

Creating a Network of *Integral Music Controllers*

R. Benjamin Knapp
BioControl Systems, LLC
Sebastopol, CA 95472
+001-415-602-9506
knapp@biocontrol.com

Perry R. Cook
Princeton University
Computer Science (also Music)
Princeton, NJ 08540
+001-609-258-4951
prc@cs.princeton.edu

ABSTRACT

In this paper, we describe the networking of multiple Integral Music Controllers (IMCs) to enable an entirely new method for creating music by tapping into the composite gestures and emotions of not just one, but many performers. The concept and operation of an IMC is reviewed as well as its use in a network of IMC controllers. We then introduce a new technique of Integral Music Control by assessing the composite gesture(s) and emotion(s) of a group of performers through the use of a wireless mesh network. The Telemuse, an IMC designed precisely for this kind of performance, is described and its use in a new musical performance project under development by the authors is discussed.

Keywords

Integral Music Control, Musical Control Networks, Physiological Interface, Emotion and Gesture Recognition

1. INTRODUCTION

The Integral Music Controller (IMC) [1] is defined as a controller that:

1. Creates a direct interface between emotion and sound production unencumbered by the 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.

This paper describes the networking of multiple IMC's, to enable not just one, but many performers to use an IMC and to interact with each other in three ways:

1. The "normal" perceptual path – the performers see, hear, and sometimes even haptically feel the other performer. .
2. The controller interaction path – the performers physical gestures and emotional state, as assessed by the IMC, are used to another performer's electro-acoustic instrument.
3. The integral control path – an entirely new path whereby the emotions or gestures of one performer, as measured by the IMC, are combined with the emotions and gestures of other performers to create an assessment of group gestures and emotions and this is used to control music creation.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

NIME 06, June 4-8, 2006, Paris, France.
Copyright remains with the author(s).

2. REVIEW OF INTEGRAL MUSIC CONTROL (from [1])

The term "integral" in "integral music controller" refers to the integration into one controller of the pyramid of interface possibilities as shown in Figure 1. Using an IMC, a performer can move up and down through the interface possibilities.

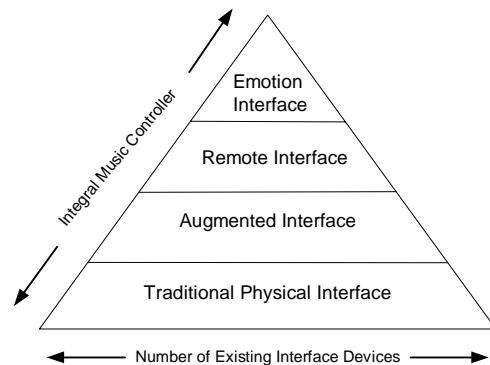


Figure 1: Pyramid of interfaces for controlling a digital musical instrument (categories loosely adapted from [2]). Note the decreasing number of existing interface devices as you move up the pyramid. The integral music controller (IMC) has elements of all interfaces.

As shown in Figure 1, the introduction of direct measurement of emotion to digital musical instrument control represents the completing of the pyramid of possible interfaces. Only with a direct interface to emotion is a truly integral controller possible.

The use of a direct emotional interface also introduces one new feedback path in a musical performance that was never before possible. Figure 2 shows three layers of feedback that can be achieved in musical performance. Layer 1 is the emotional layer. The emotional state of the performer initiates and adjusts the physical gesture being made. This emotional state might or might not be reflective of the intention of the performer. Also, the perception of the sound that is created from the physical gesture elicits an emotional response in the performer and, based on this; the performer may alter the physical gesture. Layer 2 is the physical interface layer. Feedback is achieved through visual cues and proprioception [3]. Layer 3 is the sound generation layer. The physical gestures cause a sound to be created which is heard and possibly used by the performer to adjust the physical gesture [4]. The introduction of a direct emotional interface means that a performer's emotions will directly control the sound generation without passing through the physical interface. The sounds created will effect the emotion of the performer [5] and thus a new feedback path is created.

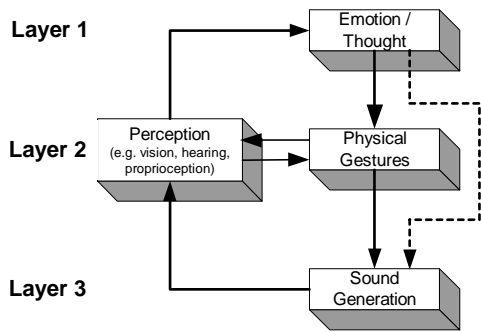


Figure 2: The three layers of performance feedback using an IMC. Layer 1 represents the internal emotion and thoughts of the performer. Layer 2 is the physical interface layer. Layer 3 represents the consequence of the gesture - the creation of music.

There is an extensive body of literature on defining, measuring and using emotion as a part of human computer interaction and “affective” computing (see [6][7][8][9] for a good overview). The emotional reaction to music is so strong that music is commonly used as the stimulus in emotion research [10]. The understanding of the emotional reaction to music, not the categorization or labeling, is critical in using emotion as a direct performance interface. It is clear [3] that this emotional reaction is highly individualistic and thus any synthesis model that uses emotion as an input must have the capability of being customized to an individual performer.

There are many techniques [9] 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 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. Physiological signals have been used many times as a technique of human computer interaction in music [11][12][13] for example). Their responsiveness to both motion and emotion makes them an ideal class of signals that can be used as part of an IMC.

3. THE NETWORKED CONTROLLER

The inclusion of networked interaction in electro-acoustic instrument performance introduces a new path for performers to communicate. Networked music controllers can be thought of as a subset of multi-user instruments (see [14] for a summary of such instruments). There are numerous examples of networked controllers used in performance including the MIT Media Lab’s Brain Opera [15] and Toy Symphony [16]. In the latter, the BeatBugs [17] allowed the players to enter musical material, then play it or modify it by manipulating sensors on the bug, and/or pass it to another player by pointing the bug at them. A subset of networked controllers is so-called “wearables” and include networked jewelry and clothing [28].

The Princeton Laptop Orchestra (PLOrk) [18] is a recent experiment in constructing an orchestra of sensor connected laptops and speakers. Various composing / performing / conducting paradigms have been investigated, including passing synchronization and other messages related to timbre, texture, etc. over 802.11G using Open Sound Control (OSC). The language Chuck [19] is one of the primary programming mechanisms used by PLOrk, as there is a rich provision for

low-latency (10-20 ms.) asynchronous messaging built into the language.

Figure 3 is a block diagram of networked IMC’s showing the networked interaction path separated into a physical gesture path and an emotion path. (Note the already existing perceptual path which symbolizes the performers’ ability to see, hear, and even feel each others performance.) These new networked interaction paths create a way for performers to collaborate with each other at the controller level before the sounds are actually created. Each performer’s physical gesture or emotional state is recognized and converted into a control parameter(s) that can be combined with the control parameter(s) of other performers to create a rich and complex means for group performance.

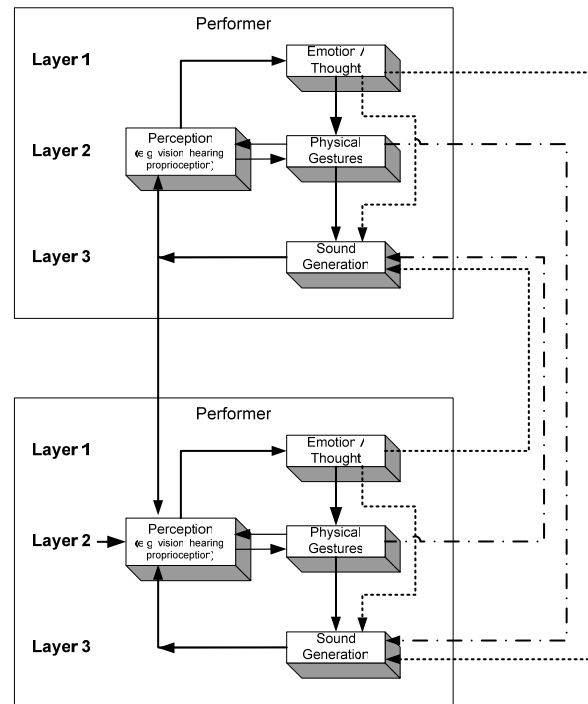


Figure 3: The networking of multiple IMC’s. The solid line between each performer represents at the perceptual level. The dashed-dot line shows the physical interaction at the controller level, i.e., how the physical gesture of one performer can effect the sound generation of another performer’s instrument. The dotted line shows the emotional interaction at the controller level, i.e., how the emotion of one performer can effect the sound generation of another performer’s instrument.

4. THE INTEGRAL CONTROL PATH

Unlike standard networked instruments, the integral control path seeks to combine the physical gestures and emotional state of multiple performers *before* they are categorized and processed into control parameters. The purpose of this is to assess a composite emotion or gesture of multiple performers first, and then to use this as a control input. As shown in Figure 4 this requires a mesh computation of composite signals. Only a completely self-forming, self-aware mesh network topology would enable sets and subsets of different performers to interact with sets and subsets of instruments in real-time.

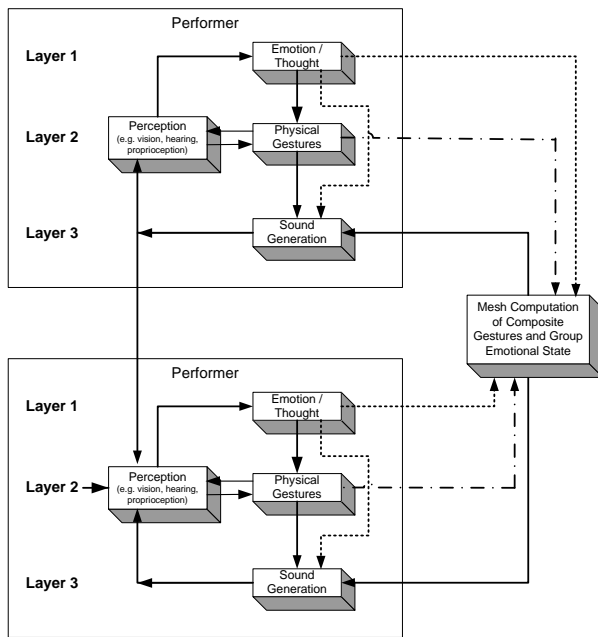


Figure 4: The networking of multiple IMC's using an integral control path as part of a mesh. In this mesh, any performer's physical gestures and emotional state can be composited with any other's.

Both forms of networking can be combined to create a network of integrally networked IMC's. Thus, for example, a performer's emotional state can be assessed by the IMC, combined to create with other performer(s) to create an overall combined emotional state, this state can be used to control the output of a controller within a network of controllers. A detailed example of this will be discussed in section 6 of this paper.

5. IMPLEMENTATION: THE TELEMUSE

There are many systems that wirelessly transmit physiological data including the BodyMedia's SenseWear [20], NASA's LifeGuard[21], and MIT's LiveNet[22]. There are several sensor systems that use wireless mesh networking, and more specifically, a network layer protocol known as ZigBee™. ZigBee™ is designed to use the IEEE 802.15.4 standard, a specification for a cost-effective, relatively low data rate (<250 kbps), 2.4 GHz or 868/928 MHz wireless technology designed for personal-area and device-to-device wireless networking [23]. There are several companies that have ZigBee™-based sensor units including those made by Crossbow [24], Dust [25], and MoteIV [26]. Harvard's CodeBlue [27] uses Crossbow's ZigBee™ compliant motes to create a mesh network of physiological sensors. None of these interfaces are designed specifically as human-computer interfaces, let alone musical instrument controllers, and therefore none of the designers incorporated the use of an integral control path.

The TeleMuse system shown in Figure 5 integrates physiological signal sensors, motion sensors, and a ZigBee™ wireless transceiver into one band designed for human-computer interaction and music control. The TeleMuse can be worn:

- on the limbs to measure muscle tension (EMG), Galvanic Skin Response and motion (dual axis accelerometers)

- on the head to measure brain activity (EEG), muscle tension (EMG), eye motion, and head motion (dual axis accelerometers)
- on the chest to measure heart activity (EKG) and respiration

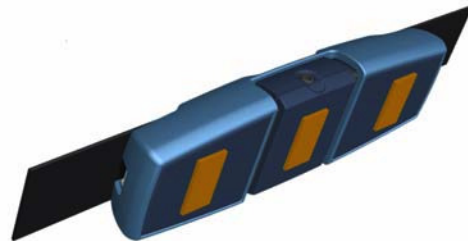


Figure 5: The TeleMuse Wireless Mesh Network IMC

The TeleMuse is the next generation of Integral Music Controller replacing the Wireless Physiological Monitor (WPM) [29] in a smaller more ergonomic design. Like the WPM, the TeleMuse uses dry electrodes to sense physiological data. Unlike the WPM, each TeleMuse is its own node in a mesh network and can communicate with any other node in the network. Computation of physical gestures and emotional state, based on physiological signals and accelerometer data, can be distributed among any of the nodes and any computers on the network.

6. VACHORALE: A PIECE FOR PLOK, TELEMUSE AND SINGERS

One use of networked IMC's will be investigated in a project entitled the Virtual/Augmented Chorale (VACHorale). The Virtual/Augmented Chorale (VACHorale) project will investigate the compositional and performance opportunities of a "cyber extended vocal ensemble" and will use the Princeton Laptop Orchestra (PLOrk), the TeleMuse, and ChuckK.

6.1 Augmenting The Singer

The VACHorale project will outfit a small choir of (eight) singers with several Telemuses and microphones, coupling each human singer to a laptop, multi-channel sound interface, and multi-channel hemispherical speaker. As an obvious first step, the system will use digital signal processing to modify and augment the acoustical sound of the singers. Further, we will use networked Telemuses to control various algorithms for modifying and extending the choral sound. The most revolutionary component will be using the TeleMuse to control various sound (primarily voice/singing) synthesis algorithms, in order to extend, and even replace the acoustic components of the choir. The singers will thus be able to "sing" without phonating, controlling the virtual choir with facial gestures, head position, breathing, heart rate, and other non-acoustic signals. An assessment of each singer's emotional state, as well as the choir's composite emotional state will be used as well.

We plan to fully realize the IMC concept, with the physical gestural "instrument" being a singer, and we will create an ensemble of multiple IMC-outfitted Virtual/Augmented singers. The continuum from the "dry" choral sound, through the digitally augmented acoustic sounds

of the singers, to the completely virtual sound of the biological sensor-controlled synthesized singing, will provide a rich compositional and performance space in which to create new music.

6.2 Building The “Instruments”

The first goal of the project is to integrate the existing hardware and software systems for biological signal acquisition and processing, acoustical signal processing, and voice synthesis, with the PLOrk (Princeton Laptop Orchestra) “workstations” to create a new augmented singer “instrument.” These instruments, hereafter called VACHS (Virtual/Augmented Choral Singers, pronounced “vax”) will be identical in technical capability, but can take on various forms based on configuration, programming, and control.

First, the PLOrkStations provide the basic computational and acoustical technical capabilities. Built with support from the Princeton University Council on Science and Technology, the Princeton Freshman Seminar Program, the Princeton departments of Music and Computer Science, the Princeton School of Engineering and Applied Science, and Apple Computer, each of the 15 existing workstations consists of a 12” Mac Powerbook, an Edirol multi-channel FireWire digital audio interface box, six channels of amplification, and a custom-built six-discrete-channel hemispherical speaker.

Second, the TeleMuse will couple each singer in the ensemble to a networked hardware workstation. Physiological signals will be captured and processed by each TelMuse node and shared with the rest of the mesh network. As mentioned previously, these signals can be used to determine not only singing gestures, but also the emotional state of the performers. Additionally, each box contains a two-axis accelerometer, so head/body tilt and orientation can be measured. Additional sensors can be used to measure absolute body and head position and orientation.

ChucK was specifically designed to allow rapid, on-the-fly audio and music programming and will be used to synchronize the multiple composited controller streams.

6.3 Virtualizing the Singer

The physiologically-derived emotion signals of the IMC can be mapped to signal processing such as adding echoes and reverberation, shifting pitch, controlling spatial position, etc., and compositional processes such as note generation and accompaniment algorithms. But the IMC can also be mapped to the parameters of physical synthesis models, creating a truly integral controller. In fact, indirect emotion mapping already exists in many acoustic instruments. The nervousness of a singer or violin player already shows in the pitch jitter and spectral shimmer of the acoustical instrument. The heartbeat of the singer modulates the voice pitch because of modulation of lung pressure.

Synthesis by physical modelling lends naturally to control from physical gestural parameters. Signals such as those that come from an IMC can easily be detected and mapped to similar, or totally different (brightness, spatial position, etc) parameters in a physical synthesis model. With higher-level control and player-modeling inside the model, emotional parameters might make even more sense than raw gestural ones. A large variety of parametric physical instrument synthesis models exist in ChucK, with many holding much promise for control from singer gestures and emotional parameters. The models that hold the most interest for this

project, however, are those that mimic the human singing voice.

Older proven models for voice synthesis, such as formant filter synthesizers, and articulatory acoustic tube models, already exist in ChucK as native “unit generators.” As such it will be easy to perform a number of different mapping experiments, and produce a variety of human-like (and quite inhuman) sounds based on control from the singer sensors.

New models of the human voice such as Yamaha’s Vocoloid (constructed with UPF Barcelona) allow for control of vocal quality parameters such as growl, breathiness, and raspiness, and more semantic qualities such as “bluesyness” and “sultriness”. These also seem completely natural for control by emotional parameters, and will be exploited in the Virtual/Augmented Chorale project.

6.4 The Performance

The goal of the Virtual/Augmented Chorale project is to compose and rehearse a number of choral pieces, aimed at the production of several concert performances. The repertoire will range from traditional early music augmented by the virtual acoustics of the VACHS, through some contemporary a capella vocal literature, but with the human ensemble augmented by virtual singers, and one or two brand new pieces composed specifically to exploit the maximum capabilities of the Virtual/Augmented Chorale.

7. REFERENCES

- [1] Knapp, R.B. and Cook, P.R., The Integral Music Controller: Introducing a Direct Emotional Interface to Gestural Control of Sound Synthesis, In *Proceedings of the International Computer Music Conference (ICMC 2005)*, Barcelona, Spain, September 5-9, 2005.
- [2] Wanderley, M.M. “Gestural Control of Music”, IRCAM – Centre Pompidou, 2000.
- [3] Askenfelt, A. and Jansson, E.V. “On Vibration Sensation and Finger Touch in Stringed Instrument Playing,” *Music Perception* 9(3), 1992, pp. 311-350.
- [4] P. Cook, “Hearing, Feeling, and Performing: Masking Studies with Trombone Players,” *International Conference on Music Perception and Cognition*, Montreal, 1996.
- [5] Panksepp, J. and Bernatzky, G. “Emotional Sounds and the Brain: the Neuro-affective Foundations of Musical Appreciation,” *Behavioural Processes*, 60, 2002, pp. 133-155.
- [6] Holland NN, “The Brain and the Book, Seminar 7: Emotion”, <http://web.clas.ufl.edu/users/nnh/sem04/memo-s04.htm>, February 9, 2004.
- [7] Hudlicka, E. “To Feel of not to Feel: The Role of Affect in Human-Computer Interaction,” *International Journal of Human-Computer Studies*, 59, 1-32, 2003.
- [8] Picard, R.W., Vyzas, E., and Healey, J. “Toward Machine Emotional Intelligence: Analysis of Affective Physiological State,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23, 10, October 2001, pp. 1175-1191.
- [9] Nasoz, F., Alvarez, K., Lisetti, C. L., & Finkelstein, N. “Emotion Recognition from Physiological Signals Using Wireless Sensors for Presence Technologies,” *Cogn. Tech. Work*, 6, 2004, pp. 4-14.

- [10] Steinberg, R. (Ed.). *Music and the Mind Machine*, Springer, Berlin, 1995.
- [11] Tanaka, A and Knapp, R.B., "Multimodal Interaction in Music Using the Electromyogram and Relative Position Sensing," *Proceedings of the New Interfaces for Musical Expression (NIME) Conference, Media Lab Europe, Dublin, Ireland, 2002*.
- [12] Knapp, R. B. & Lusted, H.S. "A Bioelectric Controller for Computer Music Applications," *Computer Music Journal*, 14 (1), 1990, pp. 42-47.
- [13] T. Marin Nakra, "Inside the Conductor's Jacket: Analysis, Interpretation and Musical Synthesis of Expressive Gesture," PhD Dissertation, MIT Media Lab, 2000.
- [14] Jorda, S. "Multi-user Instruments: Models, Examples and Promises," *Proceedings of the 2005 International Conference on New Interfaces for Musical Expression (NIME05)*, Vancouver, BC, Canada, pp. 23-26.
- [15] Joseph A. Paradiso, "The Brain Opera Technology: New Instruments and Gestural Sensors for Musical Interaction and Performance", *Journal of New Music Research*, Vol. 28, No. 2, pp. 130-149.
- [16] <http://www.toysymphony.org>
- [17] Aimi, R. and Young, D. "A New Beatbug: Revisions, Simplifications, and New Directions," *Proceedings of the International Computer Music Conference (ICMC 2004)*, Miami, Florida, November 1-6, 2004.
- [18] <http://plork.cs.princeton.edu>
- [19] Wang, G. and P. R. Cook. "ChucK: A Programming Language for On-the-fly, Real-time Audio Synthesis and Multimedia," *Proceedings of ACM Multimedia 2004*, New York, NY, October 2004.
- [20] <http://www.bodymedia.com/index.jsp>
- [21] <http://lifeguard.stanford.edu>
- [22] Sung, M., Marci, C., and Pentland, A., "Wearable Feedback Systems for Rehabilitation," *Journal of NeuroEngineering and Rehabilitation* 2005, 2:17.
- [23] <http://www.zigbee.org/>
- [24] <http://www.xbow.com/>
- [25] <http://www.dust-inc.com/>
- [26] <http://www.moteiv.com/>
- [27] Lorincz, K, et.al. "Sensor Networks for Emergency Response: Challenges and Opportunities," *IEEE Pervasive Computing*, October-December, 2004, pp 16-23.
- [28] http://www.turbulence.org/blog/archives/cat_wearables.html
- [29] Knapp, R.B. and Lusted, H.S. "Designing a Biocontrol Interface for Commercial and Consumer Mobile Applications: Effective Control within Ergonomic and Usability Constraints," *Proceedings of HCI International*, Las Vegas, NV, July 22-27, 2005.