Large Group Musical Interaction using Disposable Wireless motion sensors

Mark Feldmeier, Mateusz Malinowski, Joseph A. Paradiso
Responsive Environments Group
M.I.T. Media Laboratory
1 Cambridge Center, 5FL
Cambridge, MA 02142,
U.S.A.
carboxyl@mit.edu, joep@media.mit.edu, mattxmal@mit.edu

Abstract

We have developed a set of low-cost, wireless motion sensors that enable a large group of dancers to participate concurrently in a real-time, interactive musical performance. These sensors are either worn or held by participants and transmit a short RF pulse at the extremes of limb motion. The RF pulses are received by a base station and analyzed to detect rhythmic features and estimate the general activity level of the group. These data are then used to generate music that can either lead or follow the participants' actions, thereby tightening the feedback loop between music and dancer.

Keywords

Wireless, Sensor, Accelerometer, Motion, Rhythm, Large group, Music, Interactive, Dance.

INTRODUCTION

One of the unsolved problems in interactive music and entertainment is making environments that reflect and react to the collective activity of groups with tens, hundreds, or even thousands of participants. Generating content on this scale involves many challenges. For example, how is the individual granted low latency control and a sense of causality, while still allowing for information retrieval from all participants so that the environment responds to the behavior of the entire group? These issues are particularly pertinent in the area of interactive dance. Dance interfaces [1,2] exist that allow a single 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. The problems of cost, data communication bandwidth, and system responsiveness become increasingly more 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 to a heightened sense of expressiveness.

BACKGROUND

Current work in the area of large group interaction is dominated by interactive gaming. For groups of less than a hundred participants, there exist fixed systems such as voting interfaces for game show audiences (usually pushbuttons located in the armrests of chairs). But, for participants in numbers over a hundred, hardwired solutions become costly, and do not allow the participants to be mobile. Some systems enable many participants to become engaged via wireless PDA's [3], but these are quite costly and generally not real-time. Systems that look for cues from infrared cameras [4], microphones [5], or capacitive sensors [6] can gather information over a large, mobile audience, but they do not lend themselves to direct control by an audience member. 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 amongst each participant. This has been done via machine vision (see Carpenter's red-green voting paddles [7] and Picard and Scheirer's glowing Galvactivator skin-resistance detectors [8]), but requires a line-of-sight from camera to participant and is susceptible to illumination effects and background lighting. In general, making these types of measurements using non-contact methods such as machine vision or machine listening is not as accurate as direct methods such as wearable or handheld sensors. One example of such a wearable device was used in the Sophisticated Soiree installation [9] 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. However, these systems, which measure autonomic responses such as heart rate and skin resistance, are generally not consciously controllable by participants. Dave Cliff's "HPDJ" [10] is a hybrid that envisions providing dancers with an accelerometer, heart rate monitor, perspiration sensor, and Bluetooth wireless link. This relatively expensive package would be used to gauge the general activity level of a club's crowd, choosing appropriate tracks and tempos via a genetic algorithm to keep them dancing.

By contrast, systems based on real-time kinetic inputs can provide for very causal control. For such a system to work, each participant must be supplied with a controller that has a consistent response given a particular input, is intuitive to use, and lasts many hours, if not weeks, without draining its battery. Finally, for the controllers to be viable on scales of hundreds to thousands of participants, they must be inexpensive, wireless, and have a scalable communications platform.

HARDWARE DESIGN

The chosen design [11] is a small, inexpensive, wireless transmitter that sends a short burst of RF energy whenever it senses acceleration greater than a predetermined level. These transmitters can be either worn or held by a participant, and are activated by motion. Both the strength and duration of the RF burst are kept to a minimum for purposes of data collection and energy

Presented at the 2002 International Computer Music Conference, Gothenburg, Sweden, September 16-21, 2002, pp. 83-87.

conservation. The short transmission radius creates a zone of interaction around the receiving antenna, and the short RF burst duration reduces the probability of collisions between signals. In this way, the pulses in a particular area can be added up to give a sense of the rhythm and activity of the local participants, while still receiving each participant's action as a distinct event.

In its current form, the controller consists of a trigger, debouncing circuitry, and an RF transmitter. The trigger is a piezoelectric (PVDF) film cantilever, weighted to set the level of acceleration sensitivity. Whenever the controller is accelerated past this sensitivity threshold, the PVDF triggers a dual CMOS timer. The first half of this timer produces a 100ms pulse to eliminate double triggering due to PVDF film ring down, and the second half of the timer produces a 50µs pulse that activates the RF transmitter. The 300MHz LC oscillator transmitter has a 3 - 10m effective transmission radius, depending on the RF environment. The power for the controller comes from a single 3-volt lithium coin cell. The circuit consumes less than .01µA in standby, and an average of a few microamps during the 100ms debouncing operation. At the rate of two transmissions per second, the battery would last for a month of continuous usage, and indefinitely with no usage.

To simplify the design and provide minimal use of bandwidth (minimizing the probability of interference), 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.

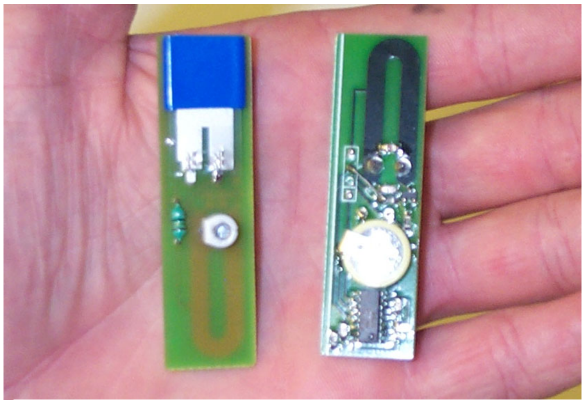

Figure 1. Prototype sensor, front and back view.

Twenty prototype controllers have been assembled and tested (see Figure 1). They measure 6.2cm x 1.6cm x 1cm and weigh 5g. These dimensions could be significantly reduced if the sensor were to be redesigned for mass manufacture. Chip-on-board technology, specially designed piezo elements, and a smaller battery could easily reduce the sensor to the size of a watch. For quantities of ten, the price is currently eight dollars a unit. This is dominated by the cost of the printed circuit board and battery. For much larger quantities, the total cost could be under a dollar per controller for parts, making it viable as a giveaway item.

INTERACTION DESIGN

We do not plan to independently ID each performer, but instead measure and react to the characteristics of ensemble behavior. The sensor's RF pulses are received by a base station, which counts the number of pulses received within a short (2ms) and periodic window. This information is sent via MIDI serial communication as channel pressure information to a computer, where it is processed in the MAX programming environment.

Sample sensor data was first taken for a non-interactive dance environment, and received data patterns were noted and compared to the music playing at that time. Fifteen participants, each holding a sensor in his or her hand, danced simultaneously for a half hour to a deejayed set of electronic dance music. The music had an average tempo of 150bpm and varied musically with ambient sections, strong beat sections, and portions with syncopated rhythms. The rate of pulse arrival for the entire event can be seen in Figure 2. A strong correlation between rate of pulse arrival and perceived energy level of music was found. In cases where the music had no beat, or had a simple beat and no melody, the received pulse rate was one third that of sections with both a strong beat and melody. A ten second sample of the received signal can be seen in Figure 3. Although it appears to have very little rhythm information, an FFT of the received signal correlated strongly to the average tempo of the music, as shown in Figure 4. An FFT of the data, taken in thirty-second increments, returned a peak frequency of 2.5Hz (150bpm), for most sections of the music. For ambient sections, during which there was no beat, the FFT could not detect a dominant frequency. It is concluded, therefore, that the system, in its most rudimentary applications, can detect both the activity level and average rhythm of its users.

This information was then used to develop algorithms that detected features of the group behavior and generated matching musical pieces. The peak frequency of an FFT was used to set the music's tempo. The rate of pulse arrival was classified into one of five levels. The first represented the lowest activity state, for which there was no beat generated, and merely low drones played. The next level brought in a simple, soft beat, and changed the drone to a more complex, higher pitched voice. The third level increased the complexity of the beat and introduced an even higher drone. The fourth level added melody and a harder beat. The final level added another melody and began distorting the beats.

In addition to these controls, each received pulse triggered a subtle chime, and if more than four pulses were received within an 80ms period, a condition hereafter referred to as clustering, a louder and longer duration chime was triggered. Also, the rate of change of the received pulse signal was used to modulate the cutoff frequency and volume of the melody voices. In this manner, we give both immediate responses for a sense of causality and control, and delayed responses that modify global parameters for a sense of responsiveness and complexity.

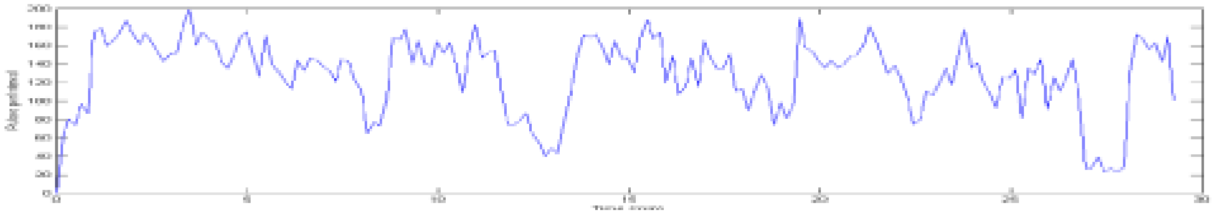

Figure 2: Non-interactive dance environment, received pulse rate, 30-minute segment

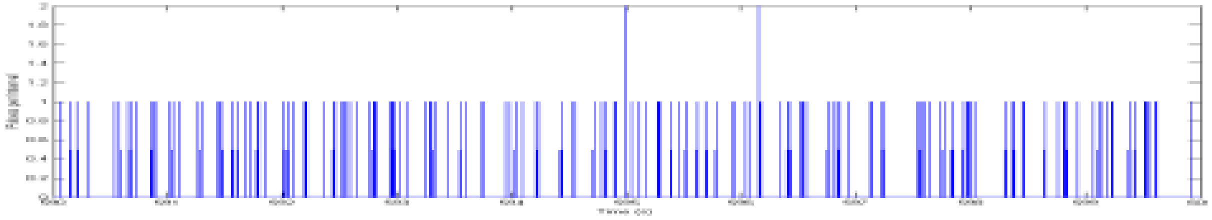

Figure 3: Non-interactive dance environment, received pulse rate, ten-second sample

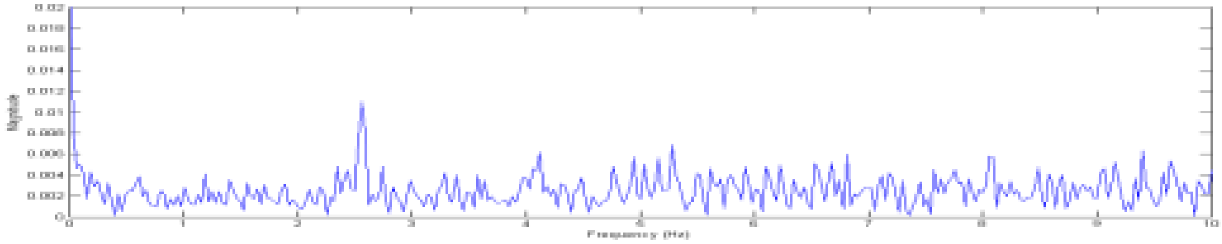

Figure 4: Non-interactive dance environment, FFT of 30-second sample

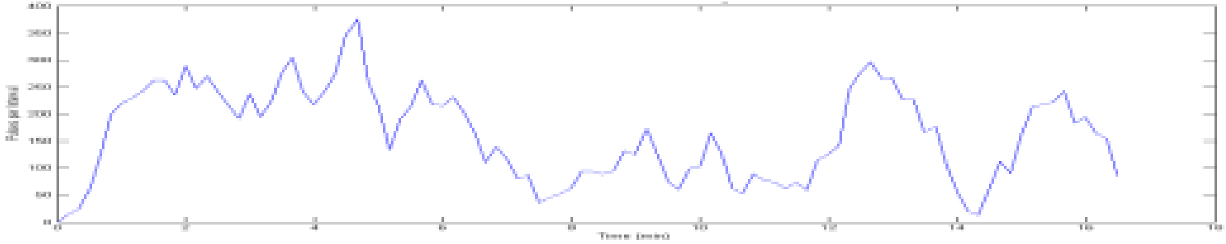

Figure 5: Interactive dance environment, received pulse rate, 17-minute segment

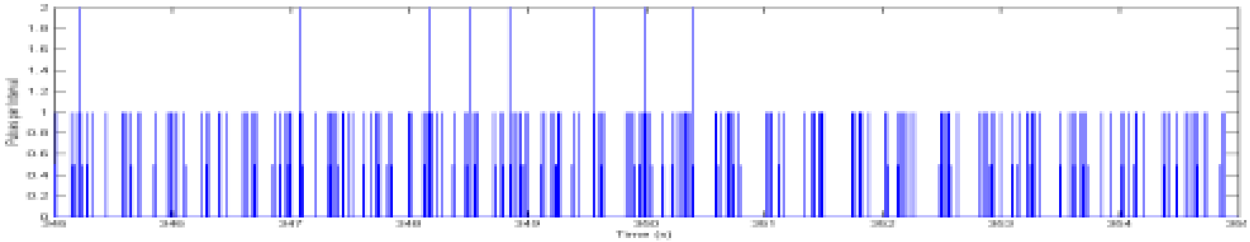

Figure 6: Interactive dance environment, received pulse rate, ten-second sample

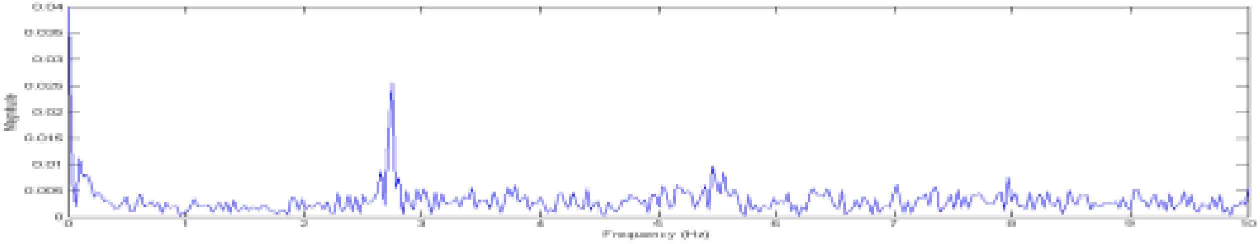

Figure 7: Interactive dance environment, FFT of 30-second sample

RESULTS

As a test of the functionality of the system, a fully interactive dance event was held. Once again, fifteen participants, each holding a sensor, danced to electronic dance music for a half hour. This time the music was not deejayed, but rather generated by the received data stream of the sensors. A laptop computer was used to receive the MIDI data stream, store the event data, perform the FFT and other data processing algorithms, and send out a corresponding MIDI data stream to a synthesizer that generated the music. The aforementioned musical mappings were used as the basis of the musical content.

The received data from this event varied significantly in a number of ways from the data for the non-interactive dance environment. The rate of pulse arrival for a 16 minute section of the event is shown in Figure 5. The rate of pulse arrival was fifty-percent higher than for the non-interactive event, denoting a higher average activity level. Also, the occurrence of data clustering was an order of magnitude greater, and the received signal showed more rhythm information, as can be seen in Figure 6. This increased synchronization of the dancers led to a higher magnitude of the FFT peak frequency, as shown in Figure 7. Finally, the tempo of the music varied from 60bpm to 195bpm, due the control the dancers had over the music.

The dancers stated that they felt that the music was responding to their motions, especially during the lower energy states when the more causal chimes could be heard, and a greater variation of music was occurring. They also felt as though 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 became disinterested after all of the various voices in the musical mappings had been exhausted. This is primarily due to the fact that the activity level was higher than in the non-interactive event upon which the mappings were modeled, so the music peaked much sooner, leaving no room for exploration.

CONCLUSIONS

These preliminary tests of the system give very promising results as to the system's ability to enable a group of dancers to control the music to which they are dancing. Perhaps the most encouraging result was merely the statements of the participants as to how much they enjoyed using the system and felt it was reacting to their motions. Our next efforts will involve changes to the hardware to increase the trigger's causality and enable the manufacture of hundreds of devices. Further tests will need to be conducted to ensure that the system scales effectively to the level of hundreds or thousands of participants. Also, more data processing algorithms will need to be developed in order to extract further patterns from the dancer's behavior. Musical mappings will then need to be written which respond to these features to direct the group behavior into particular patterns. This work will leverage the intrinsic human schooling behavior [5], especially around rhythm, making the experience complex enough to be interesting and aesthetically pleasing.

Although we have done some tests that zone the transmitter locations via amplitude discrimination between multiple basestations, this was not robust enough to warrant further work. Improvements to the base stations may produce significantly better location estimates through time-difference-of-arrival [12], hence enabling content to casually shift with the user's location and provide another expressive degree of freedom.

ACKNOWLEDGMENTS

This work could not have been completed without the contributions of our Responsive Environments colleague Michael Broxton. We are grateful for the support of the MIT Media Lab's sponsors and the Things That Think and Digital Life Consortia.

REFERENCES

[1] See: http://www.troikaranch.org/mididancer.html

[2] Paradiso, J., Hsiao, K., Benbasat, A., Teegarden, Z., "Design and Implementation of Expressive Footwear," IBM Systems Journal, Volume 39, Nos. 3 & 4, October 2000, pp. 511-529.

[3] Falk, J., Ljungstrand, Pl, Bjork, S., and Hansson, R., "Pirates: Proximity-Triggered Interaction in a Multi-Player Game," Proc. of the CHI 01 Conference on Human Factors in Computing Systems, Extended Abstracts, ACM Press, New York, 2001, pp. 119-120.

[4] Ulyate, R., Interactive Dance Club, SIGGRAPH98 Conference Abstracts and Applications, ACM Press, New York, 1998, pp. 230-235.

[5] Neda, Z., Ravasz, E., Brecht, Y., Vicsek, T., and Barabasi, A.L., "The sound of many hands clapping," NATURE, vol 403, p849 (2000).

[6] Waxman, D., Digital Theremins: Interactive Musical Experiences for Amateurs using Electric Field Sensing (MS Thesis), Massachusetts Institute of Technology, Cambridge, Massachusetts, 1995.

[7] See: http://www.cinematrix.com/

[8] Picard, R., Scheirer, J., "The Galvactivator: A Glove that Senses and Communicates Skin Conductivity," Proceedings from the 9th International Conference on Human-Computer Interaction, New Orleans, August 2001.

[9] Berger, E., et al., "A Sophisticated Soiree," Take Over, Proc. of the 2001 Ars Electronica Festival, Springer-Verlag, Vienna, 2001, pp. 352-353.

[10] Graham-Rowe, D., "Don't Stop Moving," New Scientist, v. 172, No. 23, November 17, 2001, p. 25.

[11] Feldmeier, M. and Paradiso, J.A., "Ultra-Low-Cost Wireless Motion Sensors for Musical Interaction with Very Large Groups," Presented at the Workshop on Designing Ubiquitous Computing Games, ACM UBICOMP Conference, Atlanta GA, Oct. , 2001. See: http://www.playresearch.com/workshops/ubigame.ubicomp/papers/feldmeier.pdf

[12] Pahlavan, K, Xinrong, L., Makela, J.P., "Indoor geolocation science and technology," IEEE Communications Magazine, vol.40, no.2, Feb. 2002, pp.112-118.