XTH SENSE: A STUDY OF MUSCLE SOUNDS FOR AN EXPERIMENTAL PARADIGM OF MUSICAL PERFORMANCE

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

ABSTRACT

This paper seeks to outline methods underlying the development of the Xth Sense project, an ongoing research which investigates exploratory applications of biophysical sound design for musical performance and responsive milieux.

Firstly, I describe the development and design of the Xth Sense, a wearable hardware sensor device for capturing biological body sounds; this was implemented in the realization of Music for Flesh I, a first attempt at musical performance. Next, the array of principles underpinning the application of muscle sounds to a musical performance is illustrated. Drawing from such principles, I eventually describe the methods by which useful features were extracted from the muscle sounds, and the mapping techniques used to deploy these features as control data for real time sound processing.

1. INTRODUCTION

Biosensing musical technologies use biological signals of a human subject to control music. One of the earliest applications can be identified in Alvin Lucier's *Music for Solo Performer* (1965). Alpha waves generated when the performer enters a peculiar mind state are transduced in electrical signals used to vibrate percussion instruments. Over the past thirty years biosensing technologies have been comprehensively studied [3, 8, 13, 14, 15, 18, 22] and presently notable biophysical-only music performances[1] are being implemented at SARC[2] by a research group lead by the main contributor to the Bio Muse project[3] Ben Knapp (10).

Whereas *biological motion* and *movement and music* are arising topics of interest in neuroscience research [5, 12, 21], the biologic body is being studied by music researchers as a means to control virtual instruments. Although such approach has informed gestural control of music, I argue that it overlooks the expressive capabilities of muscle sounds. They are inaudible to the human ear, but can be amplified and data extracted from these signal inside a computer may retain a meaningful vocabulary of intimate interactions with the musicians' actions.

To what extent could muscle sounds be employed musically? In which ways could the performer's perceptual experience be affected? How could such experimental paradigm motivate an original perspective on musical performance?

2. AESTHETIC PRINCIPLES

The long-term outcome of the research is the implementation of low cost, open source tools (software and hardware) capable of providing musicians, performers and dancers with a framework for biosensors-aided auditive design (BAAD)[4] in a real time[5] environment. The framework will be re-distributable, customizable and easy to set up. However, given the substantial interdisciplinary quality of such project, its realization process needed to be fragmented into more specific and measurable steps.

At a first stage, primary aim of the inquiry was to explore the musical deployment and design of biological sounds of the body in a functional context – the production of *Music for Flesh I* a solo sound performance for wearable biosensing device, which could demonstrate an experimental coupling between unheard sounds of muscle gestures and corresponding sound synthesis played back through loudspeakers. In an attempt to inform the present state of augmented musical performance and embodied interaction in performing environments, the characteristics of this pairing were identified in the authenticity of the performer's somatic interaction, the natural responsiveness of the system and the expressive immediacy and transparency of the mapping of muscle sounds to the performer's kinetic behaviour. Such work required an interdisciplinary approach embracing biomedical computing studies, music technology and most importantly sound design. In fact, as I will demonstrate later in this text, the major research issue was not a technical implementation, but rather the definition of design paradigms by which the captured biological sounds could achieve a meaningful and detailed expressiveness.

---

[1] Bio Muse Trio, GroundMe!.

[2] Queen's University, Sonic Art Research Center, Belfast, UK.

[3] A commercialized product exploiting electromyography and brainwaves analysis systems for musical applications.

[4] BAAD is a novel term used by the author to indicate a specific sound design practice which relies on the use of biological signals. Although in this context is not possible to further elaborate on this practice, its essential principles are defined in paragraph 4.1.

[5] Real time refers here to a computing system in which there exists no perceivable delay between performer's actions and sonic response.

3. METHODS: UNDERSTANDING AND CAPTURING MUSCLES SOUNDS

The project consisted of two interrelated strands of research. The first concerned the design of muscle sounds and their meaningful mapping to the somatic behaviour of the performer; the second included the design and implementation of a wearable biosensing hardware device for musical performance. Chosen research methods are discussed in the following paragraphs; however, being the focus of this paper on the research methodology, specific signal processing techniques and other technical information are not illustrated in detail, but they are fully referenced.

3.1. Xth Sense: first prototype sensor implementation

Before undertaking the development of the *Xth Sense* sensor hardware, few crucial criteria were defined:

- to develop a wearable, unobtrusive device, allowing a performer to freely move on stage;
- to implement an extremely sensitive hardware device which could efficiently capture in real time and with very low latency diverse muscle sounds;
- to make use of the most inexpensive hardware solutions, assuring a low implementation cost;
- to implement the most accessible and straightforward production methodology in order to foster the future re-distribution and openness of the hardware.

Study of the hardware sensor design began with a contextual review of biomedical engineering papers and publications focused on mechanical myography (MMG). The mechanical signal which can be observed from the surface of a muscle when it is contracted is called a MMG signal. 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 fibers at the resonant frequency of the muscle generate subsequent vibrations. The mechanomyogram is commonly known also as the phonomyogram, acoustic myogram, sound myogram or vibromyogram.

Interestingly, MMG seems not to be a topic of interest in the study of gestural control of music and music technology; apparently many researchers in this field focus their attention on electromyography (EMG), electroencephalography (EEG), or multidimensional control data which can be obtained through the use of wearable accelerometers, gyroscopes and other similar sensors. Notwithstanding the apparent lack of pertinent documentation in the studies of gestural control of music and music technologies, useful technical information regarding different MMG sensor designs were collected by reviewing recent biomedical engineering literature.

In fact, MMG is currently the subject of several investigations in this field as alternative control data for low cost, open source prosthetics research and for general biomedical applications [1, 6, 9, 20]. Most notably the work of Jorge Silva at Prism Lab; his MASc thesis extensively documents the design of a coupled microphone-accelerometer sensor pair (CMASP) and represents a comprehensive resource of information and technical insights on the use and analysis of MMG signals [19]. The device designed at Prism Lab is capable of capturing the audio signal of muscles sounds in real time. Muscle sonic resonance is transmitted to the skin, which in turn vibrates, exciting an air chamber. These vibrations are captured by an omnidirectional condenser microphone adequately shielded from noise and interferences by mean of a silicon case. A printed circuit board (PCB) is used to couple the microphone with an accelerometer in order to filter out vibrations caused by global motion of the arm, and precisely identify muscle signals (figure 1). Microphone sensitivity ranges from 20Hz up to 16kHz, thus it is capable of capturing a relevant part of the spectrum of muscle resonance[6].

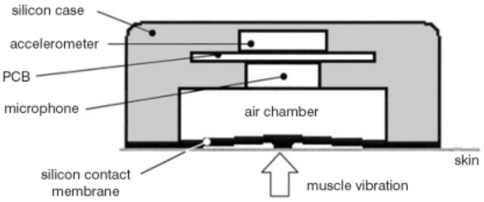

Figure 1. CMASP schematic

Although this design has been proved effectively functional through several academic reports, criteria of my investigation could have been satisfied with a less complex device. Supported by the research group at Dorkbot ALBA[7], I could develop a first, simpler MMG sensor: the circuit did not make use of a PCB and accelerometer, but deployed the same omnidirectional electret condenser microphone indicated by Silva (Panasonic WM-63PRT). This first prototype was successfully used to capture actual heart and forearm muscles sounds; earliest recordings and analysis of MMG signals were produced with the open source digital audio workstation Ardour2 and a benchmark was set in order to evaluate the signal-to-noise ratio (SNR).

In spite of the positive results obtained with the first prototype, the microphone shielding required further trials. The importance of the shield was manifold; an optimal shield had to fit specific requirements: to bypass the 60Hz electrical interference which can be heard when alternating electric current distribute itself within

---

[6] On a side note, it is interesting to observe that the biggest part of muscles sounds spectra seems to sit below 20Hz, thus pertaining to the realm of infra-sounds. Such characteristic is not being explored at the moment only due to technical constraints, although it suggests appealing prospects for a further research.

[7] Electronics open research group based in Edinburgh. See: http://dorkbot.noodlefactory.co.uk/wiki

the skin after a direct contact with the microphone metal case; to narrow the sensitive area of the microphone, filtering out external noises; to keep the microphone static, avoiding external air pressure to affect the signal; to provide a suitable air chamber for the microphone, in order to amplify sonic vibrations of the muscles, allowing to capture also deeper muscle contractions.

First, microphone was insulated by means of a polyurethane shield, but due to the strong malleability of this material, its initial shape tended to undergo substantial alterations. Eventually, sensor was insulated in a common silicon case that positively satisfied the requirements and further enhanced the SNR. Once the early prototype had reached a good degree of efficiency and reliability, the circuit was embedded in a portable plastic box (3.15 x 1.57 x 0.67) along with an audio output (¼ mono chassis jack socket) and a cell holder for a 3V coin lithium battery. The shielded microphone was embedded in a Velcro bracelet and needed wiring cables were connected to the circuit box (figure 2).

Figure 2. *Xth Sense* wearable MMG sensor prototype

4. PERFORMANCE TESTING: MAPPING AND DESIGN DEFINITIONS

At this stage of the project the creation of design paradigms for mapping muscle sounds was the major goal. The main principles and some technical implementations are illustrated below.

4.1. Sound performance and design principles

Major aim of the design of the MMG audio signals was to avoid a perception of the sound being dissociated from the performer's gesture. The dissociation I point at not only refers to the visual feedback of the performer's actions being disjointed from the sonic experience, but it also, and most importantly, concerns a metaphorical level affecting the listener's interpretation of the sounds generated by the performer's somatic behavior [2]. In this project the use of muscle sounds had to be clearly motivated in order to inform classical approaches to gestural control of music. Therefore, chosen sound processing and data mapping techniques were evaluated according to their capability of enhancing the metaphorical interpretation of the performer's physiological and spatial behaviour.

From this perspective, the essential principles of BAAD in a performing environment were defined as follows:

- to make use of biological sounds as major sonic source and control data;
- to exclude the direct interaction of the performer with a computer and to conceal the latter from the view of the public;
- to demonstrate a distinct, natural and non-linear interaction between kinetic energy and sonic outcome which could be instinctively controlled by the performer;
- to provide a rich, specific and unconventional vocabulary of gesture/sound definitions which can be unambiguously interpreted by the audience;
- to allow the performer to flexibly execute the composition, or even improvise a new one with the same sonic vocabulary;
- to make both performer and public perceive the former's body as a musical instrument and its kinetic energy as an exclusive sound generating force.

4.2. MMG features extraction

Since the project dealt with sound data, a pitch tracking system may have been a straightforward solution for an automated evaluation and recognition of gestures, however muscle sound's resonance frequency is not affected by any external agent and its pitch seems not to change significantly with different movements [17]. Whereas muscle sounds are mostly short, discrete events with no meaningful pitch change information, the most interesting and unique aspect of their acoustic composition is their extremely rich and fast dynamic; therefore, extraction of useful data can be achieved by RMS amplitude analysis and tracking, contractions onset and gesture pattern recognition. In fact, each human muscle exerts a different amount of kinetic energy when contracting and a computing system can be trained in order to measure and recognize different levels of force, i.e. different gestures. Feature extraction enabled the performer to calibrate software parameters according to the different intensity of the contractions of each finger or the wrist and provided 8 variables: 6 discrete events, 1 continuous moving event and 1 continuous exponential event.

First, the sensor was subjected to a series of movements and contractions with different intensity to identify a sensitivity range; this was measured between 57.79 dB (weakest contraction) and 89.04 dB (strongest contraction). The force threshold of each finger discrete contraction was set by normalizing and measuring the individual maximum force exertion level; although some minor issues arisen from the resemblance between the force amplitude exerted by the minimus (little finger) and the thumb still need to be solved, this method allowed the determination of 6 independent binary trigger control messages (fingers and wrist contractions).

Secondly, by measuring the continuous amplitude average of the overall contractions, it was possible to extract the running maximum amplitude of performer's gestures; in order to correct the jitter of this data, which otherwise could not have been usefully deployed, value was extracted every 2 seconds, then interpolated with the prior one to generate a continuous event and eventually normalized to MIDI range. Lastly, a basic equation of single exponential smoothing (SES) was applied to the moving global RMS amplitude in order to forecast a less sensitive continuous control value [16].

4.3. Mapping kinetic energy to control data

A first mapping model deployed the 6 triggers previously described as control messages. These were used to enable the performer to control the real time SSB modulation algorithm by choosing a specific frequency among six different preset frequencies; the performer could select which target frequency to apply according to the contracted finger; therefore, the voluntary contraction of a specific finger would enable the performer to "play" a certain note.

A one-to-many mapping model, instead, used the continuous values obtained through the RMS analysis to control several processing parameters within five digital signal processing (DSP) chains simultaneously. Being that this paper does not offer enough room to fully describe the whole DSP system which was eventually implemented, I will concentrate on one example chain which can provide a relevant insight on the chosen mapping methodology; namely, this DSP chain included a SSB modulation algorithm, a lofi distortion module, a stereo reverb, and a band-pass filter.

The SSB algorithm was employed to increase the original pitch of the raw muscle sounds by 20Hz, thus making it more easily audible. Following an aesthetic choice, the amount of distortion over the source audio signal was subtle and static, thus adding a light granulation to the body of the sound; therefore, the moving global RMS amplitude was mapped to the reverb decay time and to the moving frequency and Quality factor[8] (Q) of the band-pass filter.

The most interesting performance feature of such mapping model consisted of the possibility to control a multi-layered processing of the MMG audio signal by exerting different amounts of kinetic energy. Stronger and wider gestures would generate sharp, higher resonating frequencies coupled with a very short reverb time, whereas weaker and more confined gestures would produce gentle, lower resonances with longer reverb time.

Such direct interaction among the perceived force and spatiality of the gesture and the moving form and color of the sonic outcome happened with very low latency, and seemed to suggest promising further applications in a more complex DSP system.

Figure 3. *Music for Flesh I* first public performance, 2010

The Xth Sense framework was tested live during a first public performance of Music for Flesh I (figure 3) at the University of Edinburgh (December 2010). Although the system was still in development, it proved reliable and efficient. Audience feedback was positive, and apparently what most appealed some listeners was an authentic, neat and natural responsiveness of the system along with a suggestive and unconventional coupling of sound and gestures.

5. CONCLUSIONS

Results reported in this paper appear to disclose promising prospects of an experimental paradigm for musical performance based on MMG. The development of the Xth Sense and the composition and public performance of Music for Flesh I can possibly demonstrate an uncharted potential of biological sounds of the human body, specifically muscle sounds, in a musical performance.

Notwithstanding the apparently scarce interest of the relevant academic community towards the study and the use of muscle sounds, the experiment described here shows that these sounds could retain a relevant potential for an exploration of meaningful and unconventional sound-gesture metaphors. Besides, if compared to EMG and EEG sensing devices, the use of MMG sensors could depict a new prospect for a simpler implementation of unobtrusive and low-cost biosensing technologies for biophysical generation and control of music.

Whereas the development of the sensor hardware device did not present complex issues, several improvements to the tracking and mapping techniques can lead to a further enhancement of the expressive vocabulary of sound-gestures. In an attempt to enrich the performer's musical control over a longer period of time, hereafter priority will be given to the extraction of other useful features, to the development of a gesture pattern recognition system and to the implementation of a system for multiple sensors.

---

[8] Narrowness of the filter.

6. REFERENCES

[1] Alves, N., Sejdić, E., Sahota, B. and Chau, T. “The effect of accelerometer location on the classification of single-site forearm mechanomyograms,” *Biomedical Engineering Online*, vol. 9, 2010, p. 23.

[2] D. Arfib, J.M. Couturier, L. Kessous, and V. Verfaille, “Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces,” *Organised Sound*, vol. 7, 2003.

[3] Arslan, B., Brouse, A., Castet, J., Léhembre, R., Simon, C., Filatriau J.J. And Noirhomme, Q. “A real time music synthesis environment driven with biological signals,” *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2006, pp. 14–19.

[4] Clark, J.J. “Frequency and Pitch Shifting.” *Advanced Programming Techniques for Modular Synthesizers*. Available at: http://www.cim.mcgill.ca/~clark/nordmodularbook/nm_spectrum_shift.html [Accessed January 12, 2011].

[5] Deen, B. and McCarthy, G. “Reading about the actions of others: Biological motion imagery and action congruency influence brain activity,” *Neuropsychologia*, vol. 48, May. 2010, pp. 1607-1615.

[6] Esposito, F., Cè, E., Rampichini, S. and Veicsteinas, A. “Acute passive stretching in a previously fatigued muscle: Electrical and mechanical response during tetanic stimulation,” *Journal of Sports Sciences*, vol. 27, 2009, p. 1347.

[7] Moorer, J. “The Synthesis of Complex Audio Spectra by Means of Discrete Summation Formulas”, Journal of the Audio Engineering Society, vol. 24, 1976, pp. 717—727.

[8] Filatriau, J.J. and Kessous, L. “Visual and sound generation driven by brain, heart and respiration signals,” *Proceedings of the 2008 International Computer Music Conference (ICMC’08)*, 2008.

[9] Garcia, M.A.C., Vargas, C.D., de Souza, M.N., Imbiriba, L.A. and de Oliveira, L.F. “Evaluation of arm dominance by using the mechanomyographic signal,” *Journal of Motor Behavior*, vol. 40, Mar. 2008, pp. 83-89.

[10] Knapp, R.B. and Lusted, H.S. “A Bioelectric Controller for Computer Music Applications,” *Computer Music Journal*, vol. 14, Apr. 1990, pp. 42-47.

[11] Mäki-patola, T. and Hämäläinen, P. “Latency Tolerance for Gesture Controlled Continuous Sound Instrument Without Tactile Feedback,” *Proceedings of the 2004 International Computer Music Conference (ICMC’04)*, 2004.

[12] Manera, V., Schouten, B., Becchio, C., Bara, B.G. and Verfaillie, K. “Inferring intentions from biological motion: A stimulus set of point-light communicative interactions,” *Behavior Research Methods*, vol. 42, Feb. 2010, pp. 168-178.

[13] Miranda, E.R. and Brouse, A. “Toward direct Brain-Computer Musical Interfaces.” *Proceedings of the Conference on New Interfaces for Musical Expression (NIME05)*, 2005.

[14] Miranda, E.R. “Brain-Computer Music Interface for Composition and Performance.” *International Journal on Disability and Human Development*, 2006, Vol. 5, Issue 2, pp. 119-125.

[15] Nagashima, Y. “Bio-Sensing Systems and Bio-Feedback Systems for Interactive Media Arts,” *Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME-03*, 2003, pp. 48—53).

[16] “NIST/SEMATECH e-Handbook of Statistical Methods.” Single Exponential Smoothing, *e-Handbook of Statistical Methods*, 2003. Available at: http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc431.htm [Accessed January 12, 2011].

[17] Oster, G. and Jaffe, J. “Low frequency sounds from sustained contraction of human skeletal muscle,” *Biophysical Journal*, vol. 30, Apr. 1980, pp. 119-127.

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

[19] Silva, J. *Mechanomyography Sensor Design and Multisensor Fusion for Upper-Limb Prosthesis Control*, PhD Thesis, Master of Applied Sciences, Graduate Department of Mechanical & Industrial Engineering University of Toronto, 2004.

[20] Silva, J. and Chau, T. “Coupled microphone-accelerometer sensor pair for dynamic noise reduction in MMG signal recording,”

Electronics Letters, vol. 39, 2003, pp. 1496-1498.

[21] Stevens, C.. Ginsborg, J. and Lester, G. “Backwards and forwards in space and time: Recalling dance movement from long-term memory,” *Memory Studies*. Available at: http://mss.sagepub.com/content/early/2010/11/05/1750698010387018.full.pdf+html [Accessed January 12, 2011].

[22] Tanaka, A. “Introduction Musical Performance Practice on Sensor-based Instruments.” *Trends in Gestural Control of Music*, M. Battier and M. Wanderley eds., 2000, pp. 389-406.