Music for Flesh II: informing interactive music performance with the viscerality of the body system

Marco Donnarumma
Sound Design, ACE, The University of Edinburgh
Alison House, Nicolson Square
Edinburgh, Scotland, UK, EH8 9DF
m.donnarumma@sms.ed.ac.uk
m@marcodonnarumma.com

ABSTRACT

Performing music with a computer and loudspeakers represents always a challenge. The lack of a traditional instrument requires the performer to study idiomatic strategies by which musicianship becomes apparent. On the other hand, the audience needs to decode those strategies, so to achieve an understanding and appreciation of the music being played. The issue is particularly relevant to the performance of music that results from the mediation between biological signals of the human body and physical performance.

The present article tackles this concern by demonstrating a new model of musical performance; what I define *biophysical music*. This is music generated and played in real time by amplifying and processing the acoustic sound of a performer's muscle contractions. The model relies on an original and open source technology made of custom biosensors and a related software framework. The succesfull application of these tools is discussed in the practical context of a solo piece for sensors, laptop and loudspeakers. Eventually, the compositional strategies that characterize the piece are discussed along with a systematic description of the relevant mapping techniques and their sonic outcome.

Keywords

Muscle sounds, biophysical music, augmented body, realtime performance, human-computer interaction, embodiment.

1. INTRODUCTION

Biosensing musical technologies make use of biological signals of the human body to produce and control music. Such technologies are usually based on a medical technique called biofeedback. This technique is aimed at revealing inner, physiological processes of the body of which we are partly or completely unaware, by transforming a biological signal in something perceivable to our senses.

The first musical applications of biofeedback, later on named *biomusic*, date back to the 1960s and 1970s. In *Music for Solo Performer* and *Brainwaves Music*, Alvin Lucier and David Rosenboom [11] demonstrated the feasibility of this application in a performative and compositional context by sonifying the electrical impulses of their brainwaves. About 20 years after, the interest in such a corporeal approach to music performance rose again thanks to projects such as Knapp and Lusted's BioMuse [7], Van Raalte's Bodysynth [13], and Miranda's Brain-Computer Musical Interface [8]. The former has possibly become the most well known biomusic interface today, for several pieces have been composed for and performed with it by a number of artists, among which the same Knapp with Eric Lyon; the media and sound artist Atau Tanaka; and researchers and players such as Miguel Ortiz Perez and Sarah Nicolls.

The biological signals used by the projects above are the electrical impulses produced by physiological phenomena; such as the brainwaves (electroencephalography or EEG), the muscles tension (electromyography or EMG) and the heartbeat (heart rate monitoring, electrocardiography or ECG). Whereas Knapp, Lyon and Perez have focused their work with the BioMuse on the tracking of emotion arousal through EEG and ECG measurements, Tanaka and Nicolls, respectively a performer and a pianist, have experimented with muscle tension applications. These two approaches have produced very different outcomes. In the piece *Stem Cells* for instance, Knapp sits mostly still on a chair and garner control over synthesised music by voluntarily changing physiological states; the emotional feel diffused in the concert room seems tangible, yet the performer's effort and agency is difficult to decipher and understand, because his body does not offer visual cues. On the contrary in Tanaka's solo performances, as well as in the related work with the seminal Sensorband [2], the audience can clearly perceive and directly experience the effort of the player. This becomes of crucial importance both technically, for the player controls sound through physical muscle tension, and visually, as the viewers can appreciate a spontaneous dramaturgy of the player's gesture. Whereas installation and performances using emotion arousal are still proliferating today, the development of a bodily approach to bio-sensor music performance has suffered an apparent stand-by.

Building on the previous research, I wish to frame biosensor instruments in a "more musically interesting and contemporary context"[1]: one in which the body is not only a controller, but an actual sound generating force, for it produces the sonic material of a piece of music. Such paradigm exemplifies the idea of a new musical instrument that does not depend on quantitative analysis of the biological body, but rather on its innate expressive qualities. A musical instrument that is not designed *around* the human body, but explicitly *for* the human body.

[1]This is borrowed from a comment on the XS by STEIM's Artistic Director Dj Sniff, following my interview on the magazine Create Digital Music. For further reference see `http://marcodonnarumma.com/bio/#press`

NIME'12, May 21 – 23, 2012, University of Michigan, Ann Arbor.

2. FROM FLESH TO MUSIC

Excited muscle tissues produce an acoustic vibration, called mechanomyogram signal or MMG. This phenomenon can be observed on the surface of a muscle when it is contracted. At the onset of muscle contraction, significant changes in the muscle shape produce a large peak in the MMG. The oscillations of the muscle's fibres at the resonant frequency of the muscle generate subsequent vibrations. Several times in this article I refer to the MMG as a *sound*. Even though some may argue for a different interpretation of what a sound is, I find natural to use the term in this context, as de facto, the MMG produced by the muscle is an acoustic oscillation. As such, it can be amplified and heard through headphones or loudspeakers. Figure 1 shows the MMG sound wave of a sustained contraction captured via the XS. The sound sample is available on-line for reference [2].

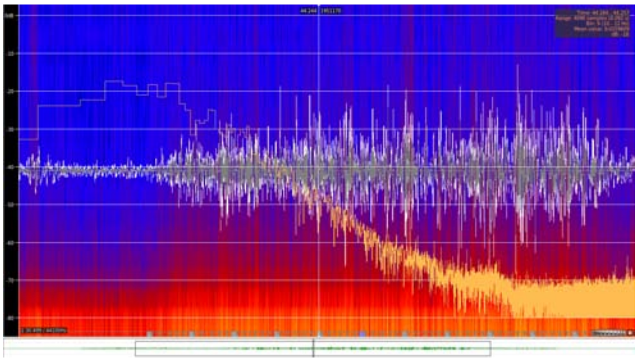

Figure 1: MMG of a sustained contraction: spectrogram (in the background); waveform (white in the foreground); and logarithmic spectrum (yellow outline)

Even though the systematic study of muscular sounds began around the 1980s [10], so far the research has been applied exclusively to biomedical engineering [12]. Yet, this little and detailed, vibrational force can be captured and computationally enhanced, so to provide a rich and dynamic bundle of musical material and control data.

2.1 Biophysical music principle

The Xth Sense (XS) is a novel biophysical system for interactive music performance and responsive milieux; it is composed of custom wearable sensors (Figure 2) and an ad hoc computational engine. The hardware and its documentation are licensed under a Creative Commons Attribution Share-Alike license, and the XS code is released as free software under a GPL license. At the best of the author's knowledge, the XS is the first open musical instrument deploying muscle sounds as musical material for an interactive system.

The XS biosensors are wearable devices that can be placed on any muscle of the performer's body; the hardware is composed of custom microphone sensors that capture the MMG sound (between 20Hz and 45Hz), with no direct contact with the skin. As opposed to bioelectric controllers (that deploy EMG signals), the XS depends on a microphone that picks up subcutaneous mechanical vibrations, or better, *sounds* that originate within the muscle fibres. The XS uses these sonic vibrations as control data and audio signal to be processed according to the same data stream. The performer controls the live sampling and spatialization of the muscle sounds, which the computer diffuses through the loudspeakers. At this point, the player perceives the surfacing sonic space, and shapes it by exerting further contractions. The creative feedback loop between her physio-somatic behaviour and the computer circuitry is closed. This is the principle that underpins the *biophysical music* paradigm.

However, the production and performance of biophysical music relies also on the design of specific compositional strategies and mapping techniques. These are discussed in the following sections within the practical context of Music for Flesh II (MFII), the first performance work for the XS[3].

For the interested reader and the perspective biophysical music performer, I would also recommend the reading of my detailed report on the technical implementation and design of the XS sensors [5], and the in-depth description of the XS computational engine [4].

3. COMPOSITIONAL STRATEGIES

Since its inception in mid 2011 MFII has toured academic conferences and specialized festivals in USA, Mexico, South Korea, Portugal, Norway, Ireland, UK, Italy and Germany.

In this work the XS is used to compose the acoustic clusters released by my muscles into music in real-time. MMG signals are continuously captured by a pair of XS wearable microphones located on my forearms and analysed by a computer to extract meaningful features. According to this data stream, the MMG sound is live sampled, processed and eventually played back through a variable array of loudspeakers.

The signal analysis and processing operated by the XS software is designed to seamlessly enhance the inherit interactions that bonds the kinetic body and the outwarding sonic form. By nature, a sudden and strong flection/extension of the limb naturally produces a loud sound with a sharp attack and very short release. Strength of the contraction and perceived loudness of the MMG are tightly related, therefore a specific mapping technique can extend that relationship by adding multiple dimension to it. The dynamics of each MMG sound is used as a continuous event to manipulate the qualities of the outwarding sound.

Figure 2: The XS biosensors

In order to ensure a fair amount of complexity and richness, up to 8 simultaneous sampling dimensions are available to the player. In this way the interrelation of agency, musicianship and musicality remains transparent throughout the piece. I have previously elaborated on the importance of such interdependency in the context of biotechnological performance in [6]. A link to an audiovisual recording of this work is provided for reference[4].

[2] http://marcodonnarumma.com/publications/media/xth-sense_mmg-arm_lifting-weight.wav

[3] Another XS based work have been produced, while collaborative projects with dancers and traditionally trained players are in the making. See: http://marcodonnarumma.com/works/hypo-chrysos

[4] Music for Flesh II, interactive music performance for enhanced body: http://vimeo.com/20889787

3.1 Encoding expressivity: features extraction

The computer learns about the emergent physiology of my body by extracting discrete and continuous features from the MMG signal. Each sensor produces one analog signal output; this is digitalized and passed through an array of algorithmic functions designed to meaningfully shape the incoming biosignal into diverse control features, namely: Natural (N), Soft (S), Linear (L), Tanh (T) and Maximum Running Average (MRA). The next section dissects the technical processes behind the features extraction and explains some of the modalities by which each feature is used to produce a specific musical result.

3.1.1 Methods and aesthetic

The N value is computed in two steps. First, the initial data that describe the dynamics of the muscular motion are generated; the XS software tracks the root mean square (RMS) of the MMG output using a Hanning window of 512 samples. This could be thought of a useful event on its own, however, my experience with the system showed that it can prove far too linear to constitute an expressive and detailed control event. Therefore, the RMS is passed through a custom function so to sculpt a subsequent continuous event that imitates the elastic, and sometimes jittery contraction of the muscular tissues. This can be compared to the bending of a rubber band: when the muscle is flat, N is equal to 0; at the onset of the first contraction N increases proportionally to the amount of energy released. However, when the contraction ceases N does not fall back immediately to 0, but it bounces back and forth as the muscle tissue recovers its static position. Such behaviour is quite interesting, for it causes control data to be produced also after a gesture is completed. In MFII I use this method to involuntary excite the machine processor, and so provoke aural *echoes* of my gestures[5]. From the audience perspective this represents a rupture of the direct interaction between performer and machine. Nonetheless, rather than contributing a negative *feel*, such disruption unveils a real and unpredictable dialogue between the player and the computer, which prompts the spectators to immerse more mindfully within the performance imagery. The gestures becomes better recognizable as the echoes widen the auditory space around it.

The S feature is obtained by passing N through a single exponential smoothing (SES) function [9], which results in a much softer continuous event. S is used to drive changes in time and room size of a series of reverb effects, located at the end of the processing chains, provoking subtle textural permutations.

L is calculated by converting the direct MMG audio signal in control value every 20ms; this time interval proved to be the best compromise between high resolution representation of the biodata and computer processor performance. L is the most used feature in MFII as it can be deployed to produce the perception of a neat and colourful coupling of the player's kinetic behaviour and the musical forms emanated by her body.

L is then passed through a SES function so to obtain T. This feature presents a minimal dynamics, it is therefore used to delicately control musical processes that require a careful and minimal lead, such as brief glissando and minimal sound spatialization.

Eventually, L is reiterated through a sub-process which produces the RMA. The computation consists of four steps: first, L is observed in order to identify the running average; then, the last maximum (LM) of the stream is extracted every 2s; finally, the discrete LM value is normalized and interpolated with its previous instance so to generate the RMA. This is a continuous event that moves away from the micro level of the single gesture, and reflects instead the average amount of energy that is being released by my body in a wider time window. Similarly to the mapping of the N feature, when applied improperly, the use of the RMA can disturb the audience perception of mutual interaction between the player and the machine. Nonetheless, a clever mapping implementation can outline the performer's agency by placing emphasis on an articulated series of actions, rather than an isolated gesture. The fifth movement of MFII is almost completely based on the RMA. Here the auditive space is fulfilled, nearly saturated. In this case the mapping definitions are quite complex; polyphony is obtained by playing back the sound generated by both my forearms. Simultaneously, the sonic matter undergoes a drastic processing lead by the RMA of multiple, bustling movements. I rely on the RMA to control the parameters of a pitch-shift based delay, a granular delay [6], a bandpass biquad filter and a cosine panner. Additionally, the position of the grains within the sonic field is subtly manipulated through T. Table 1 illustrates the control array.

Table 1: Mapping definitions in movement 5

Feature	Left arm	Right arm
RMA	delay line to pitch shifting	grain size
RMA	pitch-shift delay mix	grains delay time
RMA	not mapped	granular delay mix
RMA	not mapped	pitch-shift del. time
RMA	not mapped	filter freq. cut-off
T	not mapped	cosine panner

3.2 Visceral embodiment: sound-gestures

The features on their own are raw data. In order to use them musically, I conceived the notion of *sound-gesture* (SG). A SG is a compounded interpretation model that bonds a given feature mapping to the designated performer's gesture. The nature of a SG is twofold. On one hand, it is a gesture dictated by a neural impulse, that generates a given muscular excitement (i.e. a specific MMG sound). On the other, a SG relies on specific mapping definitions that live inside the circuits of the computer to achieve effectiveness and expressiveness. Hence, the SG can be seen as a techno-epistemic enactment of a dormant sonic capability of the body system.

In this sense a SG performed within the context of the XS is an extended and anomalous *instrumental gesture* [3]. Wanderley and Cadoz exclude the empty-handed gesture from the above category, for it owns only the *semiotic* function of the human gestural channel, that of communicating information toward the environment. They explain that this kind of gesture lacks of the *ergotic* and the *epistemic* functions, respectively the existence of a direct contact with the instrument, and the performer's use of her "tactile-kinaesthetic perception" to play the instrument. However, in the case of the XS, the instrument that a performer manipulates is not an external object, but the muscle fibre of her own body. The basic capability of the XS to musically deploy the muscle sounds produced by a performer challenges the nature of an instrumental gesture; the player does not act upon the external environment, but rather within

[5]An audio clip demonstrating this process is available on-line at `http://marcodonnarumma.com/publications/media/xth-sense_gesture-echoes_sample.wav`

[6]Included in the Soundhack collection by Tom Erbe. See: `http://ucsd.academia.edu/TomErbe/Papers/861787/SoundHack_Delay_Trio_Manual`

her own intimate, bodily milieu. One can therefore observe that a performer produces "specific (physical) phenomena" by mastering the tension of her own body (the *ergotic* function), while experiencing the enactment of a higher muscular and articulatory sensitivity (the *epistemic* function).

3.2.1 On the embodied musical outcome

At the micro level the sonic interaction is straight-forward: a single SG, such as the twitch of the wrist generates a single sound form. The meso (i.e. intermediate) level instead relies on the articulation of multiple gestures within what I call a sound *scene*; a scene consists of all the modification parameters available to the performer. By exerting a continuous amount of muscular energy, the player can choose which stages of the sound processing to activate and control. Finally, the macro level consists of the overall structure of the piece; diverse and independent SG definitions can be loaded at any given time into the system. The XS deals with this by combining machine learning (ML) [1] and a sensing timeline. ML is the design of algorithms that enables a computer to understand general behaviours according to empirical data. First, the computer learns (offline) four different performer's behaviours, that are labelled: *still*, *moving*, *fast gesture*, and *slow gesture*. During the live performance, the machine computes a real-time cross-comparison of the incoming data stream against what it learnt, and eventually identifies those behaviours. Then, key points are added to the timeline to indicate a global change of the scenes being played; as time passes and a key point in time is reached, the computer stands by and wait for the player to stand *still* before loading a new set of scenes. This strategy enables the solo performer to switch among rather complex arrays of sound processing chains, avoiding unpleasant clicks and artefacts that would otherwise be triggered. There is no need for a sound engineer, score following software or a direct contact with the machine; the XS takes care of all stages of the performance, leaving to the player the challenge and pleasure of delivering a successful performance.

For instance, during the fourth movement of MFII strong and wide contractions of my left forearm consistently repeated for more than 30 seconds prompt the computer to playback the muscle sound in its purest form: that of a deep, low frequency vibration between 20Hz and 40Hz. At the same (logical) time, the machine samples the nascent muscle sound and slightly transposes it up to 60Hz so to enhance its physical impact; finally, according to the dynamic features of my physio-somatic behaviour the computer recodes the MMG audio sample through granular synthesis, delay lines and pitch bending. The subcutaneous, low rumble of my flesh is amplified and made audible through subwoofers; simultaneously, a new textural layer appears: the grave, muscular sound wave mutates in high pitched grains that I can scatter and spatialize by nervously contracting my wrist. Then, I suddenly stop for about ten seconds; the break allows the machine to enter a condition of stand-by. In a couple of seconds I reach the required concentration to release my muscles completely, avoiding involuntary tension. At this point, all feature values gradually fall down to 0, triggering a drastic, yet continuous change in the duration of the granular delay lines. With the next contraction I begin to mangle the sound grains, I deform their aural image until a harsh and glassy bundle of mid high frequencies emerges, rapidly moving over a wide stereo field. The sustained exertion of my limbs, causes the machine to steadily increase the loudness and density of the sound output, until my body stands still once again, and finally, no sound is produced.

4. CONCLUSIONS

The work discussed in this paper can hopefully contribute a compelling statement on the practicability and richness of the biophysical music performance model. MFII demonstrates that the coupling of visceral musical embodiment and techno-epistemic processes can lead to convincing compositional strategies. Moreover, the paradigm of SG introduced here prompt for new challenges in the understanding of musical gesture, for they advance an envision of the body as an instrument with actual, and not merely metaphorical, sonic capabilities.

Future development will be dedicated to the composition of further works, and to the dissemination of the biophysical music paradigm through collaborations with other musicians, choreographers and artists. It is hoped that the openness of the XS technology will prompt for the birth of a community, which could build on the existent research so to explore different methods, perspectives and context of application.

5. ACKNOWLEDGEMENTS

Figure 2 is kind courtesy of Chris Scott. I would like to thank Martin Parker and Atau Tanaka for the thoughtful conversations that inspired and informed this research.

6. REFERENCES

[1] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, Feb. 2007.

[2] B. Bongers and Sensorband. An interview with sensorband. *Computer Music Journal*, 22(1):13–24, Apr. 1998.

[3] C. Cadoz, M. Wanderley, et al. Gesture-music. *Trends in Gestural Control of Music*, pages 71–94, 2000.

[4] M. Donnarumma. A pd framework for the xth sense: enabling computers to sense human kinetic behaviour. In *Proceedings of the 4th Pure Data Convention*, 2011.

[5] M. Donnarumma. Xth sense: a study of muscle sounds for an experimental paradigm of musical performance. In *Proceedings of the ICMC, International Computer Music Conference*, 2011.

[6] M. Donnarumma. Xth sense: recoding visceral embodiment. In *Proceedings of the CHI workshop on Liveness (in process)*, 2012.

[7] R. B. Knapp and H. S. Lusted. A bioelectric controller for computer music applications. *Computer Music Journal*, 14(1):42–47, Apr. 1990.

[8] E. R. Miranda and A. Brouse. Interfacing the brain directly with musical systems: On developing systems for making music with brain signals. *Leonardo*, 38(4):331–336, 2005.

[9] NIST/SEMATECH. e-handbook of statistical methods. http://1.usa.gov/8Hol6O, 2003.

[10] G. Oster and J. Jaffe. Low frequency sounds from sustained contraction of human skeletal muscle. *Biophysical Journal*, 30(1):119–127, Apr. 1980.

[11] D. Rosenboom. *Biofeedback and the arts: Results of early experiments*. Aesthetic Research Centre of Canada, 1974.

[12] J. Silva and T. Chau. Coupled microphone-accelerometer sensor pair for dynamic noise reduction in MMG signal recording. *Electronics Letters*, 39(21):1496–1498, 2003.

[13] C. Van Raalte. Old dog, new tricks. biofeedback as assistive technology. In *Proceedings of the CSUN, Los Angeles*, 1998.