

MARCO DONNARUMMA

Ominous: Playfulness and emergence in a performance for biophysical music

Abstract

This article tackles the issue of playfulness in biosignal-based performance with digital musical instruments. The practical context is that of an interactive sound sculpture performance, entitled *Ominous*, for the *Xth Sense*, a biophysical musical instrument. This work posits that a prominent quality of the body physiology is its emergence. Developed by cultural theorist Brian Massumi, the notion of bodily emergence is useful to inform playfulness in musical performance for it can be used to understand a performer's body as being continuously changed by physiological and autonomic processes. This article presents a strategy for musical interaction that uses the relations in time and intensity of two biosignals to make the musical instrument adapt to the performer's physiological states. This strategy is put into practice by implementing a system that, drawing upon findings in biomedical engineering, brings together biosignal feature extraction, multidimensional mapping and digital neural networks.

Body and Biosignals in Music Performance: A brief History

Biosignals are electrical or mechanical signals produced by the human body as a result of physiological processes. There are a number of artistic works that use biosignals (Ortiz, 2012). Each proposes a different perspective on what a body is or can do. This section briefly tracks the artistic perspectives on the body in the field of biosignal-driven music and performance and characterises the kind of body that is the subject of this work.

Within early works using biosignals, a common performance strategy was to immerse the player's body in its own subjective experience. In the work of musicians such as David Rosenboom and Alvin Lucier (Rosenboom, 1976), the player sat still in a meditative state while brainwaves – electrical signals from the brain – were used to modulate sounds generated by analog synthesisers or percussive instruments. Rosenboom and Lucier were interested in the power of the human mind to extend perception of and control over sound events. Today, the player's physical stillness may recall to some an idea of the human body as inert matter that cannot speak back.

Performance artist Stelarc used biosignals to stage a connected body. His well-known aphorism 'the body is obsolete' implies that it is not our physical presence that makes us human, but rather our connectivity with the environment and the beings or things that inhabits it (Donnarumma, 2012a). In his body suspension series, Stelarc used brainwaves, electrical signals from his muscles and blood flow sounds to create soundscapes that would directly link his physical body to the space that surrounded it, so that they would resonate. In his Internet-based works, Stelarc enabled online users to remotely activate an electrical stimulation system that forced convulsive contractions onto his body.

More recently, music performer Atau Tanaka and myself have produced work on the physicality of the body in biosignal-based musical performance (Tanaka, 2010; Donnarumma 2012; Donnarumma & Tanaka, 2014). We have described embodied views of music performance, where the emphasis is on the role of the body in shaping the way a

performer plays, perceives and composes music. In Tanaka's work, electrical biosignals from limb tension are used to control digital sound modulation, an interaction modality known as biocontrol. Here, the body materiality is integral to the successful entanglement of player, technology, music, and audience; yet, the physiological processes are not thought of as purely biological, but rather as a source of musical expressivity. What is the expressive quality of bodily processes? How can that quality be used to design a musical piece?

In my recent work with biotechnologies for body-based performance I focus on the notion of bodily emergence as a prominent quality of physiological processes. The cultural theorist Brian Massumi has used this quality to indicate a continuous transformation of the body originating from unconscious visceral mechanisms (Massumi, 2002). This idea describes a body that never reaches a fixed condition. Rather, it is a body that shifts endlessly between one state and another. By looking at the body from this viewpoint, one can see that to simply collect streams of data from physiological processes is a limited strategy. Physiological data are affected by the changing states of the body, changes that often happen independently from the performer's conscious actions and yet, they still affect the musical delivery. One could argue that the data can be adequately smoothed, interpolated and normalised to eliminate 'noise' and produce a consistent representation of bodily processes. The counter argument is that the result of such data manipulation is only an approximation, a scattered array of digits that cannot represent continuous and emerging processes such as those characterising the physiology of a performer's body.

This article presents a performance strategy whereby the musical instrument continuously adapts to the performer's physiological state. As a practical example, I will discuss the piece *Ominous*, where that strategy is put into practice through the use of biosignal feature extraction, multidimensional mapping and digital neural networks. As it will be discussed in the remainder of this article, the continuous changes of the player's physiological state are crucial to all components of the performance. The gestural vocabulary, sound processing, time structure and compositional strategy are progressively shaped in real-time through the performer's effort in mediating physiological processes. The unpredictability of these processes is what makes the performance both playful and challenging.

In the following sections I will discuss the musical instrument and the time domain features it extracts from the performer's biosignals. I will then go on to explore the aesthetic and performative issues leading to the technical development of a custom multidimensional mapping method and describe the motivation behind the adoption of a digital neural network system and its practical application to musical performance. Finally, I will conclude with some final observations on playfulness and delineate recent on-going research that has emerged from this work.

Tracking gesture salient traits using muscle sounds

Ominous is a real-time gestural music performance for incarnated sound. That is, music is created in real-time using amplified muscles sounds (technically called mechanomyogram or MMG). Muscle sounds are a type of biosignal that consist of acoustic vibrations emitted by muscular tissues. During the performance of *Ominous*, I use the Xth Sense¹, a biophysical musical instrument, to compose real-time music with the clusters of sound released by the muscles (Donnarumma, 2011).

The muscle sounds are captured by a pair of Xth Sense sensors located on the forearms and analysed by dedicated software that extracts continuous high-level features. The

features are summarised below. An in-depth description of the feature extraction is beyond the scope of this article. The interested reader can refer to (Donnarumma, 2012).

- *Natural (N)*
RMS-based value analysed with a Hanning window of 512 samples;
- *Soft (S)*
value obtained by passing N through a single exponential smoothing function (SES);
- *Linear (L)*
value obtained by converting the MMG audio signal into control rate message every 20ms;
- *Tanh (T)*
value obtained by passing L through an additional SES function;
- *Maximum Running Average (MRA)*
value obtained by interpolating the two last maximum values of L every 2s.

According to the unique traits of the body's muscular tension represented through these features, the muscle sounds are digitally processed and eventually played back through eight subwoofers and eight loudspeakers. A basic characteristic of muscle sounds is that their loudness increases with the strength of the contraction (Bolton et al., 1989). For instance, a sudden and strong arm contraction produces a loud sound with a sharp attack and a very short release. In this piece, a specific mapping technique extends that relationship between strength and loudness by adding multiple dimensions to it. The dynamics of the muscle sounds becomes a continuous stream of data that drives the processing of the sound output. In order to ensure a fair amount of complexity and richness, up to eight simultaneous sampling dimensions are made available to the player. In this way the interrelation of agency, musicianship, and musicality can remain transparent throughout the piece.

This is a performance model that I called biophysical music. The model has been discussed extensively in several publications and so the topic will therefore not be discussed in the present article. Rather, in the following sections I will discuss how biophysical music can serve as the ground to explore the design of a musical composition that depends upon, and adapts to the emergent nature of a performer's body. Those readers looking for an introduction and a detailed discussion of biophysical music performance can find useful references in the publications list available online at my portfolio.²

Playfulness

Before the audience, *Ominous* embodies the metaphor of an invisible and unknown object enclosed in my hands. This is made of malleable sonic matter. Similar to a mime, I model the object in the empty space by means of whole-body gestures. The muscle sounds produced by the contractions of my muscle tissues are amplified, digitally processed and played back through a sound system. The natural sounds of my muscles and their virtual counterpart blend together into an unstable sonic object. This oscillates between a state of high density and violent release. As the listeners imagine the object's shape by following my physical gestures, the sonic stimulus induces a perceptual coupling. The listeners see through sound the sculpture that their sight cannot perceive.

The realisation of *Ominous* took a large amount of time. One of the main concerns was how not to repeat the performance modalities of the previous works I had composed for

the Xth Sense. At that time, around the beginning of May 2012, I felt as if my earlier compositional strategies had become a constraint. After all, at the time I had been performing with the Xth Sense for about two years. My technique had improved, and this could, perhaps, have helped me to investigate a mode of interaction that I had not been able to devise before.



Figure 1: A sequence illustrating multidimensional gestures in *Ominous*. Photographs: Ugo Dalla Porta, 2015.

Multidimensional sound-gestures

In my previous compositions, each arm generated individual sounds that were processed separately. The software analysed the muscle sounds of both arms as if they were completely independent from each other. In most cases, however, a limb flexion tends to enact sympathetic vibrations of the adjacent limbs. This means that the muscle sounds emitted by a limb are the joint result of the flexion of the observed muscle and the subsequent vibrations of the other limbs. In short, the muscle sounds analysed at the

capture point are the sum of interrelated limb vibrations. This implies that a skilful coordination of multiple limbs results in a fine control over muscle sounds dynamics. In addition, scientific studies on the spectral variances of muscle sounds found that, as muscular force increases, the biosignal frequency spectrum becomes broader (Orizio et al., 1990). This points to the fact that the muscle sounds spectrum can be modulated by varying the intensity of a contraction. By improving whole-body coordination, a player can weigh muscular force so as to produce specific spectral results. Such observations have suggested that an examination of the relationships underlying simultaneous biosignals streams may be fruitful. As a result, a model for a multidimensional sound-gesture (see Figure 1) has been implemented. This relies on reciprocal time and intensity relations of two synced biosignals. To implement that model I looked at traditional musical instruments.

When playing a traditional instrument, limb coordination is critical to both the quality of the music and the pleasure of performing. In the case of a chord instrument, for instance, synced gestures cause complex timbre variances. The initial plucking defines the amplitude and rate of the chords oscillation: a gentle gesture provokes a quiet vibration, while a conceited one introduces distortion and harmonics. The fingers' movements in turn determine pitch changes, modulating the sound dynamic and causing resonance. For a player, being able to create a specific sonority by skilfully articulating such gestures is gratifying. With a traditional instrument in mind, the Xth Sense's original configuration was refined to enable a player to produce a sound with a limb contraction, and modulate that sound with a synchronous flexion of another limb. This model is described next.

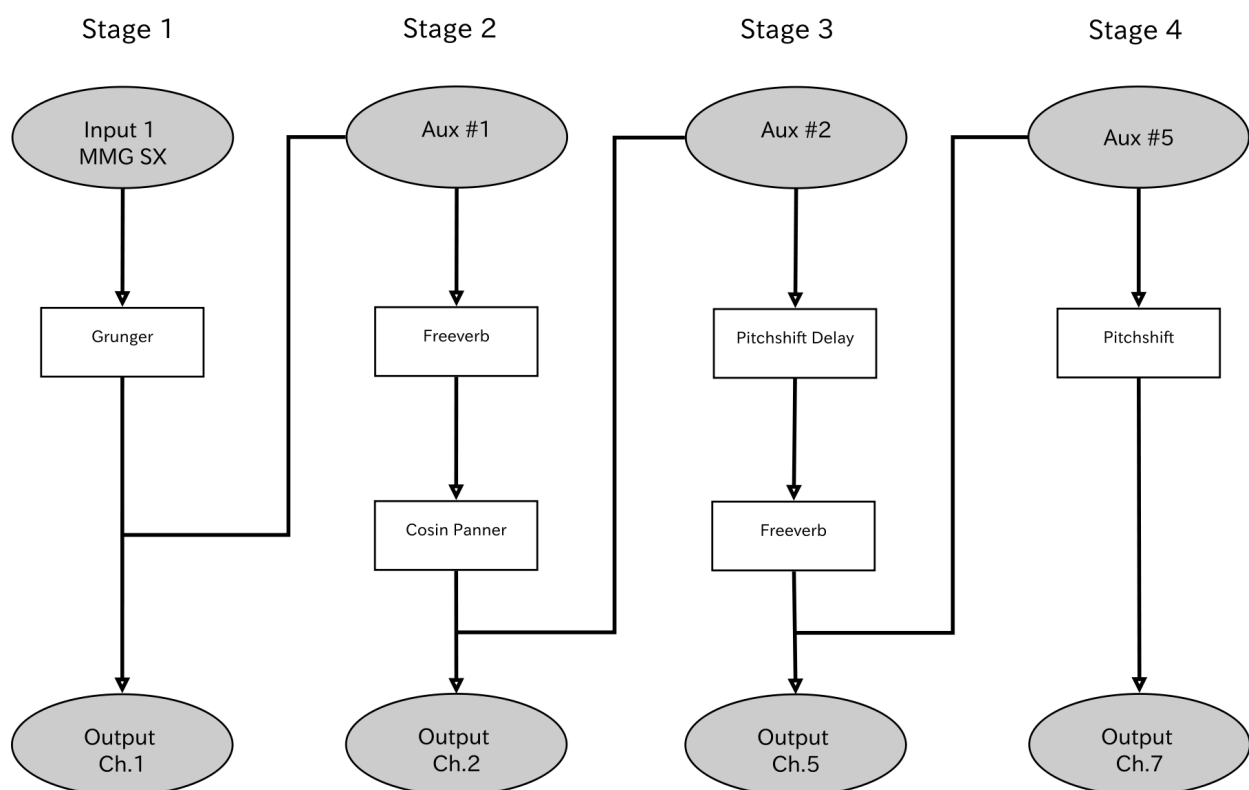


Figure 2: Block diagram of the four-stage digital signal processing system used in *Ominous*. Modulation of the input signal in the first stage affects all subsequent stages. The muscle sounds of the left bicep flows through a four-stage digital signal processing (DSP) system (Figure 2), whose parameters are driven by synced contractions of the right

arm posterior muscle. Instead of using a DSP system with one global output, each DSP stage outputs its resulting signal through the sound system. This strategy enables the playful creation of a multi-layered sound flow. The enjoyment lies in that disparate sonic forms can be shaped by coordinating and fine-tuning whole-body gestures, which address one or multiple DSP stages at once. Next, I describe the mapping technique and the signal routing strategy.

Synchronous mapping of two muscle sounds

At the beginning of the piece, the left fist is rhythmically open and closed. The muscle sounds of the posterior muscle, or flexor, is amplified, filtered and distorted. By accentuating the onset of a contraction I can intuitively broaden the biosignal spectrum and modulate the amount of distortion of the higher partials (which sit between 40-45Hz). As the sound texture and colour defined by the first processing unit characterise all the subsequent sound forms, being able to precisely modulate this DSP stage is critical. At the same time, the right arm must be held still in order to avoid the activation of the other DSP processes.

After a few seconds, the torso is bent forward and the right arm is lifted. Then, the hands are slowly brought together, so as to enclose the invisible object. As the torso is bent towards the legs, the muscle sounds becomes louder. The loudness increment is not caused by a stronger contraction of the arm muscles but rather it appears spontaneously because of the coordinated tension of the torso and shoulder muscles, which are now stretched. The muscle sounds, in the form of a deep sound wave, vibrate the bodies of the audience members. The S feature of the left arm is mapped to the narrowness (Q factor) of a resonant filter. A minimal flexion of the fingers drastically reduces the Q factor, and the resonance bandwidth becomes wider. The resulting signal is fed first to a transposition effect and then to a delay line, which eventually outputs piercing high frequencies that 'cut through' the sonic field with a sweeping movement.

The second DSP stage receives the low frequency sound supplied by the first stage, and feeds it to a stereo reverb and a cosine panner. By twisting the left wrist, the reverb size is increased and the sound is spatialised. As the reverb signal increases, it flows through the third DSP stage. Here, the sound is transposed once more and passed to a feedback delay line. The arms are slowly opened maintaining the muscle tension very intensely; a constant contraction force of the right arm produces a high and steady MRA, which in turn, triggers the saturation of the feedback delay. At the same (logical) time, the MRA of the left arm modulates the sound spatialisation.

First, the muscle sounds 'explode' in a grave resonant rumble and then they mutate into a heavy and persistent metallic rattle. The sound density is intense. As the arms are increasingly opened, the sound becomes harsher. The sequence is frantically repeated a few times until the object 'appears' ten times bigger than its initial size. It becomes evident that it can hardly be contained, and soon after, as the hands are completely separated, the object energy is released in the air (see Figure 3). In this moment, the torso is stretched backwards and the muscles are relaxed. The biosignal features fall down to zero and the sound rapidly fades out.



Figure 3: The closing gesture of the first section of *Ominous*. The arms are opened and muscular tension is about to be released. Photograph: FILE / Hypersonica Festival, 2012.

Machine Learning

As a result of the development of a cascade DSP system and a synchronous mapping technique, the sound of the Xth Sense is richer, and performing is more playful. A player can modulate sound density, spectrum, timbre, loudness, timing, and spatialisation by fine-tuning a muscle contraction and coordinating the tension of the adjacent limbs. On one hand, the use of the subtle nuances of whole-body muscular tension enables a higher degree of freedom in performance. On the other, this makes the Xth Sense more difficult to play. The challenge does not lie in the player's technical skill, but rather in the unpredictability of the body. Although a player can acquire remarkable skills by training, the body is subjected to autonomic processes (involuntary physiological processes), which affect the performance. For instance, the loudness of the muscle sounds can become very unstable, even when cleverly modulated by the performer. During a performance, the veins often bloat with blood because of the increasing heart rate. As the sound of the blood flow becomes louder, quieter muscle sounds become less audible; this flattens the overall loudness and should therefore be avoided (unless it is caused on purpose). In this case, the forearms should be lifted and brought close to the chest, for this position makes the blood flow down the arms and lets the veins recover their original shape. By combining this movement with a slower breathing rate, within about ten seconds the body enters a state of relaxation and then it is possible to play normally again. Despite the fairly large amount of time I have been playing with the Xth Sense (at the time of this writing about five years), I still need to train such mechanisms. Training alone does not enable reciprocity between the spontaneous mechanisms of the inner body and a biosignal-driven digital musical instrument. How can the instrument autonomously recognise this given changes of the biological body? Which are the musical strategies that could be enabled by such behaviour?

Instrument adaptation

In the field of biology, the term adaptation is commonly defined as the process of change by which an organism or species becomes better suited to its environment. The Xth Sense can be thought of as an (computational) organism, and the performer's body as the Xth Sense environment. This analogy suggested me to explore the modalities by which the instrument could identify the body's muscular state, and autonomously adapt to it. In the light of this goal, some features of the Xth Sense were reconsidered. In the previous compositions, the temporal structure was controlled by a timeline. Triggers located at key points in time loaded a given pre-set scene; the temporal structure of the piece and the actions of the performer on stage were completely independent. As long as one could memorise cue points by rehearsing regularly, this approach would have proved functional. However, it became evident that the lack of agency over the structure of the piece was a cause of distress during performance. This kind of interaction was not a viable, long-term solution.

Following a conversation about Machine Learning (ML) with artist and researcher Ben Bogart,³ this area of study became integral to my work. ML is a branch of Artificial Intelligence (AI). It is the design of algorithms that enable a computer to identify and 'learn' generic patterns within large datasets (Bishop, 2006). ML is currently used in a number of different fields, such as robotics, bioengineering, computational finance, videogame, and music performance. During a first test where the Xth Sense was combined with the free ML software Wekinator (Fiebrink, 2011), the instrument was able to recognise different muscular states by learning from the muscle sounds features. This prompted the development of an integrated ML system for the Xth Sense. This is based on Artificial Neural Networks (ANN). An ANN is a mathematical model inspired by biological neural networks, in which a group of interconnected artificial neurons decodes similarities and establishes statistical relation among streams of data. The details of the ML algorithm are not elaborated here as they would not fit the scope of this article. The interested reader can refer to the work of (Caramiaux, 2013) for a comprehensive review. The next section describes how the integrated learning system was applied to the performance of *Ominous*.

Identifying muscular states

There are several ML algorithm types and each of them has advantages and restrictions.⁴ In this case, the choice was to work with ANN models based on supervised learning. During supervised learning the computer analyses training examples offline. Each example consists of input data and a desired output, which is indicated by a label. By training multiple times, the algorithm defines and generalises the patterns, which correlate the input data and the output label. In this way, the algorithm learns recurrent behaviours that can be identified afterwards in real time. Using this method, the previous timeline was extended to make the Xth Sense capable of responding to the changes in the muscular state of the player.

The application of this method in the context of the Xth Sense is structured as follows. First, a player executes the performance gestures, and the instrument monitors the different muscular states offline. The features extracted from the muscle sounds (N, S, L, T, and MRA) are fed to the ANN. This identifies four different states, which are labelled: still (no activity), slow (low activity), moving (medium activity), and fast (high activity). During a real-time performance, the Xth Sense 1) compares the stream of muscle sounds features with the patterns it has learnt offline, 2) detects the current muscular state, and 3)

outputs the related label (see the code in Figure 4). Key points are then added to the timeline in order to indicate the change of the scene being played. As the time passes and a key point in time is reached, the computer stands by until it detects that the player is still; only then a new scene is loaded and specific DSP chains are activated or deactivated. When a new scene is triggered, the Xth Sense automatically fades out the volume of the sound output and restores it after a few seconds. This method allows one to bypass unpleasant clicks and artefacts caused when DSP arrays are switched on and off.

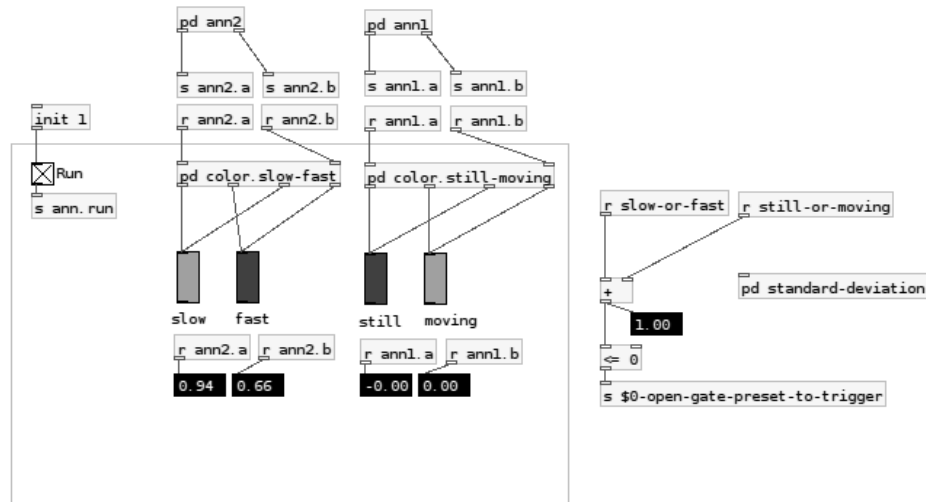


Figure 4: The Xth Sense integrated Machine Learning system. It uses a digital neural network based on supervised learning to identify four muscular states of the performer's body in real-time and configure the instrument accordingly.

Conclusion

In the performance with digital musical instruments, the digitisation of a performer's movement is the first link between a performer and a computer that needs to be established. The way by which such a link is created determines the subtleness and playfulness of interaction with the instrument. Questions on how physiological data can represent the performer's physical movement and the expression it could convey are aesthetic and technical challenges that are at the core of musical performance with biosignal-based instruments and digital instruments in general. In Joel Ryan's words, '[e]ach link between performer and computer has to be invented before anything can be played' (Ryan, 1991: 5). Indeed, the abstraction of a computer system has to be confronted with the physicality of musical performance to achieve a playful musical interaction.

By continuously adapting its configuration to the emergent physiology of the performer's body, the Xth Sense leaves to the player the challenge and the pleasure of delivering an exciting musical experience. A fascinating outcome lies in that the instrument can adapt to different players. Using this as a starting point, it is possible to imagine a future development where the gesture-to-sound mapping is not pre-programmed, but rather generated by the instrument according to the specific traits of a player's performance style. In its current state, the Xth Sense does not have this capability. In order to develop such a system, a more complex computational representation of a physical gesture is being investigated. To that end, on-going research is focusing on creating a multimodal musical system that combines muscle sounds with the electromyogram or EMG, an electrical biosignal produced by muscle tension. The muscle sounds and the EMG signals provide

complementary information on movement articulation that consistently represents a physical gesture (Caramiaux et al., 2015). By extracting and selecting complementary features of the two biosignals, it is possible to obtain a detailed computational model of a physical gesture. The model can then be used to identify, by means of unsupervised or reinforcement machine learning, particular aspects of a physical gestures. Such a system could learn how to differentiate a given player from another and then create personalised gesture-to-sound mappings that the player can explore, evolve and manipulate, and even 'break' simply through physical engagement.

In this scenario, music would be the result of both the player's and the instrument 'intention'. Perhaps in this way we would be able to move 'beyond pushing buttons and activating sensors [...] isolating gestures and mapping data' (Waisvisz, 2006) towards new strategies of musical interaction with technology; strategies where the physicality of a performer can affect the configuration of a musical instrument, and vice versa, the configuration of an instrument can affect the physicality of the performer.

Acknowledgement

Ominous is homage to artist Alberto Giacometti. The piece is an interpretation of a recurrent topic in his work, that of 'a constant irrational search and movement towards an unknown object'.⁵ This theme is embodied in the threatening, bronze-casted sculpture *Hands Holding the Void*, which is the inspiration for this performance.

Notes

1 <http://xth.io>

2 <http://marcodonnarumma.com/texts>

3 <http://www.ekran.org/ben/wp>

4 An exhaustive report of all existing ML algorithms is beyond the scope of this text. The interested reader might look at: supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning, transduction and multi-task learning.

5 This quote is taken from a bulletin of the St. Louis Art Museum (US) in which the author (apparently unknown) rephrases a description allegedly attributed to André Breton. See (Saint Louis Art Museum, 1967: 2).

References

Bishop, Christopher M. (2006) *Pattern recognition and machine learning*. New York: Springer.

Bolton, C. F., Parker, A., Thompson, T. R., Clark, M. R., & Sterne, C. J. (1989) 'Recording sound from human skeletal muscle: Technical and physiological aspects', *Muscle & Nerve*, 12(2): 126–134.

Caramiaux, Baptiste and Tanaka, Atsu (2013) *Machine Learning of Musical Gestures*. Proceedings of the International Conference on New Interfaces for Musical Expression, pp. 513–518. KAIST: Seoul.

Caramiaux, Baptiste, Donnarumma, Marco and Tanaka, Atau (2015) 'Understanding Gesture Expressivity through Muscle Sensing', *ACM Transactions on Computer-Human Interactions*, 21(6).

Donnarumma, Marco (2011) 'Xth Sense: A study of muscle sounds for an experimental paradigm of musical performance', *Proceedings of the ICMC, International Computer Music Conference*. Huddersfield.

Donnarumma, Marco (2012) 'Incarnated sound in Music for Flesh II. Defining gesture in biologically informed musical performance', *Leonardo Electronic Almanac*, 18(3): 164–175.

Donnarumma, Marco (2012a) 'Fractal flesh - alternate anatomical architectures: Interview with Stelarc', *eContact! Biotechnological Performance Practice / Pratiques de performance biotechnologique*, 14(2).

Donnarumma, Marco and Tanaka, Atau (2014) 'Principles, challenges and future directions of physiological computing for the physical performance of digital musical instruments', *Proceedings of the Conference on Interdisciplinary Musicology*. Berlin.

Fiebrink, Rebecca A. (2011) *Real-time Human Interaction with Supervised Learning Algorithms for Music Composition and Performance*. Ph.D. Thesis. Princeton University.

Massumi, Brian (2002) *Parables for the Virtual: Movement, Affect, Sensation*. Durham: Duke University Press.

NIST/SEMATECH (2003) *e-Handbook of Statistical Methods*. Retrieved from <http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc431.htm>, accessed on 21 January 2015 at 16:52 GMT.

Orizio, C., Perini, R., Diemont, B., Maranzana Figini, M., & Veicsteinas, A. (1990) 'Spectral analysis of muscular sound during isometric contraction of biceps brachii', *Journal of applied physiology*, 68(2): 508–512.

Ortiz, Miguel (2012) 'A Brief History of Biosignal-Driven Art From biofeedback to biophysical performance', *eContact! Biotechnological Performance Practice / Pratiques de Performance Biotechnologique*, 14(2).

Oster, G., & Jaffe, J. S. (1980) 'Low Frequency Sounds from Sustained Contraction of Human Skeletal Muscle', *Biophysical Journal*, 30(1): 119–127.

Rosenboom, David (1976) *Biofeedback and the arts: Results of early experiments*. Vancouver, BC: Aesthetic Research Centre of Canada.

Ryan, Joel (1991) 'Some remarks on musical instrument design at STEIM', *Contemporary Music Review*, 6(1): 3–17.

Saint Louis Art Museum (1967) 'Recent acquisitions', Available at: <http://www.jstor.org/stable/40715426?origin=JSTOR-pdf>, accessed on 21 January 2015 at 16:52 GMT.

Tanaka, Atau (2010) 'Mapping out instruments, affordances, and mobiles', *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 88-93, Sydney.

Waisvisz, Michel (2006) Panel Discussion moderated by Michel Waisvisz, 'Manager or Musician? About virtuosity in live electronic music', Proceedings of the 2006 International Conference on New Interfaces for Musical Expression, p. 415, Paris.

Biography

Marco Donnarumma is a performer, sound artist, musician and writer based in London and New York. He uses biomedical and sound technologies, computer software, actuators, sensor and transducers to create ways in which human bodies and machines can extend, transform or disrupt each other. He conceives of both human bodies and machines as materials. His live performances, concerts and installations are renown for combining a seemingly simple and minimalistic aesthetic, rigorous science, technical sophistication, and strong critical concepts.

<http://marcodonnarumma.com>