THE MEASUREMENT OF PERFORMER AND AUDIENCE EMOTIONAL STATE AS A NEW MEANS OF COMPUTER MUSIC INTERACTION: A PERFORMANCE CASE STUDY

R. Benjamin Knapp
Sonic Arts Research Centre
Queen's University, Belfast
b.knapp@qub.ac.uk

Eric Lyon
Sonic Arts Research Centre
Queen's University, Belfast
e.lyon@qub.ac.uk

ABSTRACT
This paper describes one of the first computer music pieces to use not just physical gestures, but also physiological indicators of emotion as a means for performer and audience to interact with a computer music composition. The piece, *Stem Cells*, was originally composed for solo laptop performance and served as an excellent vehicle to explore the concept of Integral Music Control (using gesture and emotion to control digital musical instruments) by replacing the laptop with the Biomuse (an instantiation of an integral music controller). Physical gestures of the performer and the emotional changes of both performer and audience created an entirely different means of interaction than simply a solo laptop and consequently, created an entirely new interpretation of the piece. This paper describes how integral music control was used in this reinterpretation and the results of this new performance, both qualitative and quantitative.

1. INTRODUCTION
While research on the introduction of quantitative measurement of emotion as a component of human-computer interaction has been ongoing for many years (a good collection of articles can be found in [1]), the concept of integral music control, the capability to use both physical gesture and emotional state to control digital musical instruments has been around a comparatively short time [2][3][4]. This paper describes one of the first pieces to use not just gesture, but also physiological indicators of emotion as a means for performer and audience to interact with a computer music composition. The piece, *Stem Cells*, was originally composed for solo laptop performance by Eric Lyon and premiered at ARS Electronica in 2005. It served as an excellent vehicle to explore Integral Music Control by replacing the laptop with the Biomuse (an instantiation of an integral music controller). Physical gestures of the performer and the emotional changes of both performer and audience created an entirely different means of interaction than simply a solo laptop and consequently, created an entirely new interpretation of the piece. This paper describes how integral music control was used in this re-interpretation and the results of this new performance, both qualitative and quantitative.

2. REVIEW OF INTEGRAL MUSIC CONTROL
Integral Music Control (IMC) is defined in [3] as “a controller that:
1. Creates a direct interface between emotion and sound production unencumbered by a physical interface.
2. Enables the musician to move between this direct emotional control of sound synthesis and the physical interaction with a traditional acoustic instrument and through all of the possible levels of interaction in between.”

Figure 1 shows the standard technique of controlling sound generation: a thought creates a gesture which then controls a sound generator. Both the sounds and the proprioception from the physical interaction of creating the sound are then sensed by the performer creating a direct feedback loop. The concept of integral music control opens up the possibility for the addition of direct measurement of emotion as another means of interaction.

![Figure 1](415)
As can be seen in Figure 1, even direct measurement of the audience’s emotional state can be used to manipulate sound. The question then becomes, what techniques can be used to directly measure motion and emotion during live musical performance to enable this kind of integral control. There are many techniques for measurement of emotion including visual recognition of facial expression, auditory recognition of speech, and pattern recognition of physiological signals. For most musical performance environments, visual and auditory recognition systems would not be appropriate. Thus, physiological signals are the most robust technique for determining emotional state for direct emotional control of a digital music instrument. Coupled with kinematic sensors such as accelerometers, the responsiveness of physiological sensors to both motion and emotion makes them an ideal candidate that can be used as part of IMC.

3. MEASUREMENT OF GESTURE AND EMOTION: THE BIOMUSE SYSTEM

For the performer, the BioMuse system was chosen to measure both gestural input as well as emotional state. The BioMuse is composed of body worn sensors (both kinematic and physiological) that enable unencumbered movement during live performance. Two Bluetooth transmitters made by Infusion Systems were used and allow for up to sixteen external sensor inputs (two per transmitter). Figure 2 shows the data path from the BioMuse worn on the performer through the wireless link to a PC running the real-time signal processing software, Eyesweb. The processed data is then sent to Max/MSP via Open Sound Control (OSC).

3.1. Gestural Measurement

In re-imagining Stem Cells, it was decided that precision gestures of the arms would be used to explore many of the sonic elements of the piece. To implement gestural control, Electromyogram (EMG) sensors were placed on the front and back of each forearm. In addition a bi-axial accelerometer was also placed on the back of the forearm underneath the arm band of the EMG sensors (see Figure 3). Thus, each arm had a total of 4 degrees of freedom of control: 2 degrees of freedom from the accelerometers (pitch and roll) and 2 degrees of freedom from the EMG (flexion and extension of the forearms). Additionally, the movement of the head was measured using a bi-axial accelerometer worn underneath a headband. This yielded a total of 10 degrees of freedom of gestural control.

Figure 3: Stem Cells Performance. The performer is wearing kinematic and physiological sensors on the forearms, head, and chest.

3.2. Emotional State Measurement - Performer

To measure the emotional state and emotional changes of the performer, various physiological indicators of emotion were used. These signals included:

- Galvanic Skin Response (skin impedance) measured by electrodes on the finger tips
- Electrocardiogram (heart rate and heart rate variability) measured using a chest strap
- Respiration (amplitude and frequency) measured by a chest strap
- Electroencephalogram measured on the occipital (rear) portion of the head using a head band
- Facial EMG measured with sensors built in to the same headband

It should be made quite clear that these physiological indicators are not only measures of emotion. Indeed, as described in [5] there are many reasons other than emotional changes why these physiological signals might vary. However, the primary alternative reasons

![Image](image-url)
for variation such as changes in environment and changes in physical activity do not apply in this performance and consequently the reliability of these physiological signals as an indicator of emotional change is presumed to be high.

3.3. Emotional State Measurement - Audience

Members of the audience were connected to custom circuitry that measured GSR and ECG signals (see Figure 4). These two signals were chosen because of their capability to measure changes in emotional state while still being relatively unobtrusive.

![Figure 4: Sensors worn by audience members measure GSR and ECG. Heart rate and heart rate variability are then derived from the ECG.](image)

As shown in Figure 2, the processed audience emotion signals were then digitized and sent via MIDI to a Macintosh (MAC) computer running Max/MSP. All of the data from all of the audience members were then aggregated and sent via OSC to another Mac running Max/MSP. The reason for this separation of processing tasks into two MACs was due to the DSP requirements of the primary sound synthesis Max/MSP patch. Separating the processing into two MACs also enabled a dedicated GUI to be created for the audience MAC which allowed real-time monitoring of the data and the removal of any noisy data streams. Typically, in any performance there were always one or two audience members that had removed their sensors or in some way manipulated the sensors so that the data was not useable. The audience data processing patch was always monitored by a member of the performance team and a noisy stream could be removed at a moment's notice.

4. SENSOR SIGNAL PROCESSING

As shown in Figure 2, all of the data from the Biomuse system worn by the performer is sent wirelessly to a PC running the Eyesweb signal processing software[6]. The SARC toolbox for Eyesweb, designed for feature extraction and pattern recognition of multimodal signals[7], was then used to process the physiological and kinematic data into specific gestural and emotion features. Some of the signal processing blocks used include heart rate extraction from the ECG, EEG brain wave filtering, and EMG envelope following. An extensive taxonomy of OSC routing messages were used as descriptors to send controller information from the PC running Eyesweb to the primary sound synthesis MAC running Max/MSP. Example messages include /right.forearm.front.env.emf which describes a controller stream which is the envelope of the EMG coming from the right front forearm or /head.roll.raw.acc which describes the raw data coming from the accelerometer worn on the head.

This processed data was then used as controller data for the primary sound synthesis and control Max/MSP patch. Additionally, a set of specific dynamic emotional and gestural piece-specific patterns are recognized within the Max/MSP patch itself. To clarify this concept by way of example: when the piece called for a relaxing breath to trigger a change of section, a "one-shot" object was created that specifically sent a "bang" only on detection of a single deep breath.

One of the critical elements of the Max/MSP patch was calibration of all of the different data streams. All of the gestures and physiological indicators of emotion had to be calibrated specifically for the performer. Most difficult was the calibration of emotional change which required several performances of the piece to optimize. However, one of the surprising outcomes of the work on this piece was that once the calibration for the performer was completed, it did not need to be recalibrated again. Thus, no calibration was needed before a performance.

While calibration for audience physiological sensors also did not need to be performed for each performance, it took quite a while to determine optimal GSR/ECG sensor mappings. Only when the piece was performed in its entirety in front of an audience could the audience sensors be calibrated. The first time the re-imagined Stem Cells was performed in front of an audience, the audience data completely saturated the mappings. Only after several performances was the data range understood (although it is still continually being adjusted for optimal mappings in relation to the piece itself).

5. RE-IMAGINING THE PIECE

5.1. The Performance

As previously mentioned, Stem Cells was composed by as a solo laptop piece and was performed at several venues over four years including Ars Electronica and Huddersfield. The original version of the piece defined a large scale form, divided into individual sections that were articulated as Max/MSP patches. The central compositional idea was to start with simple sonic materials and gradually enrich them through iterative processes. In this arrangement for IMC, the goal was to explore the possibility of performing Stem Cells in a new way, transmuting movement, but also emotional states into musical control. It was decided that the piece would be performed much the same as Alvin Lucier's Music for Solo Performer. A solo performer would sit on a chair (see Figure 3) with no other instruments visible and only a computer monitor sitting on the floor in front. The
piece has been performed both with and without a projection of the various sensor signals behind the performer depending on the venue and the desired aesthetic of the performance.

5.2. The Structure of the Piece

It was conceived that the transition between each section of the piece would be initiated by the evolution of emotional state of the performer. Within each section, many options would be available to the performer who would choose among the possibilities and, using physical gestures, performs them with a considerable degree of interpretive freedom.

As shown in Figure 5, in its final version, there are seven sections of the piece (and therefore, the Max/MSP patch). The initiation of the piece begins by the performer changing emotion from serenity to anger. As this occurs and is recognized by the pattern recognition software, the sound becomes ever more complex. Once achieving the emotional state, a sequence of increasingly angry gestures are performed. After these, the performer takes a deep breath and relaxes and the piece moves into the second section.

In the second section, melodic sequences, filters, etc are all controlled by arm gestures both large and subtle. The performer can explore this sonic space for as long as wished. In this section, the difference from the original laptop performance is twofold:

1) Unlike the laptop performance, the gestural performance is visible to the audience and cause and effect are clear. There is thus a natural connection between performer and audience.

2) As was described previously, the gestural performer can control 10 degrees of freedom with ease. The laptop performer has only a mouse and keyboard where it is quite difficult to control more than the two dimensions of a mouse and the button clicks simultaneously.

To move to section 3, the performer must again relax and take a deep breath. The performer then uses gestures to sequentially explore different scales. To traverse into the fourth section, a sudden movement of the head to the right is used. This causes a voice loop to emanate from the right speaker. This gesture makes it appear as if the performer is reacting to a sudden voice (not the other way around). This sequence is repeated for the left speaker and many voice loops are now heard. Then, a gesture toward the audience is made and the emotional state of each audience member is assigned a voice loop. The performer stands up and then walks out among the audience. This causes the audience members to suddenly become nervous. As their emotion changes, their voice loop increases in speed and intensity. At this point true co-creation between performer and audience is achieved.

Figure 5: The Stem Cells Max/MSP Patch. There are seven sections of the piece. Four of the sections allow for exploration of the sound space using physical gestures. Three of the sections are a progression that advances as the prescribed emotional state progresses. The transition between each section is based on a change of emotional state as well. (See text)
The performer returns to the seat on the stage and begins to relax and the piece traverses into the fifth section. In this section, the performer slowly becomes happy and in so doing causes the sound to evolve and a transition into the sixth section. This state of happiness requires visualization by the performer. As with the other emotionally controlled sections of the piece, the danger is if this emotional change does not happen, the piece stalls.

The sixth section of the piece is again controlled by arm gestures. The transition and flow through the final section of the piece is in anti-symmetry to the beginning of the piece. The performer slowly transitions from an agitated state (naturally caused by the previous section) to a final state of complete relaxation.

6. PERFORMANCE EVALUATION
Before re-creating Stem Cells for the Biomuse, it was hypothesized that that using integral music control would enable:

1) New dimensions of control that could not be achieved using the laptop.

This has clearly been this case. As described previously, the performer has control over 10 degrees of freedom of physical gesture as well as the ability to visualize emotional states as a means of controlling the piece. Whether these degrees of control are "superior" is not an issue. The fact that this created a new and compelling performance of the piece is what is important.

2) A more "visible" instantiation of the piece. That is, that a connection between performer and audience could be created through watching the gestures and changes of emotion of the performer.

Stem Cells has been performed many times throughout the world including New York, Sydney, England, and Ireland. In each of the performances, the audience was asked if they understood the relationship between the physical gestures being performed and the sound created. In all of the performances, the overwhelming answer was yes. When asked if they understood that some changes in emotion of the performer were used to interact with the piece, not everyone did. However, many of those that did understand the relationship spoke of a profound empathy with the performer. To test whether this empathy or even "emotional contagion" could be quantified, GSR and heart rate data was compared between performer and audience. As shown in the example of GSR in Figure 6 there was a strong correlation between some audience member's physiological indicators of emotion and the performer's. On-going data collection will show how universally true this is across all audience members across all performances.

![Figure 6: GSR of performer (Ben) vs. GSR of one audience member during a 5 minute segment of the performance of Stem Cells. Note both short and longer term correlations in the data.](image-url)
7. CONCLUSION
In this paper, we have described the adaptation of a computer music piece, originally designed for laptop performance, to integral music control. The piece explored the continuum between controlling a digital musical instrument using physical gestures and controlling a digital musical instrument using the changes of emotional state of performer and audience. In so doing, the concept of integral music control was brought to life, co-creation between performer and audience was realized, and a new expressive technique for musical performance was added to the spectrum of possibilities available to composers and performers.

8. REFERENCES