Xth sense:
a biophysical framework for the application of muscle sounds to interactive music performance

Marco Donnarumma
Department of Arts Culture & Environment, Sound Design, The University of Edinburgh, Scotland, UK
M.Donnarumma@sms.ed.ac.uk
http://marcodonnarumma.com

Proceedings of Korean Electro-Acoustic Music Society's 2011 Annual Conference (KEAMSAC2011)
Seoul, Korea, 22-23 October 2011

Abstract. The Xth Sense (XS) is an interactive system for the biophysical generation and control of music. It makes use of muscle sounds[i] produced by a performer as both raw sonic material and control data. Presently the XS technology consists of low-cost, wearable biosensors and a Pure Data[ii] based application for capture, analysis, real time processing and playback of human muscle sounds.

After an introduction on the nature of the interaction fostered by the XS technology and the aesthetic motivations underpinning the project, I focus on the development and design of the XS wearable sensor device; this was implemented in the realization of Music for Flesh II, a musical performance for enhanced body. 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.

Computers that sense and act

Interactive music relies on a flowing exchange between a human being and a computer. As Winkler puts it, "nothing is more interactive than a good conversation"; both parties are engaged, they both share ideas and respond to each others inputs. A good conversation is usually open ended, yet bounded by an intangible frame of rules and cultural contingencies, that are - more or less equally - shared by the participants. It is a delicate balance which can be achieved only by setting a "consistent context" that produces "a feeling of mutual understanding without being predictable" [Winkler, 2001]. On the other hand, for Rokeby the computer is "objective and disinterested", and therefore the interaction (e.g. experience) "should be intimate". Computers are not smart. Man creates their intelligence, Man forges them as interactive agents, and ultimately defines their degree of interaction with the real world. Man can breathe an understanding of human intimacy into the computer "tiny playing fields of integrated circuits" [Rokeby, 1886-1990].

At the heart of the XS project stands a twofold motivation: to investigate the modalities by which Man's intimate, bodily energy can become digitally tangible; to develop a context of rules and algorithms which would enable computers to sense the varied nuances of the body potential, and act accordingly. Here, I purposely refer to the term acting; it might be argued that it does not matter how much complex a computing system is, it will always be pretending to be able of making sense of human behavior. Its understanding is an artificial product of Man's choices.

The following sections seek to outline the aesthetic and the related techniques (e.g. the developer's choices) that drives the mutual and creative exchange between the user and the machine within the XS interactive framework.

Understanding and capturing muscle 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 behavior of the performer; the second included the 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.

First prototype sensor implementation

Before undertaking the development of the XS 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 [Alves et al., 2010; Esposito et al., 2009; Garcia et al., 2008; Silva and Chau, 2003]. 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 [Silva, 2004]. 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).

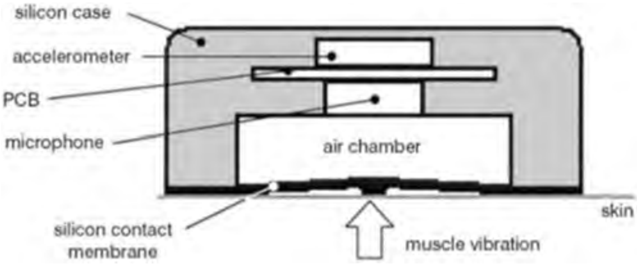

Figure 1. CMASP schematic

Microphone sensitivity ranges from 20Hz up to 16kHz, thus it is capable of capturing a relevant part of the spectrum of muscle resonance[iii]. 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[iv], 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 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

Interrelating kinetic behaviour with musical performance

This section describes the aesthetic principles, the features extraction and the mapping techniques that enable the composition of a performance based on the somatic behavior of a performer. How to achieve a seamless, real time interaction with a computer software for DSP, which would guarantee richness of color and sophistication of forms?

Performance design and principles.

Major aim of designing with MMG audio signals is 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 [Arfib et al., 2003]. The use of muscle sounds in this project 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 behavior.

From this perspective, some essential principles 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.

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 [Oster and Jaffe, 1980]. 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 [NIST/SEMATECH, 2003].

Mapping muscular energy to control data. This paragraph describe some mapping models developed during the composition of *Music for Flesh II*[v], a solo sound piece for the XS (figure 3).

Figure 3. Music for Flesh II live at Inspace, Edinburgh, May 2011

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 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 which can provide a relevant insight on the chosen mapping methodology; namely, a DSP chain which included a SSB modulation algorithm, a lo-fi 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[vi] (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.

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 XS and the composition and public performance of Music for Flesh II 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.

References

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

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

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

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

NIST/SEMATECH. (2003). Single Exponential Smoothing. *e-Handbook of Statistical Methods*. Retrieved from http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc431.htm

Oster, G. and Jaffe, J. (1980). Low frequency sounds from sustained contraction of human skeletal muscle. *Biophysical Journal*, vol. 30, pp. 119-127.

Winkler, T. (2001). *Composing Interactive Music: Techniques and Ideas Using Max*. Cambridge, Massachusets; London, England: The MIT Press, 2001.

Rokeby, D. (1886-1990). *Very Nervous System*. Retrieved from http://homepage.mac.com/davidrokeby/vns.html

Silva, J. (2004). *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.

Silva, J. and Chau, T. (2003). Coupled microphone-accelerometer sensor pair for dynamic noise reduction in MMG signal recording. *Electronics Letters*, vol. 39, pp. 1496-1498

[i] Technically called mechanical myography (MMG)

[ii] A programming language for real time signal processing and computer music.

[iii] 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.

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

[v] See: http://marcodonnarumma.com/works/music-for-flesh-ii/

[vi] Narrowness of the filter.

[Abstract in Korean | 국문 요약]
제10의 감각: 근육 소리의 상호작용 음악 공연에의 응용을 위한 생물학적 구성 체제

마르코 돈나룸마

Xth Sense는 생물물리학적으로 음악을 생성하고 조종하는 상호작용 체계interactive system이다. 이 장치는 한 행위자가 만들어내는 근육 소리를 변형 없이 그대로 소리의 재료와 조종 데이터로 사용한다. 현재의 Xth Sense는 저가의 착용식 바이오 센서들과, 순수 데이터를 토대로 하여 사람의 근육 소리를 포착, 분석, 실시간 처리 및 재생하는 하나의 응용 프로그램으로 구성되어 있다.

이 글은 이러한 구성으로 만들어지는 상호 작용성과 이 연구의 밑바탕이 되는 미학적 의미에 대한 논의를 서두로 하고, Xth Sense 착용식 센서기계의 개발과 고안에 대한 것을 이어지는 논의의 중심으로 삼았다. 이 센서는 고도화된 신체로 연주하는 형태의 음악 작품인 <육체음악 II>를 공연하는 데 사용되었던 것이다. 그 다음으로, 근육 소리를 음악 연주로 적용시키는데 관련된 원리들을 나열하고 설명한다. 본고에서는 이러한 원리를 예시함으로써, 근육 소리에서 유용한 특성을 발견해 내는 방법과 그 추출된 특성을 실시간 처리 데이터로 활용하는 방법에 대해 밝히고 있다.