Mark Feldmeier and Joseph A. Paradiso
Responsive Environments Group
MIT Media Laboratory
20 Ames Street, E15-327
Cambridge, Massachusetts 02139 USA
carboxyl@mit.edu
joep@media.mit.edu

An Interactive Music Environment for Large Groups with Giveaway Wireless Motion Sensors

Few existing systems effectively enable a large number of participants to collaboratively control a real-time, centralized interaction. This is particularly true in the area of interactive dance. Wearable dance interfaces (e.g., Siegel and Jacobsen 1998; Paradiso et al. 2000; Aylward and Paradiso 2006) allow a single dancer or small group of dancers to control music with their actions, but these do not scale to allow for hundreds of participants to interact concurrently. Various vision-based tracking systems are able to extract considerable nuance from dance ensembles (e.g., Wechsler, Weiss, and Dowling 2004; Obermaier 2004; Camurri et al. 1999), but they generally exploit a highly structured and stable stage environment with tight lighting constraints. The problems of cost, data-communication bandwidth, and system responsiveness become increasingly difficult as the number of participants increases. A system that could effectively give control to a large number of dancers offers the possibility of environments with extremely responsive music and lighting, engaging users with a heightened sense of expressiveness.

To address these issues of large-group musical mapping, we have developed a scalable system first introduced in Feldmeier (2002); Feldmeier, Malinowski, and Paradiso (2002); and Feldmeier and Paradiso (2004) that can effectively gather data over an essentially unlimited audience size. The system consists of wireless sensors that are given to audience members to collect rhythm and activity information from the crowd that can be used to dynamically determine musical structure, sonic events, and/or lighting control. (A block diagram of the system is shown in Figure 1.) The sensors are small and lightweight, and they can therefore be either worn or held by a participant. To detect the participant's motion, they have radio-frequency (RF) transmitters that send a short pulse of RF energy whenever they encounter a dynamic acceleration greater than a predetermined level. Finally, they are inexpensive enough to be viable as disposable, "giveaway" items for large crowds.

The sensors' RF pulses are collected by receiver base stations that have limited sensitivity, enabling the development of zones of interaction around each one. In this manner, multiple base stations can be used in a venue to create distinct areas where the controller takes on new functions. This zoning information can also be used to direct the music and lighting to respond to the participants' actions in that area, localizing the response to a smaller group of proximate dancers. The pulses received by the base stations are then sent to the MIDI converter, which counts the number of pulses detected in each zone and transmits this information at regular intervals via MIDI serial communication. These MIDI signals are then received by a Macintosh G4 computer, where they are analyzed to detect activity levels and rhythmic features of the audience.

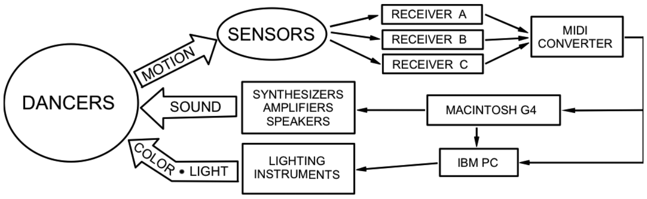
These parameters are then available to be mapped to musical content and/or lighting control information. For our applications, all data analysis and musical mapping is done in the Max programming environment, and sound generation is performed off-board with dedicated hardware music synthesizers. Lighting content is generated with an IBM-compatible personal computer, and control information is sent to the lighting instruments via DMX serial communication (Randall 2002). The sound and lighting changes are then realized, to which the audience in turn responds, allowing the experience to build upon itself and giving the users an increased connection to the music.

Background

A summary of several projects exploring electronic musical interfaces that facilitate group interaction

Computer Music Journal, 31:1, pp. 50–67, Spring 2007

Figure 1. High-level block diagram of the deployed system.



can be found in Blaine and Fels (2003) and Weinberg (2005). Past projects such as the Brain Opera (Paradiso 1999), the SIGGRAPH 98 Interactive Dance Club (Ulyate and Bianciardi 2002), and the ADA installation at Expo 2003 in Neuchatel (Eng 2004) have explored the use of rooms with a vast sensor and interface infrastructure for facilitating large-group entertainment. Although a goal of research into Ubiquitous Computing (Weiser 1991) is to make such infrastructures economical and commonplace, these installations still tend to be extremely expensive, with their deployment requiring up to millions of dollars.

Current developments in the area of large-group interaction are dominated by interactive gaming. Networked computer games (Helin 2003; Alexander 2005) can allow hundreds of physically distributed participants to interact or compete with each other, with data delays on the order of 100 msec or longer. For physically situated groups of about a hundred participants, there exist fixed systems such as voting interfaces for game-show audiences, usually pushbuttons located in the armrests of chairs (e.g., the StarRider Digital Dome Theater System made by Evans & Sutherland, used in planetariums). But, for participants in numbers over a hundred, hardwired solutions become costly and do not allow the participants to be mobile. Some situated gaming systems enable several contestants to become engaged via wireless PDAs (Falk et al. 2001), but these are also quite costly and generally do not operate in real time. Lower-end interactive systems have also been developed that allow groups of several hundred users to interactively vote via an RF (Laibowitz et al. 2006) or infrared (Cutts et al. 2004) link. Although these are less expensive than PDAs, they are still prohibitively expensive for our application, typically costing US$ 10 or more each, and they generally lack the speed and interface features that are needed for controlling interactive music. Likewise, group gaming and interaction (often for remote players) is becoming a commonplace application for cell phones (e.g., Paavilainen 2003; Akkawi et al. 2004), but again, their response time and control abilities limit their penetration in real-time interactive music—although the speakers of many cell phones carried by attendees in a large audience have been used to realize performances with distributed audio output (Levin 2001).

Systems that look for crowd cues from thermal infrared cameras (Ulyate and Bianciardi 2002), microphones (Freeman 1986), or capacitive sensors (Baxter 1996) can gather bulk information over a large, mobile audience, but they do not lend themselves to direct control by an audience member. Generally, the participant has no sense of which action will dictate the desired response. For this to happen, there must be an effective way of measuring a particular action taken by each participant. To date, the most effective methods of measuring and adding audience members' actions have been done via machine vision.

Loren Carpenter's Cinematrix Interactive Entertainment System (Carpenter 1993, 1994) enables audiences of thousands of people to compete concurrently in electronic games. The system functions by giving a retro-reflective paddle to each participant. One side of the paddle is red, and the other is green. By rotating the paddle, audience members can signal to a camera which direction they want an agent to move, or cast a vote between two outcomes. In this manner, the sums of red and green pixels in a given area determine the direction or outcome. Each audience member is given a direct and causal input, and, in turn, feedback in the form of a large display indicating the group interaction is provided to inform audience members that their responses have been received. The nature of the input is very intuitive, and so the paddle needs little explanation for its use. Finally, the reflectivity of the paddles aids the selectivity of the machine vision, allowing the system to be used under a variety of lighting conditions.

Another scalable method of tallying an audience's actions is to give each participant a very simple active device. This is demonstrated by Rosalind Picard's and Jocelyn Scheirer's glowing Galvactivator skin-resistance detectors (Picard and Scheirer 2001), which have been used on an audience of 1,200 people. The Galvactivator is a glove worn by participants that has two electrical contacts that measure the change in conductivity of their skin, ideally corresponding to their level of affective arousal. This change in conductivity is displayed as brightness of a light emitting diode (LED) on the glove. The aggregate brightness of the audience's Galvactivators can then be viewed with a video camera, allowing the audience's excitement or stress level to be measured. The glove itself consists of a minimal component set, making it an inexpensive item, capable of being given away to participants. Direct feedback from each participant can be obtained given a line of sight, and—except for calibration of the LED brightness—the user need not put any effort into the device or the interaction.

Dan Maynes-Aminzade's camera-based audience interaction work (Maynes-Aminzade et al. 2002) is unique in its ability to sense the audience's actions without tethering its members to any input device. A machine-vision algorithm merely looks for global characteristics of those motions made naturally by audience members, making the system inexpensive and intuitive to use. For example, to control a driving simulation, the audience leans in the direction it wishes the car to turn. Or, to control the motion of an agent, the system uses pixel differencing to detect the aggregate crowd motion, increasing the agent's motion accordingly.

As indicated by previous work, machine vision is an extremely low-cost method of gathering data over a large group of people. It has the advantage of not tethering the audience while being able to detect direct inputs by audience members. Despite these benefits, machine vision suffers from requiring a line of sight from the camera to all participants, and it is susceptible to illumination effects and background lighting. Also, many of the required input methods, such as voting paddles or glowing gloves, can restrict the motions of the dancer to those that placed the interface in view of the camera. For work such as Maynes-Aminzade's, which does not include a contact-sensing method, the user does not have a direct input to the system and thus represents merely a portion of the net activity.

In general, making these types of measurements using non-contact methods, such as machine vision or machine listening, is not as specific as direct methods such as wearable or hand-held sensors with RF links, which do not suffer from occlusion. One example of such a wearable device was used in the Sophisticated Soiree installation (Berger et al. 2001) at Ars Electronica 2001, where up to 64 participants were given wireless heart-rate sensors that controlled a musical stream for an experiment in large-group biofeedback. Systems that measure autonomic responses such as heart rate and skin resistance (as in the Galvactivator), however, are generally not consciously controllable by participants. Dave Cliff's proposed extension (Graham-Rowe 2001) to his HPDJ project (Cliff 2000) is an interesting hybrid that suggests providing dancers with wireless accelerometers, heart-rate monitors, and perspiration sensors. This relatively expensive package would then be used to gauge the general activity level of a club's crowd and to choose appropriate musical tracks via a genetic algorithm to keep the crowd dancing.

Sensor Design

The chosen design, as detailed in Figure 2 and shown in Figure 3, is a small, low-cost, wireless transmitter that sends a short pulse of RF energy whenever it senses a changing acceleration greater than a predetermined level. These transmitters can be either worn or held by a participant and are activated by motion. To minimize cost, power consumption, and use of bandwidth (therefore minimizing the probability of collisions among signals), the simple pulse transmissions are not coded. As a result, the system is potentially sensitive to outside interference. However, discriminating via pulse width achieves some degree of background rejection, and in tests of our system to date, we have experienced no significant interference problems. The short transmission radius (roughly 10 m) also creates an approximately bounded zone of interaction around the receiver's

Figure 2. Schematic of the basic wireless inertial sensor unit.

Figure 3. The wireless sensor (front, back, and tubed).

Figure 4. A large collection of sensors waiting to be distributed at events.

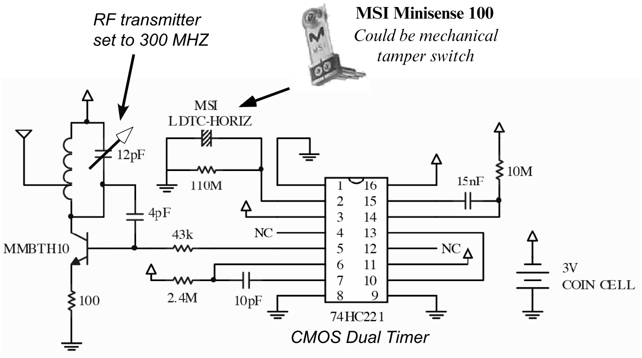


Figure 3

Figure 4

base station. In this way, each participant's action is received instantaneously as a distinct event, and the pulses in a particular area can be added across different time horizons to give a sense of the local rhythm and activity.

In its current form (see Figure 2), the sensor consists of a trigger, debouncing circuitry, and an RF transmitter. The trigger is a commercial piezoelectric polyvinylidene flouride (PVDF) film cantilever, weighted with a proof mass. Whenever the controller is accelerated past an approximately 2.5-G threshold, the PVDF generates enough voltage to trigger a dual CMOS timer. The first half of this timer produces a 150-msec pulse to eliminate double triggering due to PVDF film ring down, and the second half of the timer produces a 50-µsec pulse that activates the RF transmitter. The 300-MHz LC oscillator transmitter provides about a 10-m effective transmission radius, depending on the RF environment.

The power for the controller comes from a single 3V lithium coin cell. The circuit consumes less than 0.01 µA in standby and a few milliamps during transmission. The circuit is accordingly very long-lived. At the rate of two transmissions per second, the battery would last for a month of continuous use, 3–4 years of normal use (assuming one event/week), and up to its shelf life (10 years) with no use. Sensors of this sort that were constructed four years ago for laboratory demonstrations are still working flawlessly.

Five hundred of these sensors (see Figure 4) have been assembled, tested, and placed in protective tubes. They measure 6.2 cm × 1.6 cm × 1 cm and weigh 5 grams. These dimensions could be significantly reduced if the sensor were to be redesigned for mass manufacture. Chip-on-board technology, specially designed PVDF elements—or replacing the PVDF tab with a compact inertial switch, such as those used in toy electronic rhythm stick controllers like the Casio SS1 Soundsticks (Kashio and Yoneiki 1991) or the RockBeat RhythmStick—and a smaller battery could easily reduce the sensor to below the size of a watch. For quantities of one hundred, the price is currently US$ 10 per unit, with assembly accounting for half of this cost. For much larger quantities, more economical manufacturing and procurement procedures should bring the total cost down to under US$ 1 per sensor, making it viable as a giveaway item.

The receiver base stations consist of a commercial Ming RX-99 V3.0A receiver module that plugs into a cable driver board. To reduce the possibility of RF interference and to zone the interactions to a limited radius, the Ming receivers were modified to reduce their sensitivity. The receiver base stations send their pulses to the MIDI converter, where they are integrated at regular intervals for each zone. Although a large-scale installation would require networking many base station receivers, only three receivers can connect to the current MIDI converter's input stage. This was adequate for our implementations, which so far have occurred in fairly small spaces. The MIDI converter eliminates the possibility of double-counting a single pulse seen at multiple base stations and routes each pulse to counters dedicated to signals appearing in each zone separately or appearing together in different zone combinations. The counters accumulate the number of pulses received within a short and periodic window. By adjusting the position of a jumper, this window can be varied anywhere from 2–64 msec, in multiples of 2 msec. (A 2-msec integration time was used in the tests presented here.) Pulse counts are streamed via MIDI messages at the selected update rate. More details on the hardware implementation can be found in Feldmeier (2002).

All sensor data analysis and musical control software was written in Max and runs on a Macintosh G4 laptop, and audio was generated by a rack of commercial MIDI synthesizers. The raw data stream from the MIDI converter was logged on a separate computer. A MIDI keyboard was available for tweaking control parameters during an interactive event, such as changing the number of people that the system assumes to be participating to keep the sensor activity levels properly normalized.

Interaction Design

This section describes the musical mappings and the concepts behind these musical mappings that were used for the initial dance events during which data were collected. These mappings are focused toward generating electronic dance music ("techno"), as the lead author was familiar with making music of this style and was an experienced DJ, and hence had ready access to a large group of participants. Electronic dance music also follows a highly structured set of rules, lending itself more readily to algorithmic interpretation, although potentially making the particular mappings developed here style-specific and not readily transferable to other forms of music. (The high-level features that are derived from the sensor signals, however, could be more universally applicable.) Albeit, this does not preclude another musical or entertainment genre from being accommodated through a different mapping from gesture to music, which could meet the needs of a completely different audience.

Goals

The main objective of the system, being primarily for entertainment, is to create an engaging and enjoyable musical experience. To accomplish this, the system must be easy and intuitive to use, providing appropriate feedback to participants so that their actions will naturally follow the expected behavior assumed by the mappings. The system should also be causal, giving users knowledge of what outcomes specific actions will create. This responsiveness will allow the users to dictate the experience's direction, giving them a tool for sonic exploration and encouraging them to continue using the system.

Heuristics

Although dance is perhaps among humanity's most ancient social traditions (Wallin, Merker, and Brown 2001), there seems to be little current research in the area of large group behavior while dancing—how dancers respond to music and what parameters are most effective or causal. However, there is no reason to believe that group dance is too disparate from other human "flocking" activities and that human responses do not bear certain fundamental characteristics. A relevant study (Neda et al. 2000) on crowds of people interacting involves the human clapping response. As anyone who has been to a public performance will note, the clapping of audience members usually starts chaotic and dispersed, and then it often synchronizes very quickly so that the majority of members are clapping in unison. Typically, this combined clapping tempo will subsequently increase until it finally breaks back into chaos. This pattern of combining, building, and breaking down tends to repeat for as long as the clapping continues.

This implies a number of important things about the properties of human group interaction. First, humans tend to naturally school if given appropriate feedback regarding their current state. There is a herd instinct that encourages the individual to respond in accordance with the majority and to create order from chaos—an emergent property seen throughout many social fauna (Pikovsky, Rosenblum, and Kurths 2001). Second, humans often act in the manner of a positive feedback network. They drive each other to increase activity, each time responding to their neighbors' increases with an increase their own, building to a level that is no longer stable or pleasing. Finally, it is at this point that order breaks, and humans naturally return to a ground state, as there is no longer suitable feedback to denote what the majority is doing. It is these three assumptions that influence the philosophy of our interaction design.

For users to become aware of their actions and to direct these actions into a coherent outcome given a low participant-activity state, users must be given a direct response for each and every action. In this manner, they will realize what outcome their actions have, and they will be able to control not only the occurrence but also the timing of that outcome. They will then be able to make the decision whether or not to produce that outcome in time with the perceived average timing of all outcomes.

A limitation to this scheme, especially with this system, is that for it to work, the complexity of its responses must be limited so that an average timing can be detected. With too many responses coming from one person, as could be the case for large anonymous crowds, the sum of all outcomes could easily stay in a cacophonic state. The feedback mechanism employed in the clapping response does not intuitively translate to this system for groups of people greater than approximately twenty. With clapping, each response is created locally and is heard locally. People clapping mainly hear their own individual clapping and that of any neighbors, and they can therefore quickly synchronize with those nearby. With this system, all responses are integrated globally and in turn affected globally. The sum of all responses, as a result, can easily become chaotic, especially without visual channels (e.g., watching one's neighbors).

To eliminate this problem, a secondary form of feedback can be given. The system can detect the average timing of all outcomes, and it can give not only a direct response but also a larger response that is timed to the computationally perceived average outcome. In this manner, participants can quickly sense the average outcome and choose to either coincide or deviate.

Once the participants have become synchronized around a central outcome, the perceptual energy level of the music should increase with increased activity level of the participants. Through changing dancing patterns, the participants can then dictate whether to increase or decrease the perceived energy level of the music. This is a very intuitive interface for the dancers, as they will naturally increase their actions if the music becomes more energetic.

Because the energy level of the music cannot increase indefinitely, it will reach a maximum at some point during the interaction. When this occurs, the participants must be given some method of transitioning to a lower state. This can occur in a number of ways, but it should be done in a fashion that does not merely return the participants to the previous energy state. If this were to happen, it would have the equivalent effect of negative feedback through processing and perceptual delay, so the crowd would merely oscillate between the two highest states. Instead, we deem it preferable for this highest-energy state to grow chaotic, intuitively denoting to the users that they can go no further. This will have the combined effect of disrupting group behavior in a natural way, ensuring that the energy level decreases and setting the users' expectations of what will occur next. They will then be able to build a new structure from the ground up, adding variety and a sense of exploration to the experience.

Implementation

The main framework for our musical mappings is fairly rigid. The music is scripted into five energy levels, and within each level, it is broken into five simultaneously playing tracks: drones, melodies, high-frequency percussion (e.g., high hats, cymbals, and wood blocks), low-frequency percussion (e.g., bass drums, kick drums, and toms), and arpeggiated lines. For each track, there are 0–3 patterns, only one of which is played at a time. In some cases, a silent pattern that increases the period between repetitions of all other patterns in that track is also inserted. In this manner, the different combinations of patterns that are played each time a level is attained keep the music dynamic, creating a more complex experience.

These transitions between energy levels can occur in a number of ways. The new pattern can either fade in over 16 measures, or it can cut in on the first beat of the next measure. In the same way, the old pattern can either fade out over 16 measures or cut out on the first beat of the next measure. These transition modes are distributed somewhat randomly among patterns, creating an indeterminate state when crossing energy levels, where the participants can either move the energy level up or back down before the full transition is complete. A delay is placed between level transitions to stabilize the audience's response during these indeterminate states. After transitioning from one level to the next, the system waits for 32 measures before determining whether the activity level has changed. An exception to this rule occurs when transitioning into the highest energy level, which happens instantaneously, regardless of the previous state, to keep the participants from sliding directly back down to the previous energy level.

Within each energy level, various forms of real-time control are given to the participants, and this control increases with the energy level. This somewhat rigorous scripting of the data is done for two reasons. First, it is not yet known how to give a large group of untrained participants full control of sonic events and still achieve a structured output. Second, because the primary goal of the system is for the enjoyment of the users, it was decided to err on the side of a more ordered and pleasing experience, rather than a chaotic and potentially displeasing event, which would receive little participation.

To provide the appropriate feedback, as discussed in the previous section, the current state of the audience must be known. The aggregate receiver data is the only method available to determine this information, and as such, it becomes important to know how these data relate to such parameters as the dominant tempo and average activity level of the crowd. Because the amount of previous testing with the system was limited, some of these factors must be assumed at first and then correlated to event data to determine whether the assumptions are true.

The most basic piece of information available to the system is whether a sensor has sent an RF transmission (a condition referred to as a "hit"). Each hit indicates that a user has just crossed an acceleration threshold, e.g., by jerking the hand holding the sensor and deliberately giving input to the system. For low energy-level conditions, when few hits are arriving, these hits can be mapped directly to sonic events, giving prompt feedback to the users that the system is working and that they are contributing to its input. These sonic events also enable users to hear their neighbors' activities, allowing them to synchronize to one another.

Hits can also be integrated over various time periods, creating moving-average low-pass filters on the data. The longer the time period of accumulation, the more representative the data will be, although it will introduce a mean time lag of one half the total integration time. Shorter time periods are useful for correlating activity to current sonic events, or for detecting clustered activity, while longer time periods give an accurate view of the energy level of the crowd.

The perceived energy state of the audience is divided into five levels, with each level being proportional to the assumed maximum of four hits per person per measure. The music is then generated according to these energy levels. At levels less than one-quarter of the maximum, ambient textures are played, consisting mostly of drones, which neither discourage nor encourage activity. From one-quarter to one-half of the maximum, the drone continues, and rhythmic features and melodies are added. From one-half to three-quarters of the maximum, more complicated rhythmic patterns emerge, and arpeggiated lines are added. From 75 percent to 95 percent of the maximum, still more complicated arpeggiated lines appear, and the beat structure simplifies. The melodies and drones become less apparent, and the drum kit tends to dominate.

At activity levels above 95 percent, the music consists mainly of rhythmic tracks and arpeggiated lines, with the drum kit patches and selection of notes in the arpeggiated lines being determined solely by user input. More precisely, in this mode arpeggiated notes are played only if the instantaneous activity is above a given threshold, and the pitch of each note is higher if the instantaneous activity was higher. Other quantized percussion sounds (atop a continuous kick drum in 4/4 time) are similarly triggered by the instantaneous energy level, with a somewhat random mapping of instantaneous activity to the chosen drum voice. This mapping creates an indeterminate musical experience that can quickly deteriorate into chaos, signaling to the users that the highest energy state has been reached.

In this way, the complexity of the music increases with activity level as more patterns are included. Tracks in lower levels are assigned patterns that repeat on longer time scales and have sounds with longer attack and decay, whereas the sounds in higher-energy tracks are higher in pitch and have shorter attack and decay times. As participants increase in activity, they are given more control over the sounds in the patterns, creating a more energetic and chaotic experience.

In conjunction with these more scripted sonic events, there is also real-time timbre control. This is done to give a sense of responsiveness and allow the audience to immediately shape a portion of the musical experience. These real-time control parameters are determined by activity sums integrated over a short time period (80 msec and 500 msec) and modify filter parameters such as cut-off frequency, resonance, attack, and decay. They are also used to control the depth and frequency of low-frequency oscillators that perform both amplitude and pitch modulation of the tracks, allowing the music to undulate with the motions of the crowd.

Finally, to determine the dominant tempo of the audience, a rolling Fast Fourier Transform (FFT) of the received data signal is taken over a 30-sec window, returning the peak frequency at one-second intervals. This peak frequency can be easily converted to beats per minute (BPM) to set the global tempo of the music. To determine the percentage of the audience dancing at the dominant tempo, the magnitude of the FFT peak frequency can be found. This parameter can be used to encourage synchronization by making the rhythm more prominent as its magnitude increases. The tempo structure of the sensor data could be obtained more quickly via an autocorrelation or other techniques typically used in rhythm-tracking algorithms (Gouyon and Dixon 2005). As discussed in the next section, certain mappings augment or decrement the tempo of the generated music around the dominant tempo detected by the sensors to accelerate or dampen crowd activity.

To make the transition algorithms independent of the quantity of participants and musical tempo, the activity sums were normalized by the total number of sensors used by the dancers and scaled by the speed of the generated music. Table 1 summarizes some of the basic rules implemented in the musical mappings.

Testing and Results

Throughout the course of this work, many different stages of testing occurred, examining both the functionality of the hardware and the nature of the human response to the system. What follows is a summary of the most important tests (named after

Table 1. Summary of some Typical Rules Employed by the Interactive Musical Mappings

The activity level is adjusted by the tempo and by the number of people:

Normalized Activity = (number of hits per interval × quarter-note duration) / number of people

The style of music is set by the mean activity level:

Normalized Activity	*Quantized Activity*	*Style of Generated Music*
< 25%	level_1	ambient
25–50%	level_2	minimal techno
50–75%	level_3	house
75–95%	level_4	hardtrance
> 95%	level_5	hardcore/chaos

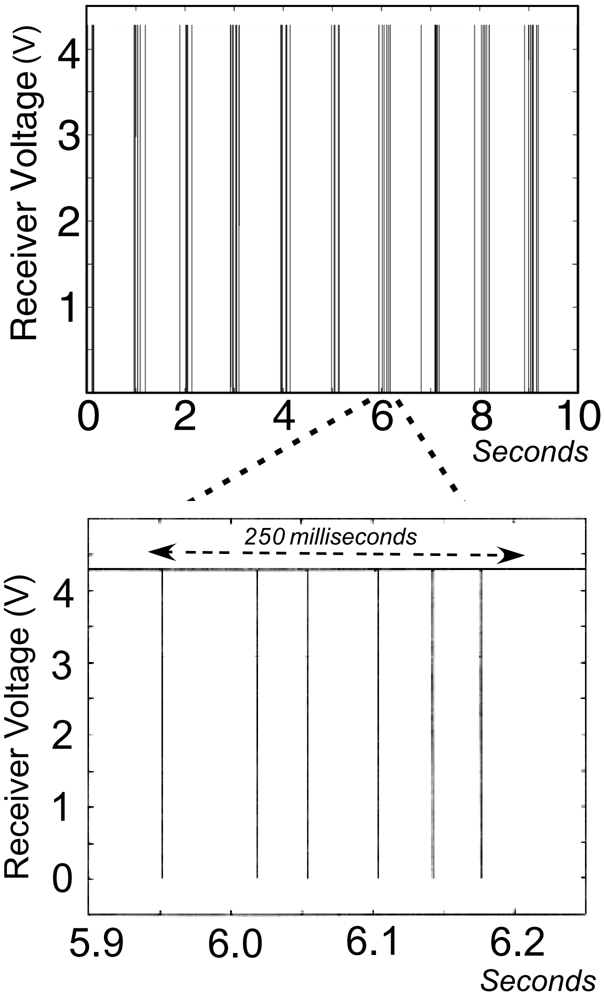
The BPM of the generated music is set to the BPM of the crowd (from the DFT) + *n*, where *n* is a mapping-specific tempo increment or decrement to suggest speed-up or slow-down (see text).

Activity levels with rolling averages over 80 msec, 0.5 sec, 1 sec, 2 sec, and 10 sec are used to modulate audio parameters such as pitch of arpeggiated lines, depth of LFO, depth of effects, filter resonance, and filter cutoff.

Transition from ambient mode to initial rhythm:

- During the ambient mode, when few hits are arriving, each hit is given a sharp, bright sound.
- As more hits enter, this dulls into a "wall of sound" that undulates with energy.
- With over 25% of sensors hitting simultaneously, a larger sound is created.
- After a succession of five of these events, a base beat at their mean interval is slowly faded in, hence the dancers build the initial rhythm.

Figure 5. Receiver voltage vs. time, showing sensor impulses arriving when six people attempt to beat together. The bottom plot is an expansion of one hit cluster, illustrating the low overlap probability for such narrow pulses.



the lab, conference, or dormitory that hosted the event) and their findings.

Media Lab Clapping Test

The first test of the system was conducted to ensure that the transmitters and data-collection systems functioned properly. This test consisted of seven people (mainly non-musicians) sitting around a table, each with a controller, attempting to clap in unison. One person was singled out beforehand as the leader, and all other participants were encouraged to follow that person's lead. In this manner, a tighter synchronization of clapping could be obtained.

A graph of the receiver output voltage versus time can be seen in Figure 5, which demonstrates that the transmitters worked well and that the sensors triggered with each clap. It also shows that such collective activity can be reliably detected, as the transmissions occur within a short time period around the leader's clap. The apparent clustering of sensor hits associated with each collective clap actually occurs over a relatively large time window in comparison to the 50-μsec transmitter pulse width. (See the expanded image in Figure 5.) Although, to the participants, the sound of the claps appeared to be in unison, the claps were actually spread out over approximately 250 msec. It is this result that supports the viability of the system for large groups; the random variation in user action reduces the likelihood of transmitter collisions to a negligible level. More quantitatively, if we model the time-arrival of clap pulses as a normal distribution with a standard deviation of 85 msec (as Figure 5 suggests), the average number of pulses that will overlap per clapping interval is only 1.8 for 100 people, 7.2 for 200 people, and 15.3 for 300 people, which is about the maximum number of people who can fit comfortably within the sensor transmission radius. (Note that the percentage of collisions rises linearly with the number of participants.) Accordingly, the sensor transmissions of a large crowd can be collected by counting individual pulses without significant loss of data from collisions.

UbiComp Sonic Tug of War

The first public application of this system was the *Sonic Tug of War,* run during the Workshop on Designing Ubiquitous Computing Games at UbiComp 2001 (Feldmeier and Paradiso 2001). As its name implies, this was a competition between two teams of users to bring the pitch of an audible tone up or down. Two base-station receivers were placed at opposite sides of a room to zone it into halves, and all players were given one wireless sensor. When a player on one side of the room jerked a sensor, the tone would ratchet up, and if a player on the opposite side of the room moved a sensor, the tone would ratchet down. The team that first moved an octave in their preferred direction would be declared the winner, and the tone would accordingly latch. Although this application was an extremely simple game, it showed that the system worked well and that the sensors could engage people with a very simple collective interaction.

Figure 6 shows the results of an experiment conducted to quantify the typical range profile for a sensor and displaced pair of base stations. The test was performed outdoors, and the base station antennae were placed 27 m apart. The sensor was moved from one base station to the other, pausing every 1.5 m, where shaking the sensor 20 times sent 20 pulses. Figure 6 illustrates the ratio of received versus transmitted signals as seen by each base station versus the distance from base station A. Aside from one transmission that was missed by base station A near the start of the test, 100% of the pulses were received within 10 m of each base station. The reception degrades between base stations across a range of roughly 20 m around the midpoint of the receive antennae. As can be seen in the plot, this transition is far from uniform, as multiple pathways from nearby buildings cause the reception to cycle from 0–100 percent. Furthermore, base station B was more sensitive than base station A, increasing base station B's relative reception area. Although using continuously received signal strength (RSSI) instead of the one-bit, received-signal gate adopted here could make this transition somewhat smoother, many factors (e.g., multipath, RF absorption and scattering by people, and differences in efficiency and tuning of transmitters and base stations) will keep the range boundary irregular, uncertain, and dynamically changing. Accordingly, although multiple base stations provide a coarse proximity suitable for general zoning, any mapping that uses the range information from this system must be tolerant to uncertainty and jitter.

Figure 6. Zoning demonstration showing the percentage of transmitted signals received by two displaced receivers (A dotted, B solid) as the sensor is moved along a line from one receiver to the other.

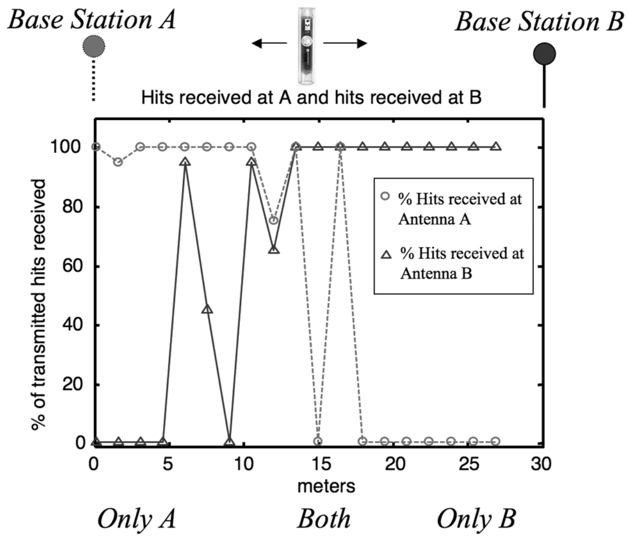
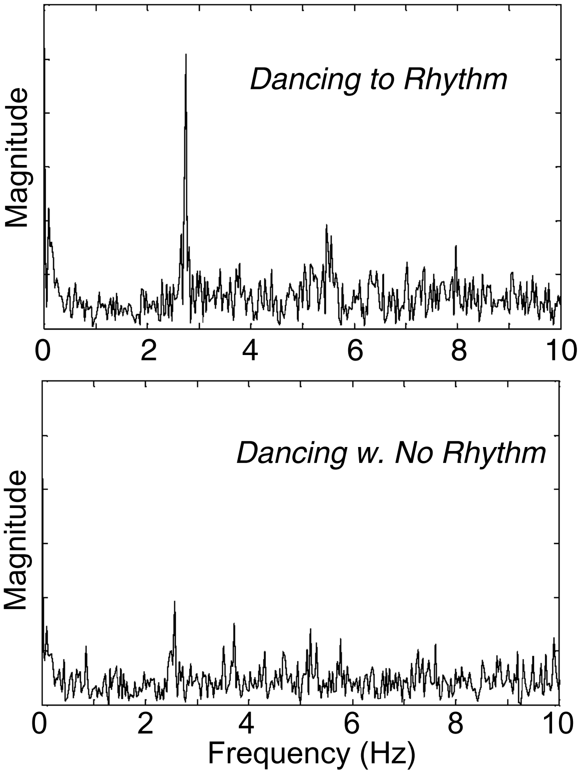


Figure 7. FFT results for sensor data from dancers responding to rhythmic music with a well-defined tempo (top) and to ambient music without a dominant tempo (bottom).



Talbot1

Because the system had not, to this point, been used for interactive dance, it was not known how users would respond to the system or what data parameters would be relevant. To answer these questions and to begin building effective musical mappings, sensor data was collected for a non-interactive dance event. Fifteen participants, each holding a sensor, danced simultaneously for half an hour to a "deejayed" set of electronic dance music. Received data patterns were noted and compared to the music played at that time. The music had an average tempo of 154 BPM and varied musically with ambient sections, strongly rhythmic sections, and portions with syncopated rhythms.

The sensor data was integrated over a ten-second period to give the rate of pulse arrival. A strong correlation between the rate of pulse arrival and the perceived energy level of the music was found. At the start of the event, the music was ambient and had no rhythm track. As a result, there was little motion by the dancers, and the rate of pulse arrival was low. When the rhythm track entered, the rate of pulse arrival increased as the dancers began to respond to the music. This pattern continued throughout the event. In cases where the music had no rhythm track or had a simple rhythm track and no melody, the rate of pulse arrival was about one-third that of sections with both a strong beat and melody.

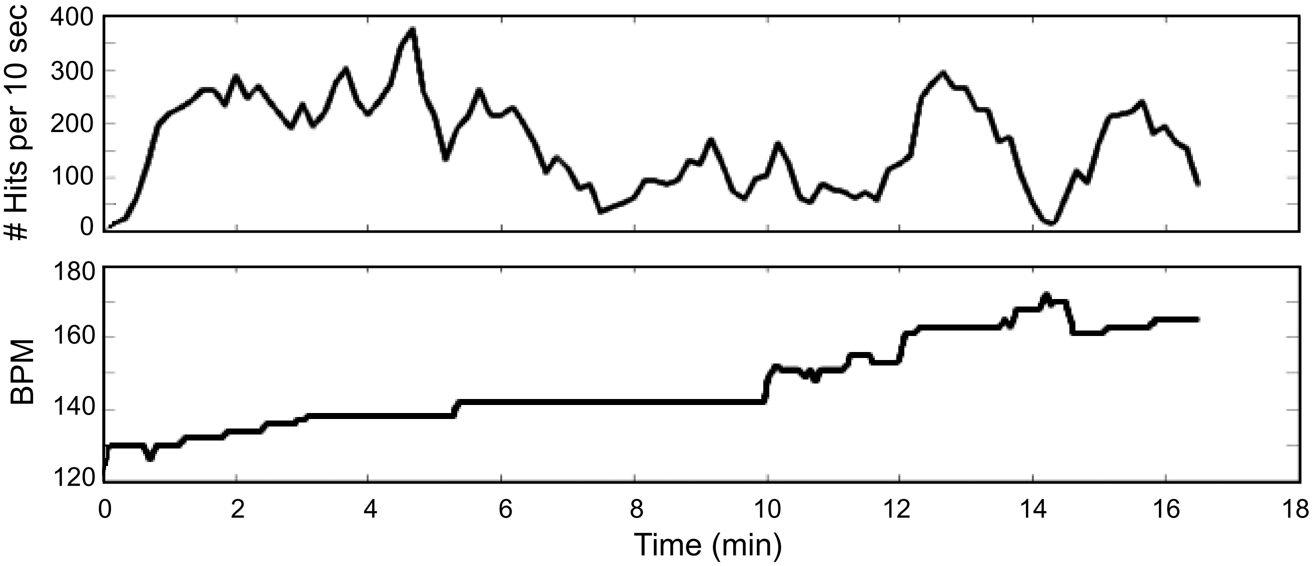
An FFT of the received signal correlated strongly to the average tempo of the music. An FFT of the data, taken in 30-sec increments, returned a peak frequency of 2.57 Hz (corresponding to a tempo of 154 BPM), for most sections of the music. The FFT could not detect a dominant frequency during periods of low activity or when the music was ambient or arrhythmic. (See Figure 7.) Despite this, we concluded that the system, in its most rudimentary applications, can detect the net activity level, the amount of coherent rhythm that is present, and dominant tempo of its users.

Talbot2

With the knowledge gained from the previous test, algorithms were developed that detected features of the group's behavior and generated matching musical pieces. Once again, 15 participants, each holding a sensor, danced to electronic music for half an hour. This time, the music was not "deejayed," but rather it was generated solely by the received data stream.

A number of hypotheses were tested with these mappings. First, the effectiveness of giving each person a sonic response per hit was evaluated as a method of giving feedback to the functionality of the system and as a tool to encourage synchronization of the participants. The participants quickly understood the causal nature of the interface as they moved their sensors, but the sound they were given had too long a decay, and as a result, the combined sounds would merely blend into each other rather than build to coherency. Despite this fact,

Figure 8. Data from Talbot2: activity level (top) and tempo as peak frequency of an FFT run on windowed sensor data (bottom), vs. time.



the participants began to synchronize at low energy levels.

Second, the effect of giving positive feedback to increasing activity level was examined. As can be seen in Figure 8, the activity tended to cycle between high and low levels. In this manner, the behavior followed the expected pattern described in the previous discussion on interaction design, suggesting that appropriate feedback was given. Finally, allowing the average motions of the group to set the tempo of the music showed that rhythmic coherency could be developed among the users. Setting the tempo of the generated music to the tempo detected by the sensors plus 1 BPM tended to encourage dancers to accelerate their timing, as seen in Figure 8, where the dominant trend is an increase in the tempo from 120 BPM to 172 BPM owing to participant control.

The dancers stated that they felt the music was responding to their motions, especially during the lower-energy states when the more causal sounds could be heard. They also stated that they felt like they were controlling the tempo, and in several instances, worked together to either raise or lower its level. Initially, they felt the music was engaging, but they became disinterested after all of the various patterns in these simple musical mappings had been exhausted.

Talbot3

Although the results of Talbot2 show the basic functionality of the system in allowing a group of dancers to control aspects of their musical environment, the musical mappings written for that event were very limited. As a result, the music peaked quickly and left little room for further exploration. Also, the control given to the dancers was confined to only tempo and energy level. Therefore, a set of less-scripted musical mappings was written, and more events were held, ranging from 25–200 participants. To recruit enough participants for these events, posters and flyers were distributed across campus (see Figure 9), and at some events the sensors were converted to dance props with the addition of a "glowstick" (see Figure 10). These events were captured on both audio and video to allow for visual verification of participant activity level and to correlate particular activity with various data parameters and the music being played. Finally, at the end of each event, a questionnaire was distributed to assess the participants' impressions of the experience. This section presents results taken from the last of these events, where the interaction mappings were the most advanced.

The increase in the amount of music written for

Figure 9. Poster and flyer to promote the interactive dance events across campus.

Figure 10. Photograph of participants dancing with sensors at an interactive event and close-up photograph of a sensor attached to a "glowstick."

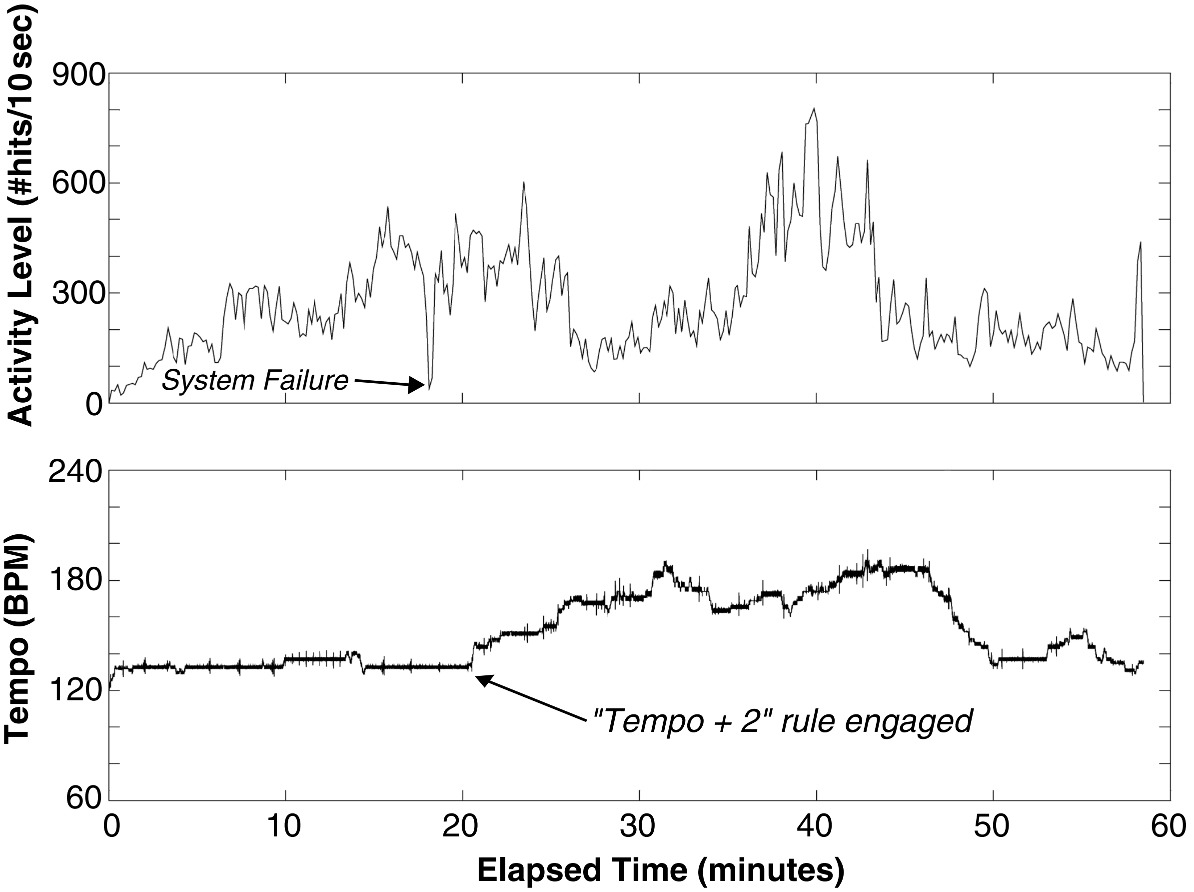
this event, in comparison to Talbot2, allowed for a greater degree of exploration by the participants. The ability to fade tracks in and out also added to the complexity. Perhaps the only potential disadvantage of the new mappings involved the modifications made to the tempo-setting function. The algorithm used in Talbot2 had taken a rolling FFT over the past 30 sec, with a 7-msec windowing function. This returned a very accurate value for the average tempo, with 2 BPM resolution, but it exhibited a 15-sec delay. As a result, the current value did not reflect the current activity. The algorithm used in Talbot3, in an attempt to improve the responsiveness of the tempo, had a windowing period of 16.4 msec, with a total sample time of 16.7 sec. This had the effect of decreasing the delay to 8 sec, but, correspondingly, decreased the resolution to 4 BPM.

The results of this change can be seen in Figure 11. Clearly, the tempo varied insignificantly for the first 21 minutes of the event, because the participants did not change the tempo enough to overcome the 4-BPM binning. Instead, they remained at their current tempo, until the tempo setting function was varied to add an increment of two to the current tempo, instead of just one. This was done at 21 min into the event, and the effect can be easily seen in Figure 11. The tempo began to ramp up quickly, as it led the dancers by a full 2 BPM. The function was then reset to +1, and the tempo plateaued.

Throughout the remainder of the event, the tempo-setting function was varied from –1 to +2. The net effects of these changes can be seen in Figure 11. At every point where the function was changed to +2, the tempo increased. For every point where the function was set to +1, the tempo plateaued. Furthermore, for every point where the function was set to –1, the tempo decreased. It seems expected that the tempo would increase when set to +2, and reasonable that it would plateau when set to +1, especially given the 4-BPM resolution. Surprisingly, however, the tempo decreased when set to –1. If the 4-BPM binning kept the participants from transitioning up in tempo when the function was set to +1, then it would presumably also keep the tempo from decreasing when the function is set to –1, especially considering that it has been shown that humans tend to increase their tempo when interacting as a synchronized group (Neda et al. 2000).

We hypothesize that people dancing to music

Figure 11. *Data from Talbot3: activity level and dominant tempo vs. time.*



whose tempo is greater than their desired tempo have an incentive to slow down their rhythm. The further they are away from this desired tempo, the greater this incentive becomes. Because the tempo continued to increase (seemingly out of control of the participants) and owing to the 4-BPM resolution, it reached a level (190 BPM) that was completely disparate from their desired tempo. At this point, we assume that desire to reduce the tempo overcame the 4-BPM tempo gap. A partial confirmation of this theory can be seen in the activity-level plot of Figure 11. As the tempo increased, between 23–33 min into the event, the activity level decreased. Furthermore, when the tempo decreased and plateaued after 30 min into the event, the activity level of the music again increased. These undulations in activity level tend to represent the participants' relative approval of the music. When the music slowed down, becoming more commensurate with their desired tempo, they would once again become engaged and raise the activity level. Accordingly, these tests indicated that the users were able to causally set their tempo with musical feedback changing in steps less than 4 BPM.

The results of Talbot3, in terms of activity level, are very similar to Talbot2, with one very important difference. In Talbot3, the generated music increases and decreases in energy from the lowest to the highest level, corroborating the finding of Talbot2 that users respond to positive feedback with increased activity until they reach an unstable state, at which time they will decrease, only to build again. However, the activity levels of Talbot3 do not vary as quickly or as much as the activity levels of Talbot2. The Talbot2 data, as seen in Figure 8, show the activity levels varying on 2–6 min

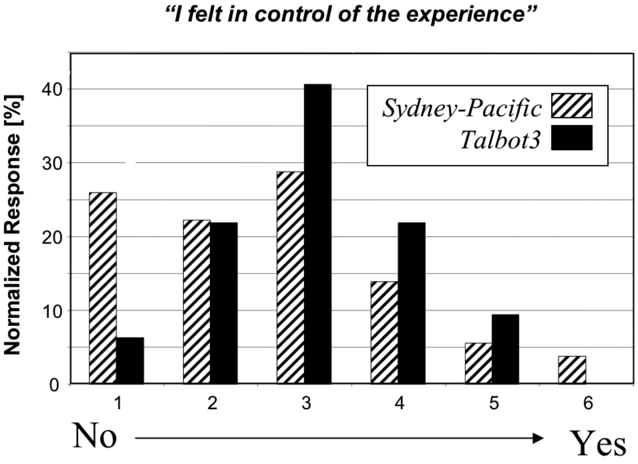


Figure 12. Degree of user engagement with Sydney-Pacific and Talbot3 events, indicating a clear preference for the latter.

time scales, spending most of the time in either the highest or lowest energy state. The Talbot3 activity levels, as shown in Figure 11, vary on 15-min time scales, with only one sharp dip happening at around 18 min, owing to a synthesizer that was momentarily set incorrectly and playing a very dissonant patch. The activity also lingers in the middle energy levels much longer, with the average time spent per level being 5 min. This is primarily owing to the 32-measure transition delay used to bring in new tracks, keeping the participants from fully receiving feedback until the system was certain that a transition was desired. The increased timbral control of the Talbot3 mappings resulted in more variations within a level, as participants shaped the music, no longer requiring a change in level for a change in the musical experience.

Perhaps the most encouraging result of this event is the fact that the participants stayed for over two hours. The majority of people who had attended danced consistently for the first one and a half hours of the event, with the last half hour being mostly loose experimentation by those still dancing. At Talbot2, the users were quickly disengaged as the music began to repeat and no longer varied with their activity, ending the event within one half hour. An inspection of the audio and video recordings of Talbot3 reveals that many grouped patterns and repeated rhythms developed, dictated solely by the users' input.

The results of a questionnaire distributed after Talbot3 indicated that the majority of participants enjoyed the event and felt that the music had responded somewhat to their actions. This is exemplified in Figure 12, which presents the normalized response to a question asking what degree of control users perceived across two events, namely Talbot3 (running with improved mappings) and Sydney-Pacific (where the mappings were mainly malfunctioning). It is encouraging to see that the users felt significantly more in control at Talbot3 (the average rating at Talbot3 was 3.4 out of 6, whereas the rating at Sydney-Pacific was only 2.6 out of 6). These results seem to imply that, given the appropriate feedback and constraints, a moderately sized group of people can assert some control over their musical environment and collaborate to create a musical outcome that they perceive as being generally pleasing and coherent. The fact that Talbot 3 (an event with about 25 people) rated, on average, barely above middle in Figure 12, however, indicates that our mappings can benefit from further development.

Conclusions and Future Work

Results of these events show the functionality of the sensor system as a useful tool for large group interaction. The handheld sensors adequately detect users' motions and transmit these data with negligible probability of collision. The low latency of transmission makes the sensors particularly well suited for use as a controller in applications like music and gaming, where causal feedback and control are required. More importantly, the sensors are unobtrusive and require no training to use, enhancing the participants' experiences. The extremely low cost of these sensors makes them feasible to be given away with a ticket to an event (or to be freely distributed to members of a special "interactive" club), and the low power drain allows these devices to stay alive on a small embedded coin cell for many years.

Along with the ability to collect simple data reflecting each individual's motion, the system demonstrates the ability to collect activity and rhythm information over groups of dancers. It also shows that this information can be effectively used to give a group of dancers causal control over the music to which they are dancing, changing, in real-time, aspects such as style, tempo, envelope, modulation, filter settings, and other timbral and voicing parameters. Although the musical mappings produced individual sounds with each arriving pulse at low activity levels, this direct link was abandoned in favor of mapping to the higher-level features introduced above as activity increased. Because it is well known that large groups of humans synchronize through prompt audio feedback—ranging from players in an orchestra through simple clapping (Neda et al. 2000)—it would be interesting to better exploit this tendency in future mappings that follow individual activity further towards constructing spontaneous structure.

Overall, participant response to the system was positive. The majority of attendees at the final (Talbot3) event enjoyed the experience and spent most of their time dancing. Although quite a few of the attendees felt that the music responded well to their motions, the majority desired more control over the experience. This poses a difficult question that is still left unanswered by this work. How can a large group of anonymous individuals be given appropriate feedback, such that each individual has a sense of close control over the central interaction, while ensuring adequate structure so that all participants find the interaction pleasing? Indeed, for large groups, the possibilities, although intriguing, may be quite limited.

The current system does not allow different sensors to be distinguished from one another. The trimmer capacitor on the current transmitter and receiver boards (see Figure 2), however, allows different RF frequencies to be set for different devices. This approach does not scale well to large numbers of IDs, but we have run this system with three base stations, each sensitive to a different frequency, giving us three distinguishable groups of sensors. Although bandwidth, power, cost, and latency limitations in large groups discourage digital transmission of individual ID codes, future low-cost hardware could transmit pulses with a much sharper edge, allowing the sensors to be better located with Ultra Wideband (UWB) techniques (Pahlavan et al. 2002) and enabling content to better exploit zoning. The noise performance of this system in our on-campus venues was adequate, as there was not much RF background noise at our operating frequency. Should this become more of a problem, one possibility for reducing spurious background hits would be to discriminate on the RF pulse width, rejecting outliers that come from other sources. Similarly, pulse width is a variable that may be able to encode a coarse analog parameter (e.g., jerk intensity), so long as the pulses are still short enough to maintain a low probability of collision in a crowded venue.

The hardware designed for our implementation could accommodate up to three RF receivers, but only two were used in the actual tests. It would be relatively straightforward, however, to design a base-station architecture around a scalable network, allowing an infrastructure of multiple receivers to cover a large venue. This also opens many interesting issues in developing content that works at both a local and global level. One implementation could involve providing tight causal audio feedback to proximate groups of fewer users, while allowing particularly strong patterns and pronounced collective gestures at the local level to influence global musical features that appear everywhere, creating perhaps something of a competitive crowd environment. Such spatial zoning could also be coupled with indicative lighting or display events, so the entire audience could become aware of dominant or even "soloing" groups of participants/dancers.

Although the examples in this article have concentrated on "techno" music, this system is amenable to enabling crowd interaction with other musical styles. For example, we have already collaborated with the jazz trombonist and composer George Lewis in a project examining musical interactions with children, and we are exploring a collaboration with composer Tod Machover for audience interaction in electronic-orchestral music. Going further, these sensors certainly appeal to other entertainment venues, for example, in interactive gaming or mass electronic "cheering" (such as group "wave" gestures) for people at sports stadiums and large outdoor events. Although developed for musical applications, these devices are an example of a system that easily crosses over into entirely different domains (Paradiso 2003). They are generally appropriate for niches where low-cost, coarse-resolution, high-longevity wireless sensing is called for—for example, creating a delivery package to trigger an alert if it falls in a truck, or scattering sensors around a smart house to track the activity of a sick or elderly occupant (Tapia and Intille 2004).

Acknowledgments

This work could not have been completed without the contributions of our research colleagues in the Responsive Environments Group, namely, Michael Broxton, Matt Malinowski, Mike Pihulic, and Josh Randall. We are grateful for the support of the MIT Media Lab's sponsors, the Things That Think and Digital Life Consortia, and Kyung Park from MSI Sensors for supplying the PVDF cantilevers. Videos and other information about this system can be seen online at: www.media.mit.edu/resenv/GiveawaySensors.

References

Akkawi, A., S. Schaller, O. Wellnitz, and L. Wolf. 2004. "Networked Mobile Gaming for 3G-Networks." *Proceedings of the Third International Conference on Entertainment Computing (ICEC 2004).* Berlin: Springer, pp. 457–467.

Alexander, T. 2005. *Massively Multiplayer Game Development 2.* Hingham, Massachusetts: Charles River Media.

Aylward, R., and J. A. Paradiso. 2006. "Sensemble: A Wireless, Compact, Multi-User Sensor System for Interactive Dance." *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2006).* Paris: IRCAM, pp. 134–139.

Baxter, L. 1996. *Capacitive Sensors: Design and Applications.* Piscataway, New Jersey: IEEE Press.

Berger, E., et al. 2001. "A Sophisticated Soiree." *Proceedings of the 2001 Ars Electronica Festival.* Vienna: Springer, pp. 352–353.

Blaine, T., and S. Fels. 2003. "Collaborative Musical Experiences for Novices." *Journal of New Music Research* 32(4):411–428.

Camurri, A., et al. 2000. "EyesWeb: Toward Gesture and Affect Recognition in Dance/Music Interactive Systems." *Computer Music Journal* 24(1):57–69.

Carpenter, L. 1993. "Method and Apparatus for Audience Participation by Electronic Imaging." U.S. Patent No. 5,210,604.

Carpenter, L. 1994. "Video Imaging Method and Apparatus for Audience Participation." U.S. Patent No. 5,365,266.

Cliff, D. 2000. "Hang the DJ: Automatic Sequencing and Seamless Mixing of Dance-Music Tracks." HP Laboratories Report HPL-2000-104. Bristol, UK: HP Laboratories.

Cutts, Q., G. Kennedy, C. Mitchell, and S. Draper. 2004. "Maximizing Dialogue in Lectures Using Group Response Systems." *Proceedings of the IASTED International Conference On Computers and Advanced Technology in Education (CATE 2004).* Calgary: ACTA Press, pp. 421–426.

Eng, K. 2004. "Designing Neuromorphic Interactive Spaces." Ph.D. Dissertation No. 15451, Swiss Federal Institute of Technology (ETHZ).

Falk, J., P. Ljungstrand, S. Bjork, and R. Hansson. 2001. "Pirates: Proximity-Triggered Interaction in a Multi-Player Game." *Proceedings of the CHI 2001 Conference on Human Factors in Computing Systems.* New York: ACM, pp. 119–120.

Feldmeier, M. 2002. "Large Group Musical Interaction Using Disposable Wireless Motion Sensors." Master's Thesis, MIT Media Laboratory.

Feldmeier, M., M. Malinowski, and J. A. Paradiso. 2002. "Large Group Musical Interaction Using Disposable Wireless Motion Sensors." *Proceedings of the ICMC 2002 Conference.* San Francisco: International Computer Music Association, pp. 83–87.

Feldmeier, M., and J. A. Paradiso. 2001. "Ultra-Low-Cost Wireless Motion Sensors for Musical Interaction with Very Large Groups." Paper presented at UBICOMP 2001, Atlanta, Georgia, 30 September.

Feldmeier, M., and J. A. Paradiso. 2004. "Giveaway Wireless Sensors for Large-Group Interaction." *Proceedings of the ACM Conference on Human Factors and Computing Systems (CHI 2004).* New York: ACM Press, pp. 1291–1292.

Freeman, M. J. 1986. "Dynamic Audience Response Movie System." U.S. Patent No. 4,591,248.

Gouyon, F., and S. Dixon. 2005. "A Review of Automatic Rhythm Description Systems." *Computer Music Journal* 29(1):34–54.

Graham-Rowe, D. 2001. "Don't Stop Moving." *New Scientist* 172(23):25.

Helin, V. 2003. "Developments of Modern Multiplayer Games." M.S. Thesis, Helsinki University of Technology (TKK), Department of Computer Science and Engineering.

Kashio, U., and A. Yoneiki. 1991. "Electronic Percussion Instrument." U.S. Patent No. 4,995,294.

Laibowitz, M., et al. 2006. "A Sensor Network for Social Dynamics." *Proceedings of the Fifth International Conference on Information Processing in Sensor Networks (IPSN 2006).* New York: ACM, pp. 483–491.

Levin, G. 2001. "Dialtones: A Telesymphony." *Catalog of Ars Electronica 2001.* Vienna: Springer, pp. 54–59.

Maynes-Aminzade, D., et al. 2002. "Techniques for Interactive Audience Participation." *Proceedings of the Fourth IEEE International Conference on Multimodal Interfaces (ICMI 2002).* Los Alamitos, California: IEEE Computer Society Press, pp. 15–20.

Neda, Z., et al. 2000. "The Sound of Many Hands Clapping." *Nature* 403:849.

Obermaier, K. 2004. "Apparition." *Proceedings of Ars Electronica 2004.* Ostfildern-Ruit, Germany: Hatje Cantz Verlag, pp. 314–318.

Paavilainen, J. 2003. *Mobile Games.* Indianapolis, Indiana: New Riders Publishing.

Pahlavan, K., L. Xinrong, and J. P. Makela. 2002. "Indoor Geolocation Science and Technology." *IEEE Communications Magazine* 40(2):112–118.

Paradiso, J. A. 1999. "The Brain Opera Technology: New Instruments and Gestural Sensors for Musical Interaction and Performance." *Journal of New Music Research* 28(2):30–149.

Paradiso, J. A. 2003. "Dual-Use Technologies for Electronic Music Controllers: A Personal Perspective." *Proceedings of the 2003 Conference on New Instruments for Musical Expression (NIME 2003).* Montreal: McGill University, pp. 228–234.

Paradiso, J. A., K. Hsiao, A. Benbasat, and Z. Teegarden. 2000. "Design and Implementation of Expressive Footwear." *IBM Systems Journal* 39(3,4):511–529.

Picard, R., and J. Scheirer. 2001. "The Galvactivator: A Glove that Senses and Communicates Skin Conductivity." *Proceedings from the 9th International Conference on Human-Computer Interaction.* Mahwah, New Jersey: Lawrence Erlbaum Associates, pp. 1538–1542.

Pikovsky, A., M. Rosenblum, and J. Kurths. 2001. *Synchronization: A Universal Concept in Nonlinear Sciences.* Cambridge, UK: Cambridge University Press.

Randall, J. 2002. *Real-Time Lighting System for Large-Group Interaction.* Bachelor's Thesis, MIT Department of Electrical Engineering and Computer Science.

Siegel, W., and J. Jacobsen. 1998. "The Challenges of Interactive Dance: An Overview and Case Study." *Computer Music Journal* 22(4):29–43.

Tapia, E. M., and S. Intille. 2004. "Activity Recognition in the Home Using Simple and Ubiquitous Sensors." *Proceedings of the Second International Conference on Pervasive Computing.* Vienna, Austria. Berlin: Springer, pp. 158–175.

Ulyate, R., and D. Bianciardi. 2002. "The Interactive Dance Club: Avoiding Chaos in a Multi-Participant Environment." *Computer Music Journal* 26(3):40–49.

Wallin, N. L., B. Merker, and S. Brown, eds. 2001. *The Origins of Music.* Cambridge, Massachusetts: MIT Press.

Wechsler, R., F. Weiss, and P. Dowling. 2004. "EyeCon: A Motion Sensing Tool for Creating Interactive Dance, Music, and Video Projections." *Proceedings of the AISB 2004 COST287-ConGAS Symposium on Gesture Interfaces for Multimedia Systems.* Brighton, UK: Society for the Study of Artificial Intelligence and the Simulation of Behaviour, pp. 74–79.

Weinberg, G. 2005. "Interconnected Musical Networks: Toward a Theoretical Framework." *Computer Music Journal* 29(2):23–39.

Weiser, M. 1991. "The Computer for the Twenty-First Century." *Scientific American* 265(3):66–75.