eMersion I Sensor-controlled Electronic Music Modules & Digital Data Workstation

Chester Udell
University of Oregon
838 Pool St
Eugene, OR 97401
cudell@uoregon.edu

James Paul Sain
University of Florida
P.O. Box 117900.
Gainesville, Fl 32611
jsain@uoregon.edu

ABSTRACT
In our current era, where smartphones are commonplace and buzzwords like "the internet of things," "wearable tech," and "augmented reality" are ubiquitous, translating performance gestures into data and intuitively mapping it to control musical/visual parameters in the realm of computing should be trivial; but it isn't. Technical barriers still persist that limit this activity to exclusive groups capable of learning skillsets far removed from one's musical craft. These skills include programming, soldering, microprocessors, wireless protocols, and circuit design. Those of us whose creative activity is centered in NIME have to become polyglots of many disciplines to achieve our work. In the NIME community, it's unclear that we should even draw distinctions between 'artist' and 'technician,' because these skillsets have become integral to our creative practice. However, what about the vast communities of musicians, composers, and artists who want to leverage sensing to take their craft into new territory with no background in circuits, soldering, embedded programming, and sensor function? eMersion, a plug-and-play, modular, wireless alternative solution for creating NIMEs will be presented. It enables one to bypass the technical hurdles listed above in favor of immediate experimentation with musical practice and wireless sensing. A unique software architecture will also be unveiled that enables one to quickly and intuitively process and map unpredictable numbers and types of wireless data streams, the Digital Data Workstation.

Keywords
Wireless Sensor Mesh Network (WSN), eMotion, eMotes, eMersion System, DIY, PnP, Digital Data Workstation (DDW)

1. INTRODUCTION
1.1 Thought Activity
Andy Russell explains an idea that is integral to this system's architecture in his article called *What Makes a Great Toy for Creative Play* [1]. Think back to the first system you used as a child to create, model, and experiment with. It was probably something like LEGOS, Tinkertoys, or Lincoln Logs. These toys are comprised of self-similar, interconnecting pieces that one can assemble, reconfigure, and literally play with to create endless variety and fulfill individual aesthetics and goals. These architectures are effective at inspiring innovation and stimulating creative play because they empower children to construct and use their own inventions. What if composers, musicians, dancers, artists, K-12 students, etc. were empowered to use sensing technology to control their digital world with no requisite knowledge in microprocessors, circuits, or programming? What implications would this have for expanding the NIME community? The eMersion system opens up this learn-by-doing ethos to a broader population that has creative skill and vision, but limited technical expertise and time. The architecture of these kinds of toys is applied in an attempt to respond to three emergent NIME issues: single-serving innovations, accessibility, and invasiveness.

NIME'14, June 30 – July 03, 2014, Goldsmiths, University of London, UK.

2. CURRENT ISSUES
2.1 Reinventing Wheels
Each newly created sensor interface invariably becomes one-of-a-kind and characteristic to individual aesthetics and projects. For example, Wanderley and Orio posit, "extended instruments have been designed to fit the idiosyncratic needs of performers and composers, but as such they have usually remained inextricably tied to their creators" [2]. One may arguably generalize this sentiment beyond "extended instruments" to include all types of NIMEs. Consider, for example, how often a particular NIME is widely adopted and used for a significant period of time (e.g. Theremin, Radio Baton, EWI/EVI) versus those that remain exclusively bound to the life of their inventors. Furthermore, the particular innovative developments these NIMEs employ commonly disseminate only in a limited number of instances (conference presentations in particular) preventing many valuable innovations and performance practices from reaching the broader community. To further detail this issue of sustainable innovation, Mark Applebaum mentions in an essay *Culture Sculpture*,

"Inventing a new instrument provides immediate gratification: one instantly becomes the world's greatest player of that instrument. The problem is that one abruptly realizes that one is also the world's worst player. So the satisfaction that comes from being novel is tempered by the fact that there is no communal standard by which to form a meaningful judgment, no cultural practice [3]."

It is important to observe, however, that the single-serving approach to instrument development is not counterproductive. This hand-made ethos is a catalyst for the rich diversity of solutions we have today. While the solution may not be to flood the community with yet more diverse technologies, identifying what we are forced to repeatedly build from scratch in most every project and creating alternatives that quickly enable one to overcome these recurrent technical barriers may help to expand our practices to include a larger community and inspire new innovations by building on older systems.

2.2 Accessibility

Despite valuable advancements in wireless networking and microprocessors, along with growing DIY (do-it-yourself) communities like ARDUINO, Sparkfun, and Adafruit, the technology is at a point that still requires at least a hobbyist's knowledge of microprocessors, programming, and sensors to even begin to experiment with novel interactive systems. While it is arguable to what degree of intensity this barrier may be (as it widely differs with individual personality, experience, and project), it's not unreasonable to posit that it does limit the use of NIME practices almost exclusively to those who can acquire a certain level of technical expertise beyond conventional musical practice.

2.3 Invasiveness

The display of wires can sometimes lend itself to a technically appealing visual aesthetic. The tradeoff (through our own experience) is the need for a performer to significantly alter their playing style to contend with shifting weight, decreased flexibility, and the risk of disconnecting sensors. One popular and long-produced sensor system, Electrotap's Teabox [4], connected sensors to a central hub using quarter-inch cable, making it challenging to attach complex networks of sensors onto moving or small sensing surfaces. Wires are still meddlesome factors even for many 'wireless' sensor interfaces that employ a central transmitter hub (like WISEAR[5] and La Kitchen's Kroonde Gamma). It would be of great value to eliminate the need for wires entirely by dedicating one wireless transmitter for each individual sensor.

By dedicating a wireless transmitter for each sensor in an environment, one can eliminate the need for wires entirely and limit the application of electronics exclusively to the desired sensing surfaces. Admittedly, this sensor-transmitter structure is not ideal in all situations (e.g. measuring the fingertips of a pianist) due to the size of transmitters today, but miniaturization of hardware could make this constraint moot in the near future.

3. EMERSION SYSTEM

3.1 Back to Toys

What do systems like LEGOS, Tinker Toys, and Lincoln Logs all have in common? They are comprised of self-similar, inter-connectable, modular components by which anyone can construct a seemingly infinite variety of combinations and configurations to suit personal desires. Robert Moog, in a paper called *Voltage-Controlled Electronic Music Modules* (the title from which this paper is derived) solved an incredibly complex issue facing synthesizer hardware [6]. Moog revolutionized the ad-hoc construction methods of synthesizers by "chunking" the complicated processes of synthesis into discrete "blocks" or units, including oscillators, amplifiers, filters, envelopes, and modulators. By interconnecting these modules, one could come up with a unique sound without having to redesign the synthesizer itself. What does this have to do with NIME? Taking a similar approach to wireless sensor interface design, eMersion is an all-purpose, modular sensor network that a user can implement to quickly and intuitively assemble wireless sensor nodes for control/interaction of computer processes including sound and visual effects. The novelty of this system with regards to all other human-computer interfaces right now resides in the network architecture. Each embodiment of this network is distinctive from user-to-user depending on the unique aesthetic and technical demands of the individual or project. Although sensor configurations and mapping may vary widely from one use to the next, the core hardware/software itself need not be redesigned. The design philosophy adheres to four general principles.

3.2 Four Principles

3.2.1 Modular

An eMersion sensor network is comprised of individual "eMotes" (a transmitter/sensor pair) and a single universal receiver that collects the sensor transmissions and sends compiled packets to the computer via USB. Each eMote is dedicated to only transmitting the data of its attached sensor and broadcasts its own unique identifier. For instance, one eMote may be dedicated to transmitting sonar data and another may be dedicated to measuring ambient light. The user is then free to utilize one, both, or neither of these by turning them on or off – attaching or detaching. Multiple eMotes of the same sensor type may also be used (for example, a matrix array comprised of a group of sonar nodes may be constructed). In this manner, one can acquire a variety sensor eMotes and use a subset of them in any combination as needed for a particular application. When one sensor is turned off (in the event of battery depletion, user preference, out of range, etc) the other eMotes on the network remain unaffected.

3.2.2 Reconfigurable

The nature of this modular interface design allows the user to easily reconfigure the placement and combination of individual sensors for any given project and instrument. This enables the user to find the most optimized sensor combination, sensing locations, and mapping scheme to meet their unique aesthetic goals in realtime – without having to redesign or reprogram the hardware itself. The sensor nodes may even be distributed amongst several instruments or non-instrumental performers, like dancers, to create novel modes of musical interactivity. For instance, a dancer may control filtering of a flute player's audio signal through their body movements.

3.2.3 Transparent

One of the issues plaguing some forms of NIMEs, especially augmented instruments, is the network of wire often required. This issue can be mitigated when each eMote wirelessly transmits its own data; localizing only short wires to the sensing surface. The eMotes are designed to be exceptionally small (the largest prototype node is 1 x 2.5 inches in area) to minimize weight and clutter on the instrument. In this manner, the sensing technology is not the object of focus or celebration, but rather the seemingly enchanted 'object' or performer.

3.2.4 Reversible

eMersion should not require any destructive alterations to musical instruments, environments, of human bodies themselves so that the user can easily revert back to the original state when finished with the sensor network. Thus, various methods including "hook-loop" fasteners, weak removable adhesives, and bendable soft materials are being explored for attaching and detaching the eMotes.

3.3 System Overview

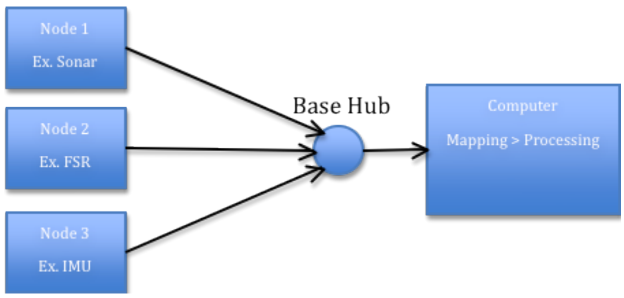

Figure 1: System overview.

eMotes independently transmit data wirelessly to the central hub. Next, the hub parses eMote data streams based on a custom addressing protocol developed by Chester Udell and sends the streams via USB to its attached computer. Then, a software client enables the user to process, mix, and map the data streams to interact with computer processes. The computer software client allows the user to convert sensor data into industry-standard MIDI and OSC protocols, DMX light control, web based UDP, or computer key and mouse commands. This enables the user to interact with virtually any hardware or software regardless of physical location. Each component of the system will be broken down and explained below.

3.3.1 eMotes

Each eMote consists of a single sensor type, an ADC, a processor, flash memory, power source, and a 2.4 GHz wireless RF transceiver. Each node also includes a red and a green status LED to provide the user with visual feedback of the transmission stream. For each data packet the node transmits, the red LED is configured to blink if it is out of range of the receiver or does not get an indication that the receiver is on. The green LED blinks when the sensor node is within range of a receiver properly transmitting data packets.
eMotes transmit data at 50Hz and have a line-of-sight transmit range of about 75 meters. The optimal transmission rate was chosen based on inversely correlational parameters of balance between battery consumption versus reaction time. While the rate is too slow to pick up the fastest of musical motions, like the 2-handed flam (around 200Hz analyzing the peaks in a sound sample), 50Hz is more than adequate at representing the majority of musical performance gestures that we can see with our eyes as they relate to actuating and shaping sound over time. This parameter will become user definable in the future. An eMote lasts more than 24 hours on a single charge and take 1 hour to completely charge from a depleted battery.

3.3.1.1 Data Hub

A single data hub can be used to receive the sensor streams transmitted by the eMotes and sends each stream to the computer. If multiple computers are to have access to the sensor data, a single computer with the hub may act as a networked host and share the data to the other computers using the software client's OSC and UDP features. Multiple data hubs may also be used in the same space on different channels. eMotes can be synchronized to a specific hub or multiple, different hubs. This creates a remarkable variety of possible configurations. The data hub includes an RF transceiver with an antenna, an asynchronous receiver/transmitter (UART), a processor, a channel button, indicator LED, charging ports for eMotes, and a USB interface for connecting to a computer.

4. DIGITAL DATA WORKSTATION

A unique software architecture was created to visualize, process, and map a dynamic and unknown variety of wireless data streams on the network at any given moment. Experimentation with different architectures (including ad hoc, patch bay, and matrix) led me to a layout that most musicians are already familiar with to intuitively visualize, process, and map audio tracks: Digital Audio Workstations, DAWs. Similar to software like Protools, Logic, or Garage Band, a Digital Data Workstation (DDW) layout was conceived for a user to manage data streams and control assignments using data "tracks." Up to the point of this writing, the author believes it is the first time anyone has used this layout as a solution to handle massive and unpredictable amounts of wireless data streams for musical applications. A Digital Data Workstation flows through 3 major processes: input (visualization), processing, output (mapping).

Figure 2: DDW Workflow

Looking at Figure 3 below, a user selects from a list of available data streams on a left-side toolbar. The data is visualized in the center, allowing the user to apply various processing algorithms to the stream including: threshold detection, smoothing, filtering, inversion, beat tracking, and rescaling number range (e.g. translating floating point 0. – 1. data range to 0 – 127 for a MIDI application). The processed data is subsequently assigned to control parameters by choosing from a list of outputs in the right-side toolbar. This method allows the user to experiment with the most optimal and intuitive mapping scheme for a sensor project in real-time. The outputs available for mapping can be virtually any software or hardware device using MIDI, OSC, over network (UDP), and DMX. A special output module called eMotion SwitchBlade emulates keypad and mouse commands, letting the user control virtually any application on a computer beyond the protocols listed above, which might be useful for web applications, video games, and PowerPoint. Custom outputs can be created and automatically populated in the output list using custom eMotion tools in MaxMSP.

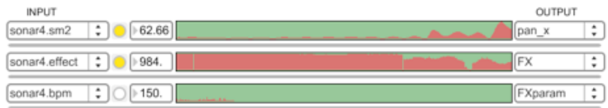

Figure 3: Three DDW data tracks showing input, visualization, and output.

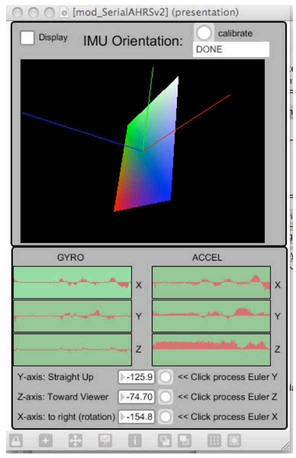

Figure 4: Sensors on the network are automatically visualized in realtime. This shows the visualization of the eMotion Twist, a 360º orientation sensor.

4.1 DDW Implications, Advantages

This architecture, in practice, has proven incredibly effective at providing an intuitive and visually appealing method for visualizing, processing, and mapping unpredictable amounts and types of data streams. While each track may have its own unique sensor input, the ability to create discrete tracks (each with its own internal processing) enables a user to use a single sensor input in many different ways. For example, perhaps a sonar data stream tracking the distance of an object will be the input of several data tracks. The user might run filtering on one track, measure beat tracking on another, and do threshold detection on the third track without having to program custom software. Additionally, each track's output may be uniquely assigned while the input all relates to a single sensor stream, enabling a user to control multiple parameters and/or programs using one axis of control, escaping the common one-to-one paradigm in favor of multi-dimensional control.

Another notable advantage of the DDW architecture is that all settings can be saved. The idea is to have a DDW file for each unique project or piece. This is a powerful concept because once sensors are placed and DDW files are configured, the user could quickly recall unique processing/mapping schemes in realtime (during a live concert or performance). Essentially, the entire behavior of the sensor network could be changed instantly by recalling unique DDW configurations. This, along with the ability to freely add and remove sensors from the network while maintaining the overall integrity of the system, offers an exciting spontaneity to the process of play, experimentation, and exploration in the process of NIME creation and performance.

The final advantage is that the DDW software is not restricted to the eMotion hardware system. Because it's built entirely on OSC at the core, one can use the DDW software for any project requiring the realtime visualization, processing, and mapping of data streams to musical or visual parameters. The implication is that this could be a powerful centerpiece for virtually any data-driven project, NIME-oriented or otherwise.

5. PRESENT STATE, FUTURE DIRECTIONS

15 sets of hardware (1 hub and 2 eMotes) were distributed to talented Beta testers all over the United States and one in Ecuador. The process lasted 4 months and many improvements to the software and hardware were implemented, including USB charging ports and vast improvements in the UI software design.

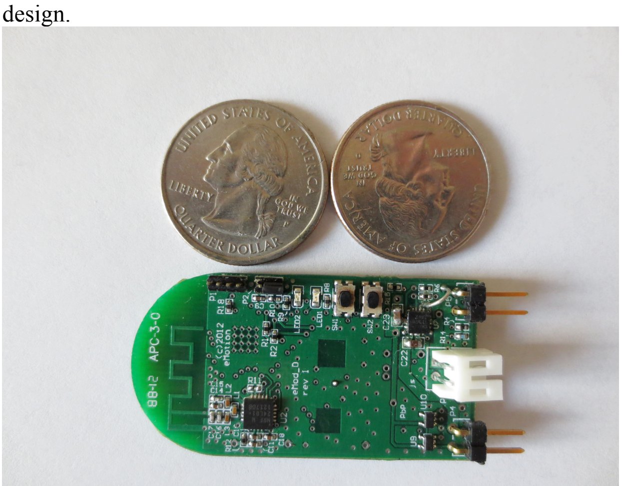

Figure 5: eMotion prototype Twist eMote, measures 360º orientation with 3 axes of accelerometer, gyroscope, and magnetometer. Displayed to scale by two US quarters.

Version 2 systems are currently in production and are expected to be available for general use August of 2014.

There are vast implications for eMotion Tech in the rapid prototyping, robotics, industry, and Arduino communities. To that end, an eMersion Arduino Shield is in development that will enable eMotes to readily communicate with the existing variety of Arduino platforms. Another add-on, the eMersion Portal, will essentially be a wireless receiver module that sends eMote transmissions to any microcontroller using a SPI, I2C (or 2-wire) interface.

In the near future, we hope to apply and evaluate eMersion systems toward developing new STE[A]M (STEM+Arts) curricula for K-12 programs. Given the system's ease of use, it seems well positioned to become a useful tool for providing a hands-on platform for educating teachers and students about using wireless sensing and how we might interact with the digital world around us.

6. ACKNOWLEDGMENTS

Many thanks to James Paul Sain, my dissertation chair throughout the duration of this dissertation [7]. I also wish to extend thanks to engineering mentors Karl Gugel and Antonio A. Arroyo for letting me, a composer, take their advanced EE courses. Special thanks is due to the Arduino and NIME communities for inspiring a generation that takes the evolution of technology into their own hands. And finally for my wife (and all NIME family-partners), who let us glue, drill, and destroy our dining room tables for the sake of building new electronics.

7. REFERENCES

[1] A. Russell, "What Makes a Great Toy for Creative Play?," *MindShift*. [Online]. Available: http://blogs.kqed.org/mindshift/2011/10/what-makes-a-great-toy-for-creative-play/. [Accessed: 30-Dec-2012].

[2] M. M. Wanderley and N. Orio, "Evaluation of Input Devices for Musical Expression: Borrowing Tools from HCI," *Comput. Music J.*, vol. 26, no. 3, pp. 62–76, Nov. 2011.

[3] M. Applebaum, "Culture Sculpture," 2003.

[4] "Teabox," *Electrotap*. [Online]. Available: http://electrotap.com/teabox/. [Accessed: 21-Jan-2014].

[5] D. Topper, "Extended Applications of the Wireless Sensor Array. (WISEAR)," in *NIME 2007*.

[6] R. A. Moog, "Voltage-Controlled Electronic Music Modules," *Audio Eng. Soc. Conv. 16*, vol. 13, no. 3, pp. 200–206, Jul. 1965.

[7] C. Udell, "Towards Intelligent Musical Instruments: New Wireless Modular Gestural Control Interfaces." 2012.

8. Appendix A

U.S. and International patents pending: No. PCT/USUS2012/040723; University of Florida Research Foundation (UFRF) Office of Technology Licensing.