Rendering embodied experience into multimodal data: concepts, tools and applications for Xenakis' piano performance

Pavlos Antoniadis, Stella Paschalidou, Aurélien Duval, Jean-François Jégo, Frédéric Bevilacqua

▶ **To cite this version:**

Pavlos Antoniadis, Stella Paschalidou, Aurélien Duval, Jean-François Jégo, Frédéric Bevilacqua. Rendering embodied experience into multimodal data: concepts, tools and applications for Xenakis' piano performance. Xenakis 22: Centenary International Symposium, May 2022, Athens & Nafplio, Greece. hal-03999834

HAL Id: hal-03999834
https://hal.science/hal-03999834

Submitted on 21 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Rendering embodied experience into multimodal data: concepts, tools and applications for Xenakis' piano performance

Pavlos Antoniadis, Humboldt Stiftung / Foundation; Department of Audio Communication, Technical University of Berlin, Germany; Sound Music Movement Interaction team, IRCAM, Paris, France, antoniadis@ircam.fr – https://pavlosantoniadis.wordpress.com/

Stella Paschalidou, Department of Music Technology and Acoustics, Hellenic Mediterranean University, Greece, pashalidou@hmu.gr – https://mta.hmu.gr/en/

Aurélien Duval, INREV research team, AIAC Laboratory, University Paris 8, France, duval.aurelien@gmail.com – https://aurelduval.xyz/

Jean-François Jégo, INREV research team, AIAC Laboratory, University Paris 8, France, jean-francois.jego03@univ-paris8.fr – www.jfcad.com

Frédéric Bevilacqua, Sound Music Movement Interaction team, IRCAM (Science & Technology for Music and Sound – IRCAM – CNRS – Sorbonne Université), Paris, France, frederic.bevilacqua@ircam.fr – https://www.stms-lab.fr/team/interaction-son-musique-mouvement/

Proceedings of the *Xenakis 22: Centenary International Symposium* Athens & Nafplio (Greece), 24-29 May 2022 - https://xenakis2022.uoa.gr/

Abstract

Iannis Xenakis' performance practice has increasingly been the object of investigation by both interpreters and musicologists. After a first generation of pioneering performers, who attempted to register and communicate what at the time was a singular challenging experience, subsequent generations have systematically kept developing practice-based research methods for learning and playing Xenakis (Kanach 2010). Similarly, musicology has been shifting its attention from Xenakis' structuralist approach (Xenakis 2002) to post-structuralist (Exarchos 2015) and ecological approaches (Solomos 1996), to Xenakis' composition as practice (Gibson 2011) and quintessentially towards performance analysis, in the wider context of a performative turn (Lalitte 2015) and a more recent embodied cognitive turn (Leman 2008), (Besada et al. 2021). The results of this double movement by performers and musicologists have exemplarily been codified in the series of conferences "Interpréter/Performer Xenakis"[1].

The reasons for the constancy of this research interest might be located in a specific trait: Xenakian performance presents us with a unique manifestation of the 'mind-body problem'. In accordance with contemporary empirical and psychological studies in performance analysis and education (Clare and Cook 2004), (Parncutt and McPherson 2002), we assert that ancient distinctions between abstract understanding, performing technique and artistic interpretation can hardly address the emergence in sound of mathematical algorithms and geometrical structures, which are communicated through dense symbolic music notation and are realized with extreme physical investment. This

1 https://musidanse.univ-paris8.fr/spip.php?article1519

tension between Xenakis' conception, notation and performance is usually framed in terms of impossibility, meta-/anti-virtuosity, athleticism, energetic striving and effort (Varga 1996) or of utter disembodiment and deconstruction of bodily reflexes (Thomopoulos 2010).

In this paper, we attempt a paradigm-shifting approach to performance analysis for the fourth generation of Xenakis' performers, considering the relation between notation and embodiment as expressed in multimodal performance data: the mind-body problem is to be addressed as a decoupling in the relation between the musical score, on the one hand, and the multimodal performance data on the other.

We will present a library of data collected over many years, as well as a wide range of applications for learning and performing Xenakis' piano works. We will provide an overview of technological means for capturing, analysing, assisting, augmenting and communicating Xenakis' performance practice, with reference to his three major works for solo piano: *Herma* (1961), *Evryali* (1973) and *Mists* (1980). In particular, we will look at how multimodal data are indispensable in addressing Xenakis' tensions codified above, but also how Xenakis' transcendental performance challenges the very notion of body rendition in itself.

We will address the following axes:

1. Capturing the Xenakian body: Affordances, constraints and invasiveness in system development

2. Analysing the corporeal subtext of Xenakis' notation: Recent work towards evaluating relationships between musical structure and multimodal data, including effort-related EMG

3. Beyond unnecessary challenge: Creating interactive systems for enabling learning through multimodal data

4. Augmenting Xenakis performance: Integration of the data in an augmented reality spectacle communicating latent layers of performance information to the public

Along these axes, a certain biopolitical trait, the notion of body rendition through data (Zuboff 2019), is counterbalanced by irreducible physical elements that "escape computation"[2]. Thus, the ancient friction between mind and body will be reframed in terms of the friction between what can and what cannot be computationally captured and manipulated in Xenakis' piano performance, in order to define new potentialities.

1. Introduction

Sixty years after the Japanese pianist Yuji Takahashi commissioned Xenakis with his first significant solo piano work, *Herma* (1961), this paper aims at updating the artistic and scientific study of Xenakian piano performance practice. On the one hand, Xenakis' music has by now become part of the contemporary music canon: It is widely performed by students and professional artists around the globe, to a degree that exhausts the tropes concerning impossibility, surpassing, athleticism, meta-/anti-virtuosity and piano heroes, still found in a wide range of scholarship and journalism, in stark contrast to Takahashi's elegantly laconic remarks (Takahashi 2008). On the other hand, both conceptual and technological advancements in the study of performance make this update urgent: First, the performative turn in musicology since the 1980s (Cook 2013) has revealed the specificity of performers' approaches to musical works as complex phenomena that live in multiple domains, the symbolic notation and the notion of the 'work' being only one of those.

2 https://www.slomoco.surf/projects/provocations/resource-collection/provocations-events

Second, the embodied cognitive turn in musicology since the 2000s (Leman 2008) has stressed the importance of embodiment in shaping the understanding of the symbolic notation itself, in learning as well as in performance. Finally, the development of technologies for the documentation of musical performance has resulted in a multiplicity of methodologies for studying not only the sonic outcome, but every aspect of the multimodal phenomenon of performance, including image, gesture, movement and touch, especially since the democratization of sensor technologies and interactive music in the 21st century.

What makes Xenakis' music particularly fit as a case-study for the artistic and scientific research of performance is what we defined in our abstract as its 'mind-body problem': On the one hand, Xenakis is effectively transcribing algorithmic and geometric structures in conventional music notation, often codified in his original complex calculations and compositional sketches. This fact generates a first decoupling between the conception of the musical work and its embodiment in notation as the main communication interface between the composer and the performer. On the other hand, the very structures themselves are consciously driven by Xenakis' own predilection for athleticism, extremes of energy expenditure and effort. This aspect invites the form of performative approach that has been defined as 'energetic striving' (Cox 2002), and even more so when the notated structures are occasionally not taking into consideration the biomechanics of human performance and the construction of the instrument. Thus, a second form of decoupling emerges, between the embodiment of the performer and the embodiment of concepts in notation. This decoupling needs to be consciously addressed by the performer, without loss of the expressive quality of transcendence.

In what follows, we propose to transpose the question of these decouplings onto the decoupling between notated symbols and captured multimodal data in performance. The reason for this is that both forms of recording, in notation and in data, are fixed in their respective media, which allows for their qualitative and quantitative analysis and for their effective communication to third parties other than the performer and the composer. Thus, the usual first-person, subjective descriptions offered by the performer can be transformed into third-person, objective ones.

However epistemologically sound such an approach might be, it reveals several aporias: First, on the developer end, the media for capturing performance are often constrained by numerous factors, which will be explored in detail. Second, on the user end, such media always require performance trade-offs depending on their body invasiveness during the act. Third, this aspect becomes even more pronounced in Xenakis' performance, with its very particular set of biomechanical problems, which take us to the domain of 'extreme users in extreme situations'[3]. Fourth, the combination of these media and the potential range of applications present their own challenges, which are to be addressed not only at a properly technical level, but also at a conceptual and methodological level, necessitating very careful rethinking of performance/interpretation in general and in Xenakis. Finally, the attempt to capture embodiment in data bears a certain biopolitical trait, what the American sociologist Shoshana Zuboff has termed 'body rendition' (Zuboff 2019), and which in an age of acute data surveillance and 'cognitive capitalism' (Moulier Boutang 2007) cannot remain unexamined: To what extent are the elements that "escape computation"[4] still desirable and what are the limits and ethical considerations of intimate body data being studied and shared?

Our paper is articulated as follows: First, we present a brief chronology of systems' development for studying Xenakis' piano performance. Then, we address the theme of affordances, constraints and issues of invasiveness in system development (section 3). Subsequently, we offer an overview of methodologies for the study of three Xenakis piano works, *Herma* (1961), *Evryali* (1973) and *Mists* (1980), from a computational musicology point of view (section 4). In the fifth section, we provide an overview of the *GesTCom* system (acronym standing for *Gesture Cutting Through Textual*

3 https://www.inria.fr/en/ex-situ

4 https://www.slomoco.surf/projects/provocations/resource-collection/provocations-events

Complexity), as a concept and tool for technology-enhanced learning and performance in Xenakis. Finally, we present an augmented performance of *Evryali,* employing motion capture, augmented reality and interactive staging.

2. Chronology

This work has followed four distinct chronological stages, corresponding to material/technological and conceptual/methodological advances:

A first phase of development of a prototype system called *GesTCom* (acronym for *Gesture Cutting through Textual Complexity*) took place at IRCAM (2014). The goal of the prototype was to enable recording of movement data of a piano player, simultaneously with audio. It was developed with the constraints of being easy to set up, transportable and affordable to be duplicated by performers. A first prototype was then built using a *Kinect* camera (colour and depth images) and two wireless custom-made 3D accelerometers and 3D gyroscopes, based on the Modular Music Objects (Rasamimanana et al. 2010), attached to the wrists. The software was built using Max/MSP[5] and particularly the package MuBu[6], enabling advanced recording and data processing. This setup was evaluated and used for recording several pieces, and in particular Xenakis' *Herma* and *Mists*. Some of the preliminary results are presented in (Antoniadis et al. 2014).

A second phase of system development and recording was achieved at the University of Strasbourg (2014-2018), in collaboration with IRCAM. The use of capacitive sensing via *TouchKeys*[7] (Augmented Instruments Laboratory) was added in order to capture the position of the hand and fingers on the keyboard. The software was also updated to visualize MIDI information. Next to these hardware and software developments, a methodology for the mappings between notation and movement and for notational processing on the basis of movement was developed in (Antoniadis and Bevilacqua 2016).

A third phase of development including full-body motion capture, augmented reality systems and interactive scenography for the enhancement of the spectator's experience has been taking place since 2020, at EUR-ArTeC, Université Paris 8, in collaboration with the labs MUSIDANSE and INREV-AIAC (Antoniadis et al. 2022).

Finally, in the context of independent collaboration with Stella Paschalidou, we combined some of the *GesTCom* capabilities with full-body motion capture and EMG recordings, in an experiment of registering and tracking multiple datasets, with the intention of studying mappings between notation and effort.

3. Capturing the Xenakian body: affordances, constraints and invasiveness in system development

In this section, we present an overview of themes that will be detailed in the following ones along three axes: affordances, constraints and invasiveness of the systems in question, both on the developer's and on the user's ends.

Affordances and *constraints* are terms used in ecological psychology (Gibson 1979) to describe the sort of action and perception that a given environment allows to (or affords for) a given organism with given abilities.

In developing systems for the capture, analysis and augmentation of Xenakis' piano performance, we needed to make choices concerning a series of issues and themes:

First, the sort of modalities and data that are to be captured. Given the democratization of sensor

5 https://cycling74.com/

6 http://forumnet.ircam.fr/product/mubu-en/

7 https://touchkeys.co.uk/

technologies and multimodal data methodologies since the 2000s, there is an abundance of systems that can be distinguished according to the modality captured (image, sound, inertial data, location on a surface, movement around a space, etc).

In the following case-studies, combinations of heterogeneous capture systems are used that afford a variety of modalities and performance data. Variety ensures the possibility of capturing multiple aspects of the performance phenomenon, but is constrained by both specific limitations of the capture systems, as well as the limitations of their combinations. The most notable one is the need for aligning the heterogeneous data-streams spatially and temporally and segmenting continuous data into discrete events and actions, or inversely, deducing higher-order structures from discrete data (MacRitchie and McPherson 2015).

Other important issues concern the question of the relationship (technically called *mapping*) between the captured data used as input in a system to output parameters, such as digital audio or, in this case, visual feedback and notation processing, the temporal or spatial structure of the mappings, etc. (Bevilacqua et al. 2011). Such issues are explicitly addressed by the development of the *MuBu* (acronym standing for *Multiple Buffer*) library.

Additional issues regarding measurement concern the inherent complexity of the performance act, notably in terms of biomechanical complexity, which is often analysed in terms of movement coarticulation (Godøy 2011), proximal or distal control of movement fluency (Gonzalez-Sanchez et al. 2019), the action-perception cycle (Leman 2008) etc., and how can this complexity be mapped upon the data-streams, that can be organized in low-level raw data or higher-level descriptors.

Talking about motion capture usually means capturing the whole body movements (generating a virtual skeleton), which can include fingers, but this requires high-end systems. Recent motion capture systems are also able to capture facial expressions in the case of performance capture mostly used in cinema or video-game industries. Motion recording quality is now quite precise but the challenge remains in creating algorithms to extract data regarding expressivity or emotions.

The complexity of the motion capture setup or suits and its invasiveness could influence the performance recording and somehow transform or even bias data. Regarding the general problems of motion capture, it is important to take care of several elements:

- the set-up requires a specific marker-set and calibration according to body measurements. It has to be done accurately in order to avoid offsets;

- if the motion capture is used offline, data synchronization with other data (such as music or video) is crucial to avoid offsets in space and/or time;

- if the motion capture is used online in real-time, the processing requires more computing and also need to be synchronized with other data-flows for the same concern about offsets.

In any case, the question of spatial and temporal alignment of multiple datasets has to be addressed.

All these aspects are accentuated by the extremity of Xenakis' performance practice in terms of energy expenditure, speed and intensity, what has been termed the *athleticism* of Xenakis' piano performance (Kanach 2010), so that the usual constraints of the systems on the users' end need to be rethought as double constraints and be treated accordingly.

4. Analysing the corporeal subtext of Xenakis' notation: mappings between notated structures and data

In this section, we present an overview of previous and ongoing work towards the evaluation of relationships and technologically-enabled mappings between musical structure and multimodal data from a computational musicology point of view. Multimodal data refers to a variety of captured inputs, including the *GesTCom* datasets (Antoniadis 2018), as well as effort-related EMG

(Paschalidou et al. 2016). Methodologically, we explore ways of discovering and comparing patterns between multimodal and symbolic notation data, visualizing their couplings and decouplings. In that sense, we seek to systematically explore what the German musicologist Martin Zenck has described as the 'corporeal subtext' of music notation (Zenck 2006).

4.1. Mists

In (Antoniadis 2018) we have proposed a methodological framework for the analysis of texture and form through multimodal performance data. We pursued the analysis drawing on Xenakis' distinction between *outside time* and *inside time* structures as exposed in (Xenakis 2002).

As far as the outside time structures are concerned, three types of algorithmic processes are defined by Xenakis: linear random walks, non-linear random walks / stochastic distributions and arborescences undergoing geometric rotations in time-space. We recognized these structures and we mapped them onto distinct types of texture and physical movement, captured and analysed through the system and methodology described in (Antoniadis and Bevilacqua 2016).

Inside time structures were consequently defined on the basis of the multimodal data analysis, revealing middle-ground structures in relation to movement and offering a performer-specific interpretation of the global musical form.

Finally, we identified a third temporal category, *inside learning time*, and we proposed types of musical score navigation, drawing on a typology initially developed in (Stefanou and Antoniadis 2009). Four dimensions of learning (*scanning*, *stratification*, *resistance to the flow* and *line of flight*) are identified and studied through multimodal data.

In the following video[8], the first two steps of our methodology are demonstrated: the recording and analysis of multimodal data through the *GesTCom* system. The video presents the playback of multimodal data, recorded during a performance of Iannis Xenakis' *Mists*, page 1. Figure 1 presents an annotated snapshot of the Max/MSP patch used for the reproduction and annotation of the data.

From top to bottom in the blue panel on the right side of the figure, there is a visual representation of the following datasets:

• Stereophonic audio;

• Twelve gestural signals from inertial sensors on the player's wrists. 3D acceleration data are shown with black, green and blue signals, for both the left hand ('LH ACCEL') and the right hand ('RH ACCEL'); 3-axis gyroscope data are shown with red, yellow and orange signals for the left hand ('LH GYRO') and the right hand ('RH GYRO');

• MIDI information from the keys and the two piano pedals. Colour coding indicates velocity;

• Capacitive data from *TouchKeys* sensors on the keys of the piano. Colour coding indicates the position of the finger on the key, clusters are traces of hand-grasps.

The black markers superimposed over these datasets indicate a segmentation, which is defined by both quantitative and qualitative data according to the methodology in (Antoniadis and Bevilacqua 2016).

The quantitative data used here are the orange gyroscope signals for both hands and the TouchKeys clusters. The orange gyroscopic signal peaks (shown with red circles in Figure 1) indicate two of many hand position changes, or hand displacements, which are visible in the video at the moment the red cursor crosses over the thick marker (00:29-00:30 in the video). This marker is thick because both hands are displaced towards the bottom of the keyboard, as opposed to thin markers, whereby only one hand is displaced.

8 https://www.youtube.com/watch?v=io9iGpVUAkI

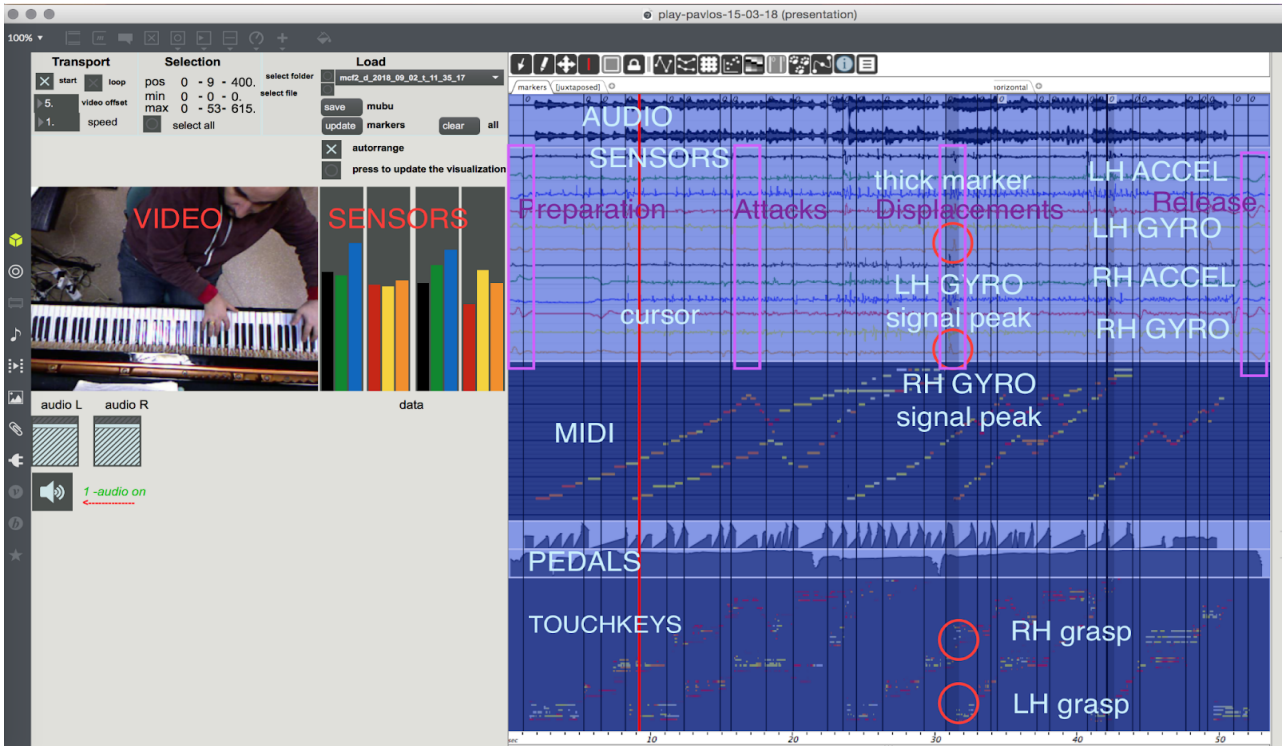

Figure 1: Max/MSP patch for the reproduction and analysis of recorded multimodal data (audio, movement, video, MIDI, TouchKeys)

The qualitative data used are performer's annotations. There is also a qualitative comparison of the gestural data to the *Kinect* video (left side in Figure 1), for visually confirming the displacement of the hands.

This pattern-matching between gyroscopic signals, TouchKeys data and annotated hand-grasps is indicative of the fact that subjective annotations can have an objective expression in multimodal data. In (Antoniadis and Bevilacqua 2016) we have described a syntax of piano movement derived from multimodal data. This syntax takes the form of 'movement envelopes' consisting of movement Preparation, Attack, Displacement and Release phases. We called them 'PADR envelopes'. In Figure 1, purple boxes indicate instances of these types of movement. Practically, attacks are accelerometer activations between the gyroscope activations that indicate displacements. So, in this particular example, attacks take place in-between the markers of displacements and the preparation and release gestures are visible before and after the sound (Figure 1).

4.2. Herma

The previous methodological framework is expanded through a current multimodal analysis of *Herma*, based on Xenakis' analysis of the piece as exposed in *Formalized Music* and adapting the theoretical framework of symbolic music to the study of physical movement.

At a first stage, we proceed with a computation of movement density, expressed in terms of displacements (the 'D' of the PADR envelope in Figure 1) per metrical unit of time in relation to the calculated pitch densities by the composer.

At a second stage, we have proposed a visualization and alignment of Xenakis' deep algorithmic structural elements, summarized by the composer as a *function flow chart* and *temporal flow chart*

(Xenakis 2002, 176), with recorded multimodal data as in this video[9].

At a third level of abstraction, we propose symbolic logical operations for movement as an equivalent to Xenakis' treatment of pitch and we explore relations and mappings.

Next to studying the relation of movement and symbolic notation, we aim at defining measures of notational and movement complexity, to be employed in a series of experiments on human technology-enhanced learning.

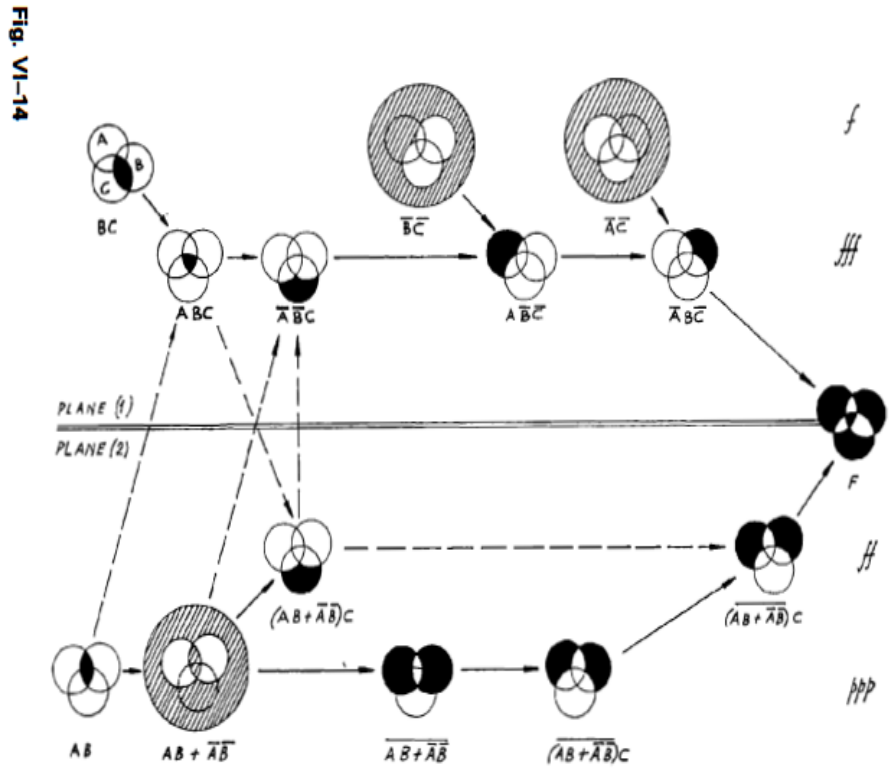

Figure 2: flow chart that directs the logical operations on two parallel planes in Xenakis' Herma (Formalized Music, p. 176)

4.3. Evryali

In more recent work, we have attempted to address issues of virtuosity and performability with respect to physical effort and energy expenditure in the context of *Evryali*. Associations of bodily and mental effort to movement features, but mostly to pitch-related information of melodic glides

9 https://youtu.be/L8pAjnAiQ9E

that invite higher levels of physical power by the performer have been found in Hindustani vocal music by (Paschalidou et al. 2016). Virtuosity in Xenakian terms has been attributed by (Solomos 1996) to measures of 'pure consumption of physical energy' rather than number of 'wrong' notes, ‘sheer physical pressure and transcendence of the performers’ limits’ by (Varga 1996). Therefore, we raised concerns about whether the notorious difficulty and even near-impossibility in performing specific passages of *Evryali* (Antoniadis et al. 2022), due to sensorimotor constraints that the textures of the dense and complex graphical notation request the performer to surpass, might be associated with equivalent levels of action power or not. For this reason, we ran a pilot study for recording and analysing surface electromyography data, that is electrical potentials produced during synergistic muscle contractions, which are supposed to reflect levels of muscle-activation or force-related information exerted by body joints.

Recordings were conducted in an ecologically valid setting at the music department of the Aristotle University in Thessaloniki, Greece and they included the acquisition of the following separate data streams for capturing motion, audio, video and EMG data on three different computer systems:

- Full-body IMU (inertial) motion tracking of a total of 32 sensors at ~60Hz (Perception Neuron v2 by Noitom), including hands/fingers, data transferred over wifi through dedicated router;

- 3 synchronized RGB video recordings of the performance scene at 60Hz each, 2 from the side and one from the top (PS3 Eye cameras by Sony, USB);

- 2 EMG & IMU recordings, EMG data running at 200Hz and IMU at 50Hz, one for each hand / wrist (Myo Armbands by Thalmic Labs, data transferred over bluetooth);

- 1 depth-video recording of the performance scene / top-view at 22.8.Hz (Kinect for Xbox 360 by Microsoft/PrimeSense, USB);

- 2 individual IMU sensors at 100 Hz, one for each hand (RIot Bitalino by PLUX wireless biosignals S.A., data transferred over wifi through dedicated router);

- 2 separate stereo audio recordings at 44.1 kHz - 16bits each, one for close miking and one for ambient recording (1. A pair of condenser microphones DPA ST 2011C / audio interface RME Fireface UFX - firewire / synchronized audio-bitalino-kinect recording in MaxMsp7/patch based on the Mubu object, 2. Sony PCM-D50 with onboard mics), comprising a limited version of the GesTCom set-up presented before.

Multiple IMU data were recorded by different devices as a means to visually align them and thus cross-verify the clap-based data synchronization. For instance, IMU data of the PN system can be aligned to the IMU data of the Myo Armbands, thus also EMG data to IMU.

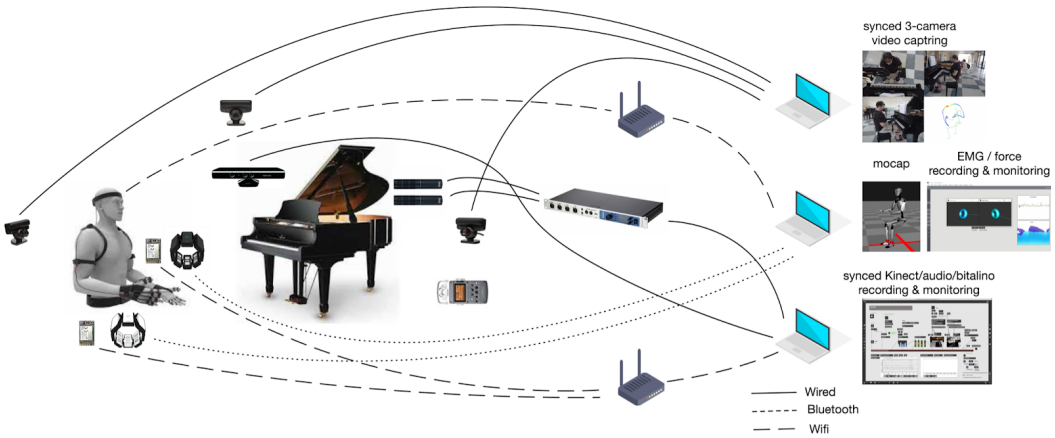

Figure 3: Hardware setup of pilot study

Several technical challenges were raised during these recordings, with a potential effect on the temporal and spatial resolution, data reliability and the reproducibility of the study.

Foremost, we were concerned with issues of synchronizing independent data streams of varying sampling rates for the time-critical analysis of combined multimodal information. As hardware solutions for synchronization (for either common time-stamping or forcing hardware devices to stay locked to a central clock or reference signal) are rarely at hand in music research, this mostly refers to two things:

- Successfully triggering common start- and end-marking points manually (by hand claps) to manually synchronize different data streams in post-processing. This means in practice producing clear signalling points that will allow to accurately trim individual data streams;

- Keeping the different data streams aligned over time. This means avoiding different data streams to progressively drift apart from each other over the duration of each recording, which is a problem that may arise by either unstable sampling rates or individual frames being dropped, and which commonly results in a mis-match of time lengths between recordings of different types of data.

Such issues arose during the recording process and led to the necessity of manual post-production work. These were mostly due to the unstable wireless PerceptionNeuron-pc communication as well as occasional frame drops of the three HD-video streams being recorded on a single computer. This led to tedious manual post-production work.

Despite advantages of ecological validity in using a IMU motion capture system which requires a less controlled environment than a passive optical marker-based motion capture system, magnetic interferences from the metal construction of the building and the piano led to occasional spatial drifts of joints and thus partially unreliable position data.

Another challenge we faced was related to the PN system's affordances and the constraints it imposed on the performer's finger movements, which were considered unacceptable for performing the work of *Evryali*. This led to coming up with original, less obtrusive (or at least destructive) solutions in using flexible gloves for attaching the sensors on the fingers (Figure 4).

As much as posing technical challenges, the use of an IMU system for recording motion data was opted against the alternative of an optical mocap system due to its flexibility, ease of setup, portability and lack of optical occlusion.

Figure 4: Flexible textile gloves for finger sensor placement

Finally, there were decisions to be made with respect to the calibration and normalization of EMG data based on minimum and maximum force values, which were addressed by an initial recording phase, whereby the musician was asked to perform a maximum tension (full muscle contraction) gesture versus a release (full relaxation) pose.

Videos and graphs displaying spatio-temporal relationships of multimodal data streams were plotted and were used in an explorative approach in the analysis, which is still in progress. Local spatio-temporal peaks and troughs of force-related data can be visually identified, as those in (Figure 5) for the hands in the first page of *Evryali*, and are cross-evaluated against the composer's complex graphical notation, in which effort-related information of action affordances are embedded in varying densities of arborescence branch sketches.

5. Beyond unnecessary challenge: motion following as basis for interactive notation learning in Xenakis' *Mists*

Beyond the documentation and visualization of the performative act in Xenakis' piano music, and beyond computational analysis, a central application has been technology-enhanced learning and performance optimization. This application resonates with the objectives outlined in the introduction: demystification of Xenakian performance practice and assistance of the learning process in the physical and mental domains. The system *GesTCom*[10] in its current state is conceived as a sensor-based environment for the visualization, analysis and following of pianist's gestures in relation to notation. It comprises four modules, implemented in the form of Max/MSP patches featuring the *MuBu*[11] toolbox and connected to *INScore* scripts[12]:

a) a module for the synchronized recording of multimodal data of a performance (refer to 4.1 of the current);

b) a module for the reproduction and analysis of the data (refer to 4.1 of the current);

10 https://medias.ircam.fr/x2253e1

11 https://forum.ircam.fr/projects/detail/mubu/

12 http://inscore.sourceforge.net/

c) a module for the processing of the notation on the basis of the data;

d) a module for real-time gestural interaction with the notation.

In the last module, we distinguish between a *recording phase* and a *following phase*: in the latter, the system follows variations of a gesture recorded in the former. This module is based on the notion of motion following (Bevilacqua et al. 2011, Bevilacqua et al. 2010), a probabilistic, one-shot Hidden Markov Model architecture that dynamically compares the two gestures, originally implemented in the Gesture Follower[13] (GF). This project's contribution to the original GF is that, through the notion of coarticulation / PADR envelopes, as described in section 4, an expansion of the range of gesture variability is attempted: not the performing gesture itself, but rather a reduced form of it (the PADR envelope) is recorded. The PADR envelope is also mapped on the processed notational representations of the third module, so that the system can follow extreme variations of the performance, both in terms of movement and in terms of notational representation. This notion of performance variability is connected to several hypotheses regarding top-down learning, multimodal feedback and prioritization processes in complex music.

13 http://ismm.ircam.fr/gesture-follower/

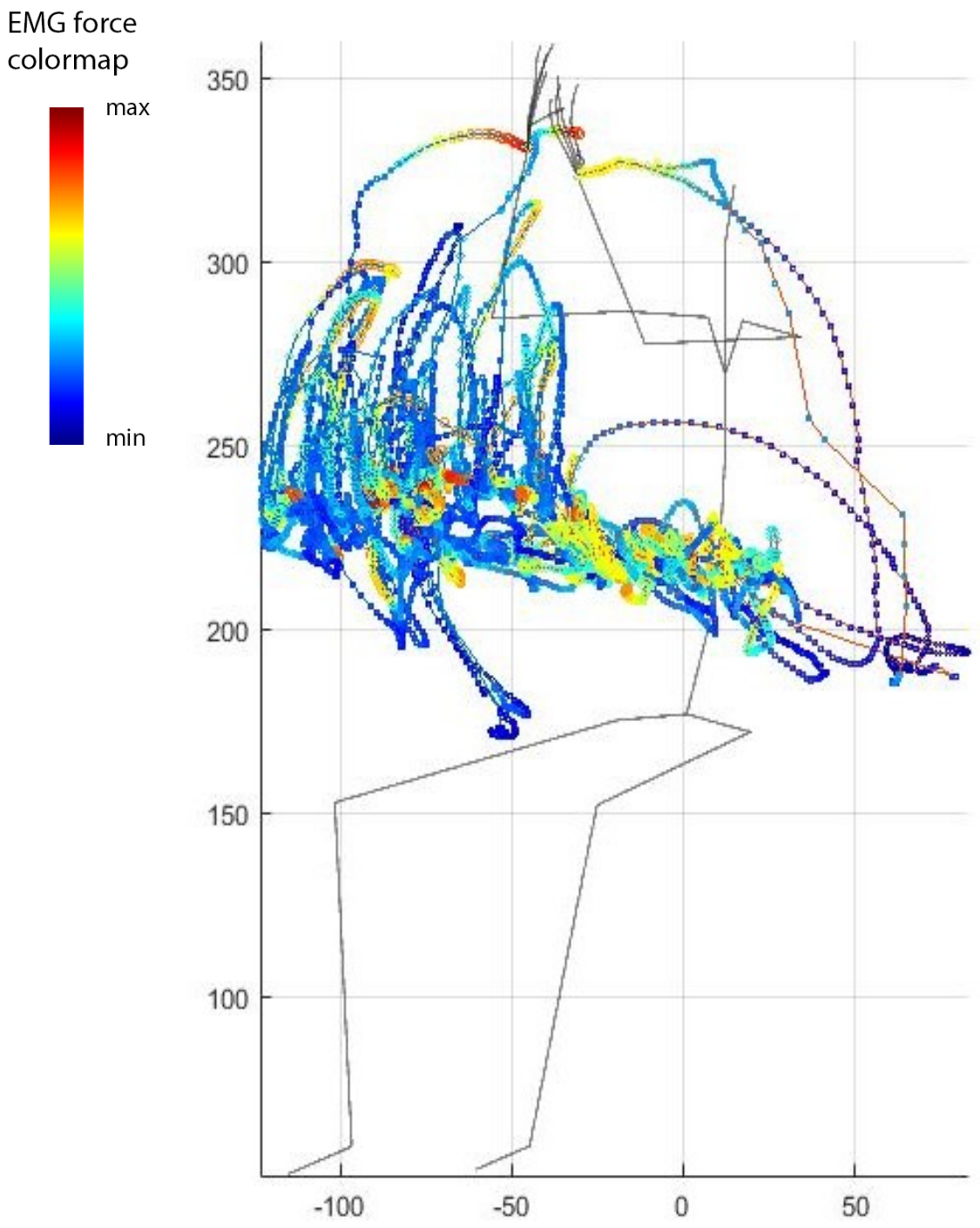

Figure 5: Spatio-temporal colored EMG graph, 3D stick-figure of Evryali's first page

In terms of the end user, the *GesTCom* features the following workflow: In the first step, the performer generates multimodal data through a recording of a performance of the original notation (module 1). In the second step, the quantitative and qualitative analysis of the data, in comparison to the qualitative analysis of the notation, results in a shared segmentation that we described as PADR envelope (module 2). On the basis of this segmentation, the original notation is processed and reshaped into a multilayered 'tablature' (module 3). Eventually, again by virtue of this common segmentation and machine learning techniques, the multilayered tablature can be trained to follow the performer in variations of the initial performance. The performer, that is, interactively controls the tablature through gesture and expressively navigates networks of notational parameters in real-time. Please note that the new output notation can be fed back into this loop, generating new performances, recordings and tablatures. Thus, the following interaction schema emerges:

Figure 6: the GesTCom interaction schema

The further development of the *GesTCom* system seeks to integrate developments from co-adaptive systems employing active learning and reinforcement learning. Co-adaptation[14] indicates that the system and the user adapt to each other during learning. In that sense, the *GesTCom* would adapt to

14 For more on the notion of co-adaptation for extreme users in extreme situations, please refer to the work conducted at the Inria (National Institute for Research in Digital Science and Technology) research lab https://ex-situ.lri.fr/ directed by Wendy Mackay.

user-specific learning pathways and would provide feedback, which would in itself facilitate both the acquisition of motor skills and the deciphering of notation. Interactive reinforcement learning indicates techniques of user feedback to the machine, which improve the process of the co-adaptation. Active learning equally involves the users by querying them to label specific examples picked by the system. The end goal is a system of mutual reinforcement learning between humans and machines. Such a system would optimize the piano performer's learning experience through longitudinal multimodal performance documentation, real-time activity monitoring with augmented multimodal feedback and guidance, and real-time adaptation of the complexity of the music notation, according to the user's developing skills.

Let's now take a closer look at this module in use for Xenakis' *Mists*, as demonstrated in the following video[15]:

It consists of one INScore representation, connected interactively to the motionfollower Max/MSP patch through another Max/MSP patch (indicated in purple as 'INScore', 'motionfollower', 'connector', Figure 7).

An INScore script generates this augmented interactive score, which consists of the following graphic objects: the reduced proportional score representation of the first page of *Mists,* a cursor and a signal, as shown in the Figure above. The reduced proportional notation has been generated automatically by the MIDI data, using command-line tools based on the GUIDO Engine[16] and developed by Dominique Fober. The cursor in red and the signal in blue are controlled through the inertial sensors in the wrists of the pianist. The signal in blue comes from the motionfollower.

The score's interactive possibilities are based on the motionfollower, an object in Max / MSP and a customized patch shown in the next Figure 8.

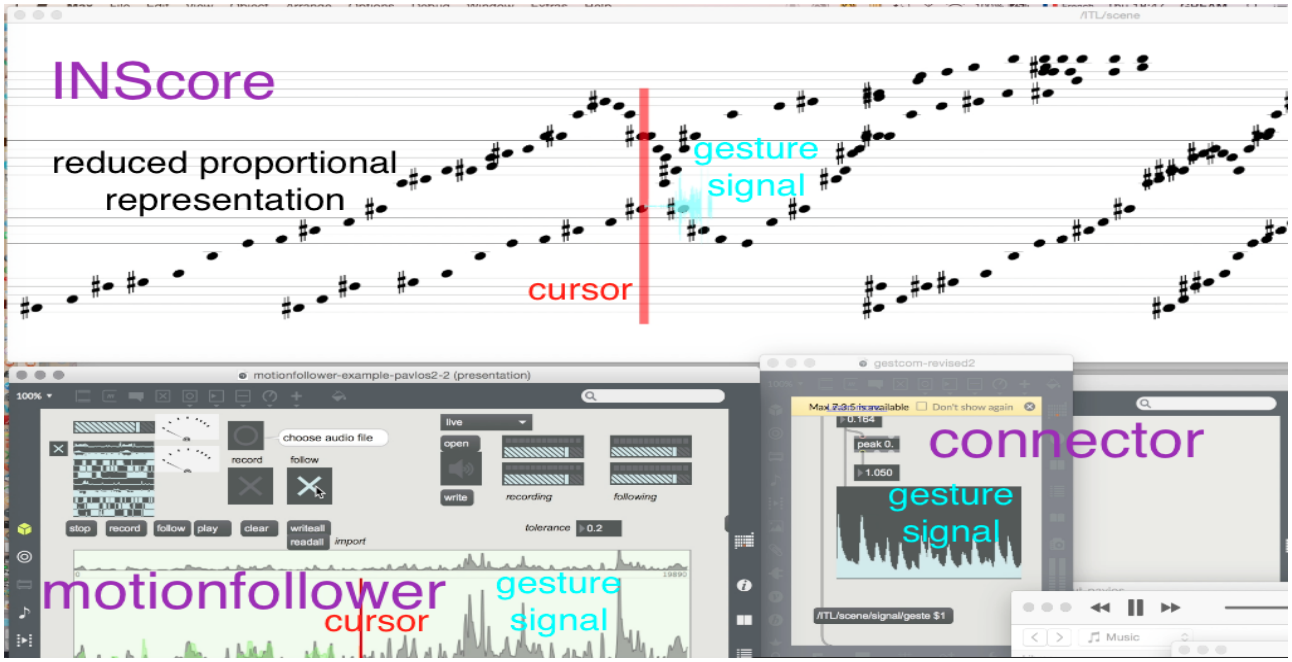

Figure 7: Augmented interactive score controlled through sensors and connected to the motionfollower through a connector Max/MSP patch for Iannis Xenakis' Mists

15 https://youtu.be/Rql732JUm5M

16 http://guidolib.sourceforge.net/doc/guidolib/

In the first phase, a gesture is recorded (as in the video, 00:00 – 00:25). This gesture is represented by the grey signal in Figure 8. This signal is the sum of the twelve signals we saw before in the recording patch (Figure 1), plus audio energy. In a second phase (00:25 – 01:24), this gesture is compared probabilistically to a new, incoming gesture, represented by the green signal in Figure 8, which is superimposed over the grey signal. The system essentially predicts the probability of the new gesture being similar enough to the recorded gesture. If this is the case, the system "follows" the player, and this "following" is indicated by the smooth movement of the cursor. If not, the cursor is moving with a certain viscosity, gets stuck, jumps abruptly forward or is waiting for the performer and so on. On top of the visual feedback, the motionfollower may offer sonic feedback, as the initial recording is of both movement and sound.

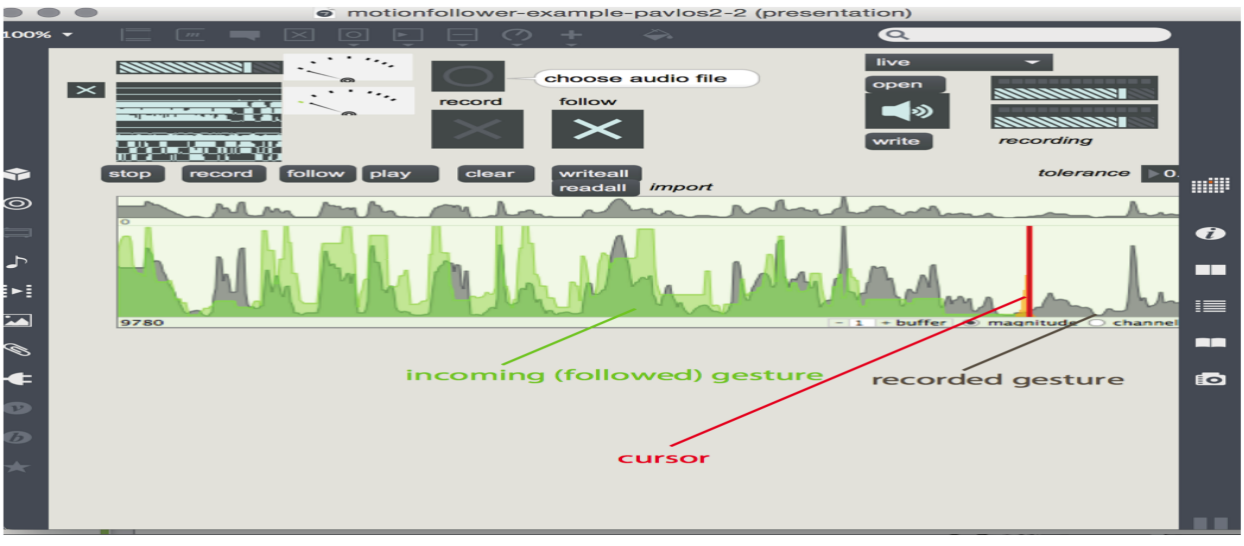

Figure 8: Max/MSP patch for motion following

The third component is a Max/MSP patch, which functions as a connector, sending the incoming new signal of the motionfollower, to the INScore tablature in the form of Open Sound Control[17] messages.

The crucial element, which allows for the motion-following to be reflected in the notation and thus become score-following, is that both the gesture and the notation are sharing the same basic segmentation.

In the recording phase (00:00 - 00:25), the user follows any mobile element of the INScore, in our case the red cursor, which is set to move at a desired speed, like a classic metronome would do. The musical score has already been graphically segmented and assigned a duration according to a specific INScore space-time formalism (explicit mapping). In this recording phase, the motionfollower learns, so to speak, the mapping from the performer (implicit mapping), who follows the mapping of the INScore (explicit mapping). In the 'following' phase, the performer can pursue highly varied performances: a faster performance (00:25 – 00:45 in video), a faster performance with softer dynamics (00:45 – 01:02), a performance with different (staccato) articulation and even mistaken notes (01:03 – 01:24). This time, it is not the performer that follows the system, but rather the system that follows the performer, given that the segmentation is correct and common in all these varied performances. Thus, the performer may control the mobile elements of the INScore tablature. The feedback of the follower has been extended to score compound representations. The gesture-following has been turned into score-following.

6. Augmenting Xenakis' performance

We entered in 2020 to a fourth phase of exploration crossing multimodal recordings and interactivity. To do so, we choose to work on *Evryali* (1973) by Xenakis for solo piano. The score composed of tree structures sketched on graph paper is considered as "extreme" since it contains multiple arborescences which are a tangle of lines in pitch-time space like a bush or a tree, undergoing rotations, expansions, deformations, etc. The piece raises several challenges for the performer and the audience to handle its complexity:

- The score made of complex and detailed graphics creates an innovative polyphony which evokes the elusive infinity of the sea or the uncontrollable energy of the waves. We proposed to decompose it as a palimpsest presenting different layers of information on a screen: from the original scores, to manual and graphic annotations, to the display of augmented information about the performer's movement and body engagement in a 3D interactive virtual environment;

- To visualize the energy and body engagement of the pianist required to perform *Evryali*, we opt for embodied data visualization in real-time of the pianist movement, focusing on the kinematics of the two hands and the head such as speed/acceleration/jerk visualized in 3D;

- The challenge of visualizing the body from different perspectives was explored by developing an avatar which allows switching interactively and continuously between first-person and third-person point of views. This control is done by a visual artist present on stage;

17 http://opensoundcontrol.org/introduction-osc

- In order to engage the audience[18], we proposed to integrate the data in an augmented reality spectacle with the goal to immerse spectators. On top of this, we offered the possibility to one spectator to come on stage and co-perform with the visual artist.

Regarding technical aspects, Xenakis' sketches were reworked and separated according to the different parts of the piece and are displayed in the main projection on stage.

In order to capture the pianist's performance, we used the Perception Neuron[19] motion capture system made of 32 inertial sensors. The gestures are recorded and reproduced in real-time in the form of digital data that can be viewed on an abstract digital twin (avatar) of the performer. Data was streamed from the Axis Neuron software to Unity 3D, a real-time and cross-platform 3D engine where we developed the gesture analysis and visualization tools described in (Jégo et al. 2019).

The scenography, inspired by the polytopes of Xenakis which are immersive events, is composed of 4 video projections on the stage, the ceiling and the public were designed to create a volume of sensory habitation, reflecting the concept of 'dwelling' as in (Ingold 2000). A large projection at the back of the stage displays the avatar of the pianist in real-time, which allows switching from egocentric to allocentric points of view and viewing the acceleration and velocity vectors of the two hands.

We developed modular interfaces to control the various visual effects, and some of them are participative, allowing a spectator to come on stage and manipulate the controls on 2 touch screens.

The visual effects of the pianist movements intend to dilate the micro-movements of the body, usually invisible to the naked eye, in particular of the hands. In this way, the spectators attend a version of the performance augmented by kinesthetic elements which contribute to making the experience of the concert multisensory or, in a prospective way, transmodal.

Since our artistic goal is to immerse spectators and to reveal elements that are usually invisible, we choose to explicitly show the technological artifice as part of the interpretation: the equipping phase of the motion capture suit, usually hidden, is orchestrated like a ritual. The technological devices and setup are not set back either, but on the side of the stage, highlighting the different actors of the performance, whether the visual artist or the spectator invited to interact.

This augmented performance invites the audience to inhabit a hybrid environment blurring the boundaries between the stage and the public, the performer and his digital avatar, the visual artist and the performer, and the physical and symbolic spaces. This immersive space, echoing the volume of the wide sea, tends here to offer a new form of access to the learned elements of Xenakis, which are multiple, complex and sensitive at the same time.

---

18 Antoniadis, Pavlos, Duval Aurelien, Jégo, Jean-François, Solomos Makis. "Augmented reality for Iannis Xenakis Evryali for piano solo". Online video of a performance and a debate. https://www.youtube.com/watch?v=iOrbrHbkqZc

19 https://neuronmocap.com/?gclid=Cj0KCQjwz7uRBhDRARIsAFqjull6Cext3xeCfXBhiByYwBk0nCf6gq6UvSCTFNa61NLpXXr2atWMCu0aAnxOEALw_wcB

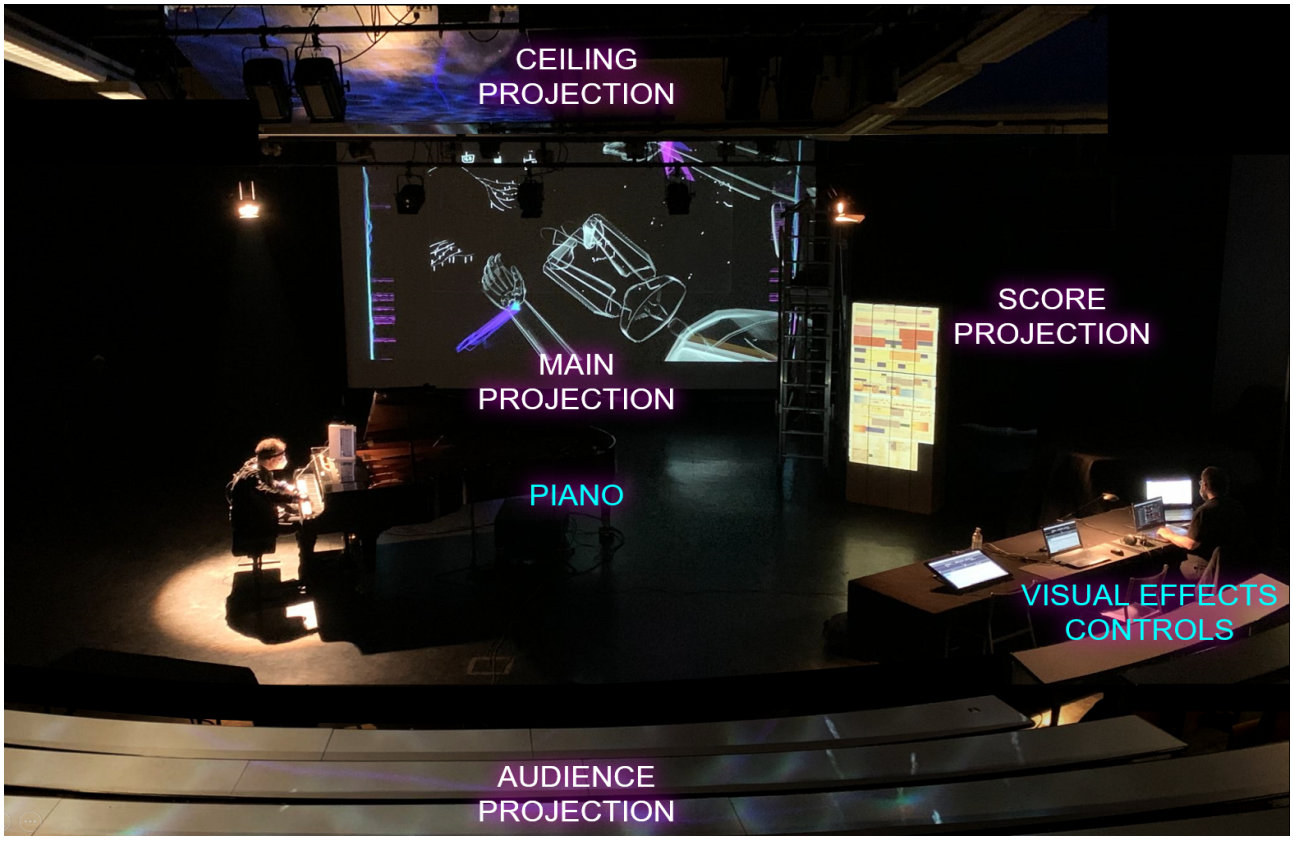

Figure 9: Scenography of the Evryali performance presenting the pianist wearing a motion capture suit, the controls and the video projections

7. Conclusion and Discussion

We have tried to approach what we identified as the 'mind-body problem' in Xenakis' piano performance as a problem of rendering human experience into multimodal data through concepts and tools that have been developed over a period of eight years in the context of different projects. The various technological breakthroughs presented above and stemming in the domain of Human-Computer Interaction are coupled with conceptual breakthroughs in the domains of performance practice, systematic and computational musicology and embodied cognition. These advances aim at demystifying Xenakian performance practice, while simultaneously enhancing, facilitating and communicating the learning process without compromising the transcendental expressive qualities in concert.

The notion of body rendition into data for research purposes opens up a domain of urgent ethical questions. 'Your body is reimagined as a behaving object to be tracked and calculated for indexing and search' (Zuboff 2019, p. 231), even if the end purposes described above might as of yet be far from the usual utilitarianism of manufacturing and sales through behaviour prediction and modification of consumers:

"Rendition describes the concrete operational practices through which dispossession is accomplished, as human experience is claimed as raw material for datafication and all that follows, from manufacturing to sales." (ibid., p.223)

Is it that far, though? Could we imagine a form of economy, where practicing Xenakis through 'smart' equipment could be the key to producing crypto-currencies or where the biometric data of musicians would be traded in online markets (such as NFT collectibles) producing profits for third

parties? Even though such speculation might seem premature, it might be key to designing ethical practices for harvesting musicians' data in the very near future.

The second issue addressed in the introduction was of a different epistemological order: Is it possible that the elements that escape computation are the vital ones in addressing human performance, and in particular the performance of Xenakis' music? What are the limits of digital materialism and is there a danger of missing the particular Xenakian transcendental expression by trying to contain the challenges? A provisional response is that by containing unnecessary effort and challenge, one can channel the performing and expressive energy more efficiently, pushing the envelope of transcendental expression even further. In any case, the only way to address these issues is by taking this line of research to its potential end.

Acknowledgments

We would like to warmly thank: Mâkhi Xenakis for kindly providing access to Xenakis' sketches for *Evryali*; Makis Solomos and Anastasia Georgaki for organising the Athens University recording session (June 2020); Stella Paschalidou, Katerina El Raheb, Christos Lougiakis and Marina Stergiou for their assistance with the motion capture during the Athens University recordings; Stella Paschalidou, Emilios Cambouropoulos and Dimitris Maronidis for organising and facilitating the Thessaloniki University recordings (July 2020); Elaine Chew and the COSMOS project for facilitating the Ircam recordings (September 2020). This project was funded with the support of Université Paris 8 and the National Research Agency - Investissements d'Avenir program – ANR-17-EURE-0008, as well as of the Alexander von Humboldt Stiftung/Foundation.

References

Antoniadis P. and Bevilacqua F. 2016. "Processing of symbolic music notation via multimodal performance data: Ferneyhough's Lemma-Icon-Epigram for solo piano, phase 1". In *Proceedings of the TENOR 2016 conference*, 127-136. Cambridge: Anglia Ruskin University 2016. http://tenor2016.tenor-conference.org/TENOR2016-Proceedings.pdf.

Antoniadis P., Bevilacqua F. and Fober D. 2014. "Gesture cutting through textual complexity: Towards a tool for online gestural analysis and control of complex piano notation processing", in Timour Klouche & Eduardo R. Miranda (eds.), *Proceedings of the 9th Conference on Interdisciplinary Musicology – CIM14*, 236-241, Staatliches Institut für Musikforschung, Berlin, Germany.

Antoniadis P., Jégo, J.-F. Duval A., Solomos M. and Bevilacqua F. 2022 (forthcoming). "Dwelling Xenakis. An augmented reality project on Evryali for piano solo". In *Proceedings of NCMM21 (NOVA Contemporary Music Meeting) International Conference*, Lisbon, Portugal.

Besada J. L., Barthel-Calvet A.-S. and Cánovas C. P. 2021. "Gearing Time Toward Musical Creativity: Conceptual Integration and Material Anchoring in Xenakis' Psappha". In *Frontiers in Psychology*, Vol. 11 https://doi.org/10.3389/fpsyg.2020.611316.

Bevilacqua F., Schnell N., Rasamimanana N., Zamborlin B. and Guedy F. 2011. "Online Gesture Analysis and Control of Audio Processing". In: Solis, J. & Ng K.(eds.), *Musical Robots and Interactive Multimodal Systems*, 127-142. Berlin Heidelberg: Springer.

Bevilacqua F., Zamborlin B., Sypniewski A., Schnell N., Guedy F. and Rasamimanana N. 2010. "Continuous realtime gesture following and recognition". In *Lecture Notes on Computer Science,*

Gesture Workshop, pages 73–84. Berlin Heidelberg: Springer.

Clarke E. and Cook N. (eds.). 2004. *Empirical musicology. Aims, Methods, Prospects*. Oxford: Oxford University Press.

Cook N. 2013. *Beyond the Score: Music as Performance*. Oxford: Oxford University Press.

Cox F. 2002. "Notes Toward a Performance Practice for Complex Music". In Claus-Steffen Mahnkopf, Frank Cox and Wolfram Schurig, eds.: *Polyphony and Complexity, New Music and Aesthetics in the 21st Century*, 70-132. Hofheim: Wolke Verlag.

Stefanou D. and Antoniadis P. 2009. "Inter-structures: Rethinking continuity in post-1945 piano repertoire". In *Journal of interdisciplinary music studies*, spring/ fall 2009, volume 3, issue 1&2, art. #0931205, pp. 77-93.

Exarchos D. 2015. "Listening Outside of Time". In: Solomos, Makis (ed.) *Xenakis. La musique électroacoustique*. Paris: L'Harmattan.

Gibson B. 2011. *The Instrumental Music of Iannis Xenakis: Theory, Practice, Self-borrowing*. New York: Pendragon Press.

Gibson J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton-Mifflin.

Godøy R.I. 2011. "Sound-Action Chunks in Music". In Solis, S. & Ng. K. (eds.) *Musical Robots and Interactive Multimodal Systems*. Berlin-Heidelberg: Springer.

Ingold T. 2000. *The perception of the environment. Essays on livelihood, dwelling and skill.* London-New York: Routledge.

Jégo J.-F. Meyrueis V. and Boutet D. 2019. "A Workflow for Real-Time Visualization and Data Analysis of Gesture Using Motion Capture". In *Proceedings of the 6th International Conference on Movement and Computing*, New York, NY, USA, Association for Computing Machinery

Kanach S. (ed.). 2010. *Performing Xenakis*. New York: Pendragon Press

Lalitte P. 2015. *Analyser l'Interprétation de la Musique du XXe Siècle. Une analyse d'interprétations enregistrées des Dix pièces pour quintette à vent de György Ligeti*. Paris: Hermann

Leman M. 2008. *Embodied Music Cognition and Mediation Technology*. Cambridge, Mas: MIT Press

MacRitchie J. and McPherson A. P. 2015. "Integrating optical finger motion tracking with surface touch events". In: *Frontiers in Psychology*, vol. 6, 2015.

Moulier Boutang Y. 2007. *Le capitalisme cognitif*. Paris: Editions Amsterdam.

Parncutt R. and McPherson G. 2002. *The science and psychology of music performance: creative strategies for teaching and learning*. Oxford: Oxford University Press.

Paschalidou S., Eerola, T., & Clayton, M. 2016. “Voice and movement as predictors of gesture types and physical effort in virtual object interactions of classical Indian singing”. In *Proceedings of the 3rd International Symposium on Movement and Computing* (1-2).

Rasamimanana N., Bevilacqua F., Schnell N., Guedy F., Flety E., Maestracci C., Zamborlin B., Frechin J.-L. and Petrevski U. 2010. Modular musical objects towards embodied control of digital music. In Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction (TEI '11). Association for Computing Machinery, New York, NY, USA, 9–12.

Solomos M., 1996. *Iannis Xenakis*. Mercuès: P.O. Editions.

Takahashi Y., 2008. “Performing Xenakis”. Note in his personal website: https://www.suigyu.com/yuji/en-text/Performing_Xenakis.html.

Thomopoulos S. 2010. “The Olympian Piano: Iannis Xenakis’ Synaphaï”. In Kanach, Sharon (ed.) *Performing Xenakis*. New York: Pendragon Press.

Varga B. 1996. *Conversations with Iannis Xenakis*. London: Faber and Faber.

Gonzalez-Sanchez V., Dahl S., Lunde Hatfield J. and Godøy R. I., 2019. “Characterizing Movement Fluency in Musical Performance: Toward a Generic Measure for Technology Enhanced Learning”. In *Frontiers in Psychology,* 10:84.

Xenakis I., 2002. *Formalized Music*. New York: Pendragon Press.

Zenck M. 2006. “Luigi Nono-Marina Abramovic: Eingeschriebene bewegte und befreite Körper zwischen Aufführungspartitur, Live-Elektronik und freier Improvisation / Performance”. In Petra Maria Meyer (Hsgb.), *Performance in medialen Wandel.* München: Wilhelm Fink Verlag.

Zuboff S. 2019. *The age of surveillance capitalism. The fight for a human future at the new frontier of power*. New York: Public Affairs.