intersect at exactly one point. Although this point appears geometrically obvious, the problem is nonlinear and has no closed solution. In reality the circles don't intersect properly because of measurement noise, and measurements become temporarily unavailable due to occlusion of a performer's sonar pulses. We use a recursive Kalman-filter-like estimator that deals with all these issues and provides smooth visual and musical dynamics.

First we filter out unrealistic jumps in the distance measurement. Since we are measuring human performers, we assume maximally possible motion and discard measurements that indicate a violation of basic inertial rules. The 'good measurements' are used in an algorithm which can be interpreted as a hybrid of a Kalman filter and a nonlinear steepest-descent search [6].

Let's consider a single performer occupying the stage position (x, y) (fig. 6). We express the six radii r_i as a function of the position,

$$\begin{aligned} r_i &= f(x, y) \\ &= \sqrt{(x - x_i)^2 + (y - y_i)^2} \end{aligned} \quad (1)$$

where x_i and y_i are the fixed coordinates of sonar receiver i . Using a Taylor expansion, we approximate this expression to first order as

$$r_i(x - x', y - y') \approx f|_{(x', y')} + \mathbf{f}'|_{(x', y')} \cdot \begin{bmatrix} x - x' \\ y - y' \end{bmatrix} \quad (2)$$

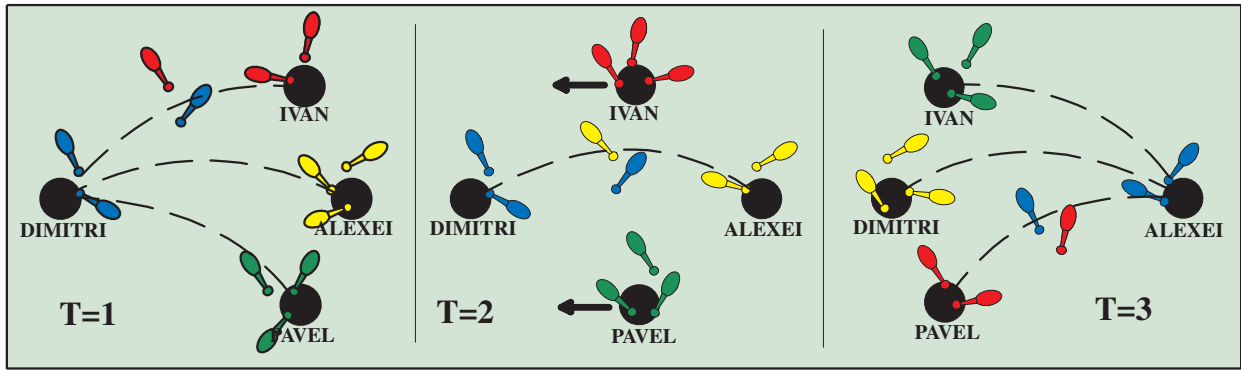


Figure 5: Jazz juggling, *Straightcut* pattern. T=1: the feeder is positioned on the left, normal jazz juggling. T=2: while feeder and middle feedee exchange clubs, left and right feedees move forward turning the former feeder into the middle feedee and the former middle feedee into the feeder. T=3: the feeder is positioned on the right, normal jazz juggling. In the course of the change of positions the clubs change colors according to the new position the performers are occupying.

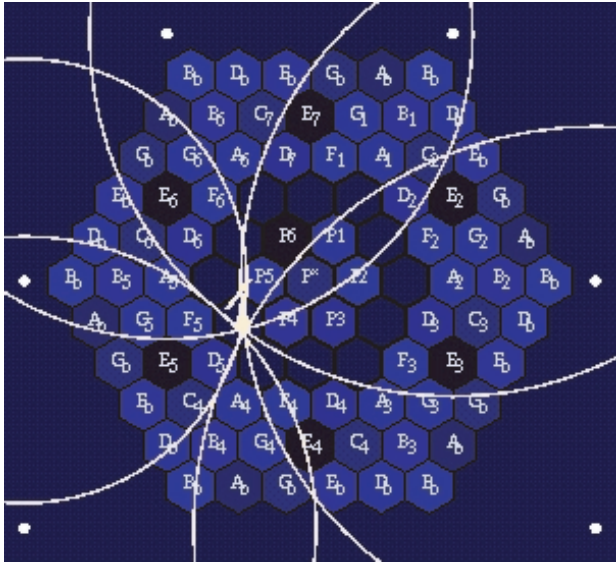


Figure 6: Hexagonal floor with receiver positions (white dots), one performer, measurement circles (white circles), and hexagonal keys.

where x' and y' are arbitrary anchor points, $[\cdot]$ indicates the inner vector product, and \mathbf{f} is defined as $\mathbf{f}' = \begin{bmatrix} \frac{\partial x}{\partial r} \\ \frac{\partial y}{\partial r} \end{bmatrix}$. Introducing the time step δ and using matrix notation we obtain,

$$\begin{aligned} \mathbf{r}(x - x', y - y') &= \mathbf{r}(x', y') + \delta \mathbf{r} \\ &= \mathbf{f}(x', y') + \partial \mathbf{F} \cdot \begin{bmatrix} x - x' \\ y - y' \end{bmatrix} \end{aligned} \quad (3)$$

with

$$\partial \mathbf{F} = \begin{bmatrix} \frac{\partial r_1}{\partial x} & \frac{\partial r_1}{\partial y} \\ \frac{\partial r_2}{\partial x} & \frac{\partial r_2}{\partial y} \\ \dots & \dots \\ \frac{\partial r_6}{\partial x} & \frac{\partial r_6}{\partial y} \end{bmatrix}. \quad (4)$$

In order to update (x, y) we invert this last expression and get

$$\begin{bmatrix} \delta x \\ \delta y \end{bmatrix} = \epsilon \cdot \partial \mathbf{F}^{-1} \cdot \delta \mathbf{r} \quad (5)$$

where ϵ is the step size of the update. The non-square matrix $\partial \mathbf{F}$ is inverted using the pseudo-inverse (SVD). Although new sensor readings become available at only 6-8Hz, we iterate on each reading orders of magnitude more often, and hence can afford a small step size ($\epsilon < 0.05$). In this application we consider the two-dimensional case only, yet the algorithms naturally extend into three dimensions.

The recursive approach can be extended to include the discovery of the receiver locations. Fixing receivers one and two (necessary to guarantee a unique solution), we include the remaining receiver positions as unknowns and obtain

$$\begin{bmatrix} x \\ y \\ x_3 \\ y_3 \\ \dots \\ x_6 \\ y_6 \end{bmatrix} = \epsilon \cdot \partial \mathbf{G}^{-1} \cdot \delta \mathbf{r} \quad (6)$$

with

$$\partial \mathbf{G} = \begin{bmatrix} \frac{\partial r_1}{\partial x} & \frac{\partial r_1}{\partial y} & \frac{\partial r_1}{\partial x_3} & \dots & \frac{\partial r_1}{\partial x_6} \\ \frac{\partial r_2}{\partial x} & \frac{\partial r_2}{\partial y} & \frac{\partial r_2}{\partial x_3} & \dots & \frac{\partial r_2}{\partial x_6} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial r_6}{\partial x} & \frac{\partial r_6}{\partial y} & \frac{\partial r_6}{\partial x_3} & \dots & \frac{\partial r_6}{\partial x_6} \end{bmatrix}. \quad (7)$$

4.2 Hand accelerometer units

The hand accelerometer units measure the acceleration of each hand of each performer in three axes at a resolution of 300 steps per g (acceleration due to gravity). The filtered update rate per sensor and performer on the stage base station PC is 60Hz.

Information from the accelerometer units is used in two distinct ways. Signals from the left hand sensors are low-pass filtered to extract the DC response of the sensors relative to gravity. Given the baseline of the accelerometers in absence of any force and given sufficiently slow movements of the left hand, the DC response from the sensors is mapped linearly to the angle of the left hand and arm relative to the body. The angle of the left arm is then used to

determine the initial volume of a triggered note and for after-touch control of a note.

The performers use their right-hand accelerometer units to trigger a note. Abrupt acceleration and de-acceleration of the hand equals three to eight times the acceleration due to gravity, which is sufficient to distinguish rapid action from the DC response. Since acceleration is measured in three axes, different directions of movement are mapped into different kinds of sounds. In particular, beating the forearm is mapped to sustained notes, while movement along the axis of the arm is mapped to staccato notes.

4.3 Musical mappings

The stage (shaped as a hexagon) is divided into 72 hexagonal *pitch* or note keys, arranged in a ring, and 7 hexagonal *patch* keys in the center rosette (fig. 6).¹ Triggering a key on the pitch keys causes a note onset, while triggering a patch key changes the sound patch of the particular performer. The patch keys can be compared to the stops on church organs used to select different sets of pipes. The pitch keys run chromatically from *F1* to *E7*. The performers are completely independent from each other.

In order for a performer to get oriented, the virtual hexagonal keys are indicated with black light reflecting paint on the stage. The performers select a key by moving their heads over the key. They trigger a note by a sharp movement of the right hand. At the same time they use slow movements of the left hand to select the initial volume of the sound, to shape the after-touch volume, and to release a note (sec. 4.2).

The position information of the performers on stage is used to create stereophonic sound. The individual sound is mapped on left and right speakers of the theater sound systems according to the left-right position of the individual player. Hence, the spectator's experience is similar to that of the audience of a symphony orchestra, where, for example, the violin section is projecting from the left side of the stage and celli and double basses are projecting from the right side.

The musical mappings are implemented with the Rogus-McBogus MIDI library, developed for interactive musical spaces and applications on the Windows operating system [2].

4.4 Dynamically colored juggling clubs

The Flying Karamazov Brothers are masters of club passing. The four performers simultaneously juggle a total of 12 clubs, constantly exchanging them among each other (fig. 4). Specifically, one performer is acting as the *feeder*. The feeder passes every right-hand club throw to one of the *feedees*, i.e. one of the three other performers, in the order, left-feedee, middle-feedee, right-feedee, middle-feedee, left-feedee and so forth (fig. 4). Hence the middle-feedee passes every second right-hand club throw to the feeder, while the two side-feedees pass every third right-hand club throw to the feeder. In between all the performers keep their own clubs moving.

The pattern is interrupted by variations, resulting in position changes of the performers. Fig. 5 illustrates one such pattern, called *straight-cut*. The two side-feedees move forward turning the middle-feedee into the feeder and the feeder into the middle-feedee. Other popular patterns are the *outside end-run*², or the *crosscut*³.

In the course of juggling these complicated patterns, the audience gets easily "lost," since it is practically impossible to track

¹The seven patches were custom designed by Doug Wieselmann and are played on an EMU ESI-4000 sampler.

²A side-feedee moves behind the other two feedees thus taking the position of the other side-feedee.

³A straightcut occurs when the side-feedees cross through the center.

each club and to understand who is fulfilling which role in the club-passing pattern. To assist the audience in its understanding of the game, we instrumented the clubs in such a way that each club can be illuminated in the color of the logical position of the performer that is juggling it. We assigned a specific color to a position, for example blue to the feeder and yellow to the middle feedee. The clubs juggled by the feeder are illuminated blue independently of who is juggling them, whereas the clubs juggled by the middle feedee are yellow.

In the previous section, a detailed description was given on how to find the (x, y) position of any of the four performers on stage. We now use this information to assign positions to each performer at any moment in time by means of a pattern-recognition algorithm. Given four (x, y) pairs we first determine the center of mass of the performers. The performer occupying the position of the middle-feedee is easily found, since the middle-feedee is always closest in distance to the center of mass. We then renormalize the position data with respect to the center of mass and transform the Cartesian coordinates into polar coordinates. Given the reference point 'middle-feedee' and the fixed order of positions, the angles relative to the middle-feedee are sufficient to assign the three remaining positions.

The desktop PC sends the color information back to the performers' Beltpacks via the 900MHz wireless link. The Beltpacks transmit color data through the Short Range Radio from the hand units to the clubs. The range of this second RF link is limited to a radius of about one meter, so that the signals remain local to the clubs of a particular performer. Given the color information the Club Board's onboard microprocessor turns on the corresponding LEDs. The audience sees each club in the color of the particular position and hence easily understands the 'meaning' of the position.

5 Artistic Evaluation

The system accurately tracks all four performers to centimeter resolution, the hand accelerometer units function as explained, and the communication between hand and juggling clubs works reliably. The musical application as well as the dynamically-colored clubs are used in a two year international tour that is now halfway complete. The stress on the system is enormous, but so is the success with audiences around the world. The performance leaves a strong impression on the audience due to the surprising technology, a nicely integrated set, fancy costumes and, most importantly, the charisma and professionalism of the Karamazovs (fig. 2).

The system uses a head-mounted ultrasonic sonar tracking system to locate the performers and to select *keys*. From a performance perspective, this is unusual since the players seem intuitively to want to use their feet to step on keys and select notes through contact. However, heads can be moved much faster and are less hindered by other physical constraints, which opens up several interesting possibilities regarding the maximal speed of note sequences and choreographic effects. For example, a performer who is arriving late on a particular note can "stick his neck out" and grab the appropriate note. A foot based tracking system could be implemented with the exact same hardware by mounting the sonar transmitters on the performer's feet, but would suffer from ground reflections and additional occlusion from elements of the set and other performers.

Unlike electro-acoustic and digital instruments that use indirect mappings of sounds, the stage instrument is merciless when it comes to wrong notes or notes that are not triggered in a timely fashion. As in an acoustic instrument such as the piano or in percussion instruments, correct timing and coordination among the members of the ensemble are crucial for the quality of the musical output. Because the instrument is so demanding, performers have to practice to learn how to use it, much as they would practice with a traditional acoustical instrument. Since the opening of the show,

the musical results have improved dramatically, mostly due to the experience of the performers, who have proven to be remarkably adaptable to a heretofore nonexistent musical instrument.

From a composer's perspective, writing music for the stage instrument is very much like writing for a classical ensemble because of the natural constraints of the instrument. For example, it takes a certain amount of time for a performer to move from the lower register to the higher register, and it takes a different amount of time to move from one note to the next within a particular register. In analogy, a pianist is unable to play the lowest, middle and highest register at the same time because he or she has only two hands. The sequence of pieces written for the stage instrument reflects the learning process on the side of the composer. The piece currently performed as part of the L'Universe show, *Floor Waltz* by Doug Wieselmann, makes much better use of the instrument than earlier works, both musically and choreographically.

6 Conclusions

The natural time constant of the sonar tracking system, limited by the speed of sound in air and providing an update rate of 6 Hz per performer, is at the lower limit of what is desirable for musical applications. The performers have to take into account the response time of the system when moving about on the stage. In that sense update speed is not a fatal flaw of the system, because practice and careful attention to musical composition can compensate for this limitation, but faster technology would make playing easier.

Many applications other than the one described here would benefit from the ultrasonic sensing system, which is low-cost, flexible, precise to sub-centimeter resolution, and increases nicely in accuracy with the number of receivers. While three receivers provide enough information to guarantee a unique solution in a noise-free environment, each additional receiver increases the robustness of the measurement, dealing with the case of occlusion of any given sonar chirps. This is due to the use of relatively sophisticated Kalman-filtering and nonlinear interpolation algorithms.

In general, electronic devices as part of a touring performance system are difficult to design and to manage. Systems have to meet unpredictable constraints, such as RF noise in venue halls, stage particularities, limited financial budgets, artists with very strong feelings regarding functionality and aesthetic appearance of the devices, and stage union rules which don't allow for last minute fixes. None of these difficulties, however, compares to that of building devices for jugglers who will take very complex electronic equipment and literally throw it around the stage. Of course the performers are supposed to smoothly catch the flying objects, but in reality many throws miss their marks and accidentally hit the floor instead. The Karamazovs refer to these undesired but frequent events as *non-conforming patterns*. Should we be surprised that the stage instrument and the juggling clubs are still working after a year of touring and daily performances? We will certainly be investigating the availability of military specified components for our next performance system...

7 Acknowledgements

The authors would like to thank Paul Magid, Howard Patterson, Mark Ettinger, Roderick Kimball, Doug Wieselmann, Doug Elkins, and Kristina Wicke from the Flying Karamazov Brothers, as well as the staff of ACT Theatre in Seattle for their juggling inspiration and support. Susan Botari of MIT provided very important logistical support during this project. This work was made possible by the MIT Media Lab's Things That Think research consortium, NEC Corporation, and Creative Technologies Corporation.

References

- [1] David Bianciardi and Ryan Ulyate. Interactive Dance Club, 1998. ACM SIGGRAPH.
- [2] Ben Denckla and Patrick Pelletier. The technical documentation for *Rogus McBogus*, a MIDI library, 1996. Available at <http://www.media.mit.edu/hyperins/rogus/home.html>.
- [3] M. Harrington E. Foxlin and G. Pfeifer. Constellation(tm): A wide-range wireless motion-tracking system for augmented reality and virtual set applications. In M. F. Cohen, editor, *SIGGRAPH, Annual Conference on Computer Graphics & Interactive Techniques*, 1998.
- [4] L. Vicci S. Brumback K. Keller G. Welch, G. Bishop and D. n. Colucci. The hiball tracker: High-performance wide-area tracking for virtual and augmented environments. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, pages 1–11, London, 1999. ACM SIGGRAPH, Addison-Wesley.
- [5] Ralf Gelhaar. Sound=space: an interactive musical environment. *Contemporary Music Review*, 6(1):59–72, 1991.
- [6] Neil Gershenfeld. *The Nature of Mathematical Modeling*. Cambridge University Press, New York, 1999.
- [7] O. Omojola, R. Post, M. Hancher, Y. Maguire, R. Pappu, B. Schoner, P. Russo, R. Fletcher, and N. Gershenfeld. An installation of interactive furniture. *IBM Systems Journal*, pages 861–879, 2000.
- [8] J. Paradiso, K. Hsiao, and A. Benbassat. Interfacing the foot: Apparatus and applications. In *Proceedings of the ACM CHI Conference. Extended Abstracts*, pages 175–176. 2000.
- [9] J. Paradiso and F. Sparacino. Optical tracking for music and dance performance. In A. Gruen and H. Kahmen, editors, *Optical 3-D measurement techniques IV*, pages 11–18. Herbert Wichmann Verlag, Heidelberg, 1997.
- [10] Joseph A. Paradiso. Electronic music interfaces: New ways to play. *IEEE Spectrum Magazine*, 34(12):18–30, 1997.
- [11] Joseph A. Paradiso. The brain opera technology: New instruments and gestural sensors for musical interaction and performance. *Journal of New Music Research*, 28(2):130–149, 1999.
- [12] Michel Waisvisz. The hands. In *Proceedings International Computer Music Conference*, pages 313–318, 1985.
- [13] M. Wanderly and M. Battier, editors. *Trends in Gestural Control of Music*. Editions IRCAM, Paris, 2000.
- [14] G. Welch and G. Bishop. Scaat: Incremental tracking with incomplete information. In T. Whitted, editor, *Annual Conference on Computer Graphics & Interactive Techniques*, pages 333–344. SIGGRAPH 97, ACM Press, Addison-Wesley, 1997.