A GEOGRAPHICALLY DISTRIBUTED PERFORMANCE MEDIATED BY A CO-CREATIVE AGENT OVER THE WEB

Thales R. P. PESSANHA (0000-0002-3046-9710) [1,2], **Pedro MARTINS** (0000-0002-3630-7034) [2], and **Jônatas MANZOLLI** (0000-0003-4257-7118) [2,3]

[1] *Institute for Interdisciplinary Research*, **University of Coimbra**, *Computational Media Design*, Coimbra, Portugal
[2] *CISUC/LASI*, **University of Coimbra**, *Centre for Informatics and Systems of the University of Coimbra, Department of Informatics Engineering*, Coimbra, Portugal
[3] *Interdisciplinary Nucleus for Sound Studies*, **University of Campinas**, Campinas, Brazil

ABSTRACT

The convergence of digital sound technologies and immersive computing presents a fertile ground for experimentation, enabling artists to craft immersive sonic experiences that push the boundaries of traditional creative processes and promoting the active involvement of both artists and audiences, fostering creative collaborations and deep cognitive and emotional connections. However, in distributed performance settings, immersive environments bring challenges in terms of involvement, since the contexts in which both the performers and the audience are inserted are distinct. This, in turn, requires the development of new strategies and techniques to engage with digital sound, since this artifact is responsible for interconnecting the contexts of such environments. Based on this, we outline the architecture of a system aimed at creating immersive environments for distributed performance using an artificial co-creative agent. Also, we discuss both the computer implementation and performance tests of the DSP calculation to measure its feasibility in real-time performances. Finally, we consider that this architecture opens up new paths for involving human-computer interaction, offering innovative strategies for sound design in distributed performance contexts and the creation of accessible digital sound systems on the web.

1. INTRODUCTION

The emergence of new technologies, marked by the exponential growth in CPU/GPU processing power and the reduction of devices, has opened up new possibilities for integrating immersive technologies in the domains of music and performance art.

These emergent technologies also enabled the rise of systems that allow the connection of a set of technologies and objects in embedded systems, which grant communication and interaction between different elements over the internet [1], and helped to integrate local environmental information through motion sensors, cameras, and microphones, since mobile devices connected to the web create a much more connected environment and supply users with easy access to multimedia content [2].

Also, the specification and implementation of Application Programming Interfaces (APIs), Websocket protocols, synchronous and asynchronous processing, and other paradigms, infrastructures, and toolkit which are integrated into the environment and development of the web applications, have fostered the web platform into a fertile playground for artists and musicians [3], simplifying and democratizing the use of devices for sound diffusion, spatially distributed rendering, and the exploration of both social and sound interactions in their performances and creative projects.

Altogether, this interconnection promoted the active involvement of both artists and audiences, furthering creative collaborations and deep cognitive and emotional connections, drawing on a wealth of perceptual, motor and affective resources, enriching the experiences of all participants involved.

Nevertheless, Networked Music Performances still presents some challenges and drawbacks. These include the requirement for robust technological infrastructure facilitating seamless communication among devices, as well as ensuring connectivity conducive to meaningful performances [1, 4]. Moreover, the scarcity of sophisticated, simple-to-use tools tailored for non-expert users have required both researchers and artists to familiarize themselves with fundamental technologies [3]. Lastly, latency mitigation represents another significant challenge, given musicians' highly sensitivity to interaction delays. In a network performance setting, such delays are not only unavoidable but also have a physical lower limit [5].

In this paper, we propose a system's architecture which, based on management guided by an artificial co-creative agent, enables geographically distributed web-based musical performances. For such, we developed a prototype system called *JANIS* (Joint Audio Networking Interactive System) to devise engagement based on symbolic interactions [6]. This enables both musicians and co-creative agents to interact, learn and progress with each other through a cross-learning process which allows them to share attributes – such as intentionality, relevance, adequacy, aim, discovery and style [6] – and structures in both micro and macro scenarios.

This discussion contributes for the understanding and development of better strategies to engage and adapt musicians in contexts of distributed performance, enhancing their perceptual aspects and experience in a Networked Music Performance (NMP) context.

The remainder of this paper is organized as follows: in section 2, we provide an explanation about the concepts of co-creativity and the use of symbolic interaction to produce relationships between musicians and artificial agent. Section 3 presents a brief description about current technologies in Networked Music Performance research field. In Section 4 we address the architecture of *JANIS* system, the structure developed to process digital audio signal, which allows artificial co-creative agent's decisions during performance, and some initial results. Finally, Section 5 presents concluding remarks.

2. CO-CREATIVITY AND SYMBOLIC INTERACTION

The definition of concepts and psychological aspects related to creativity is something that has produced an extensive literature ranging from the definition of creativity in the creative process of the arts [7], the creativity of important creators in the history of humanity [8], and the manifestation of creativity as a kind of catalytic energy [9].

In practical terms, we can define creativity as the ability to come up with ideas or artifacts that are new, surprising and valuable [10]. Moreover, as creativity is grounded in everyday abilities – such as conceptual thinking, perception, memory, and reflective self-criticism – it can be defined as an aspect of human intelligence in general [10], entering into virtually every aspect of life and being part, to some extent, of everyone's ability to be creative.

Thus, whether from ideas – such as concepts, poems, musical compositions, scientific theories, culinary recipes, choreography, jokes and so on – or from artifacts – paintings, sculptures, steam engines, vacuum cleaners, pottery, origami, penny whistles and many other things we can name – creativity happens from the correspondence to three types of surprise [10]: by making unfamiliar combinations of familiar ideas, by the exploration of conceptual spaces in people's minds, or by the transformation of these conceptual spaces, which ends up transgressing the space of creation.

In recent years, some theorists, such as Williams [11], define creativity as a dynamic and iterative process that unfolds when individuals actively engage and disengage with a multifaceted array of stimuli – ranging from interactions with fellow individuals to immersion in diverse information, ideas, and the physical environment.

Also, creativity manifests itself through the prism of novelty, utility, and a grounding in reality with discernible goals and developmental progression over time [11], extending beyond a binary perspective, encompassing a continuum that spans from small-c creativity, focusing on problem-solving and incremental improvements, to big-C creativity, which exerts profound influences on cultures.

Crucially, the social dimension emerges as a pivotal element, highlighting creativity as an inherently social activity where interactions with both the social and physical environment shape the creative process [11]. Moreover, the intrinsic nature of creativity positions it as an autotelic pursuit, where the enjoyment derived from the process itself serves as a potent catalyst for producing outcomes of heightened creative merit [11].

When we focus on this perspective of creativity in the field of computer music, we see in recent years the development of interactive systems rooted in the evaluation of historical data, in attention to the ongoing sound environment, and dynamic forecasting strategies spanning multiple time scales, being designed for real-time responsive interactions and forward-looking projections [6].

The combination, therefore, of temporal and spatial elements generated by both the algorithms and the dynamic composition, provides a creative projection space capable of sustaining and constantly reviving the musicians' interest in creating with an artificial agents.

Furthermore, the reciprocal causality fostered by the interaction between agents – whether they are real or artificial – shows us that, from a cognitive point of view, they come together in joint action to show each other the paths of performance [12, 13], with their reciprocal production also being a source of learning for others, that is, their relationships are not causal, but causative [12].

This situation of adaptive interdependence is qualified within the musical domain as *symbolic interaction* [6] and involves every level of musical representation. Whether at the level of signal, structure, real or simulated cognitive modalities, these interactions are distributed between human and artificial agents.

It is from this context, thus, that co-creativity is established, arising emergent processes that, based on intertwined musical contributions from humans and machines, goes beyond the mere sum of their individual capabilities [6]. This collaborative process gives rise to unexpected outcomes, influencing the internal states of the performers through intricate dynamics, resulting in moments of tranquility, gradual progression, or abrupt shifts.

At its core, co-creativity – whether produced by human-human or human-computer interaction – is facilitated by generative learning mechanisms that are both cross-referenced and reflexive, with each agent's input at a given time combining its own productions with those of others, all while incorporating memories of past interactions [6]. This interplay fosters a rich and dynamic creative environment, where human and machine coalesce to produce novel and evolving musical expressions.

The synergy produced by the agents articulates actions and dynamic formations unattainable through individual efforts. The emerging forms evolve in a non-linear way, defying conventional molds and norms. This underlines the ability of co-creativity to not be a mere aggregation of individual contributions, but rather the birth of new and irreducible expressions.

As we discussed in this section, symbolic interactions can provide emergent processes during performance that require adaptation from both human and artificial performers, establishing a process that we can call co-creativity.

This gives rise to elements that, from our cognitive point of view and in accordance with the definitions of creativity described previously, can be characterized as novelty.

In our proposal, the Networked Music Performance mediated by co-creative agent provides the exchange of musical information in real-time and induces adaptative actions from both geographically distributed musicians and the artificial agent itself. In this way, we project such musical experience to bring new meanings to NMP experiences.

Next, we discuss the current technological tools in NMP research field.

3. CURRENT TECHNOLOGY IN NETWORKED MUSIC PERFORMANCE

The design and development of Networked Music Performance tools can be divided into three parts: a) the design and development of systems and/or interfaces; b) the design and development of protocols to establish communication between distributed agents over the network; c) the design of sound artifacts and the performance itself;

From the late 1990s through the early years of the 21st century, researchers focused on projects that aimed to achieve high-quality, real-time Networked Music Performance [14]. This performance, in addition to raising questions from a technological point of view, also led researchers to analyze the cultural implications that such installations imposed on musicians and producers [14], establishing strategies to increase the level of interactions between musicians who collaborate through network connections.

Thus, the challenges imposed at this stage of the NMP research field resulted in three approaches to designing NMP tools [14], namely:

- **Realistic Jam Approach:** When real-time live musical interactions are crucial, requiring geographically displaced musicians to come as close as possible to feeling like they are playing in the same space;
- **Latency Accepting Approach:** It considers the internet as a decentralized and space-independent medium. Therefore, efforts are focused on finding new forms of delayed musical interaction;
- **Remote Recording Approach:** It involves producing music using the internet as a medium for remote recording sessions, thus overcoming latencies for real human-to-human interactions to occur is not the goal.

As an example of technologies and tools included in each of these approaches we can mention: *Soundjack* [1] [15], *Ninjam* [2], and *VSTunnel*.

From this scenario, and with the increasingly accelerated revolutions in processing capacity and connection between devices, we have seen in recent decades the increasingly prominent development of tools and technologies aimed at distributed and web-based music systems for performance.

Moreover, the advancement of techniques and tools for the development of web platforms – made possible by the architectures, protocols, infrastructures and frameworks that are standardized in the market – makes us divert our focus from simply exploring and probing the possibilities and opportunities inherent in distributed and web-based music systems [3], allowing us to delve deeper into the characterization and implementation of experimental platforms that can function as a shared space for artists and researchers. In that sense, *Pteroptyx Malaccae* and *Clock(s)* [3] are good examples because they allow modularization as clients of a distributed system, reusing and adapting many existing software components such as clock synchronization, scheduling, stream processing or functionalities offered by existing frameworks.

However, although these technologies aim to simplify and democratize the experimental implementation of platforms dedicated to distributed and interactive web-based musical systems for artists and researchers in exploratory tasks, they were emergent technologies and had drawbacks related to the simplification of experimentation for non-specialist developer users such as composers, researchers or computer music designers.

Another issue currently imposed on the technological scenario in the NMP field is the creation of dynamic network infrastructures capable of connecting anything – whether they are physical or virtual objects – through different means of communication. This enables the collection of any type of information that can further be communicated and controlled via the Internet. The sensitivity and ability of devices to react and change their behavior based on the context and environment in which they are inserted, as well as the way they treat and exchange data with other devices that are connected to them, is a factor that it has been explored and placed in question.

As an example of this kind of technology is *Sunflower* [16, 17], an environment design that tries to contribute to solving the recurrent problems related to standardization, interoperability, privacy, and security between devices in a context of IoT musical technologies. Also, it handles different pieces of information, such as audio, video and control, in order to expand the capacity of non-musicians to participate and assist in an artistic creative process [17].

In order to support the interaction with all sort of musical things, Sunflower's management layer was inspired by other tools that work similarly, such as *Libmapper* [18] and *Medusa* [19], and also in tools present in musical desktops, like ALSA MIDI [3] and QJackCtl [4] [16]. Nevertheless, even though *Sunflower* is a playful and intuitive platform, which allows multiple users with different objectives and skill levels to use it without any major problems, it does not necessarily guarantee a good user experience in controlling the sound artifacts or other piece information in musical activities. This, however, is out of the scope of this paper.

Recently, given the social isolation caused by COVID-19 and motivated by the challenge of creating and interacting

[1] `https://www.soundjack.eu/`
[2] `https://www.cockos.com/ninjam/`
[3] `https://www.alsa-project.org/`
[4] `https://qjackctl.sourceforge.io/`

with other artists remotely, *grainBIRD* [20] was developed having as a concept a haze of mobile devices generating granular sounds via OSC messages within a network.

grainBIRD's network communication was based on IoT, Cloud computing and Fog computing paradigms, being derived from the concept of the laptop orchestra. The design of *grainBIRD*'s app applied the notion of cloud and fog computing architecture, in order to enable both local interactions and performances over a network of mobile devices [20].

As much as this technology interacts with the possibility of co-creation through the Internet and also with the paradigm of Ubiquitous Music, its multiple performance configurations can be further explored to obtain significant artistic results to promote engagement among performers connected to the network.

To summarize, it is clear that although the area of NMP has produced an extensive literature, its efforts are still related to technical aspects inherent to the development of the tools that are proposed. However, as musical practices offer a valuable mean for exploring the connections between perception, action, and meaning within the dynamic interplay among performers, audiences, and their environment [21], we see in this gap an important contribution to researches on perceptual aspects in the context of Networked Musical Performance.

It is in this context, thus, that the *JANIS* system is proposed, fostering the engagement and adaptation of performances in a NMP context based on an architecture which will be further explained in the next section.

4. ARCHITECTURE AND DEVELOPMENT STRUCTURES

In this section, we will focus on describing the architecture of *JANIS* (Joint Audio Networked Interactive System), a prototype system that aims to manage interactive performances through the web from an artificial co-creative agent.

Also, we will highlight the applied digital signal processing and the similarity calculation between pre-recorded signals and real-time signals to guarantee the artificial agent its independence for decision making.

Finally, we combined the results obtained from the digital signal processing, along with the concept of co-creativity explained in Section 2, to highlight some preliminary results.

4.1 User interaction and system's architecture

The starting point for using the *JANIS* system is the user interface created to perform interactions based on predefined actions. This interface, as shown in Figure 1, has buttons that, when clicked by the users, trigger actions such as recording, analyzing, comparing audio signals, among others.

These actions, in turn, are perceived by the users through a chat, in which they receive messages from the agent indicating which actions it has just performed, or through the signal graph, which indicates the capture of the audio and

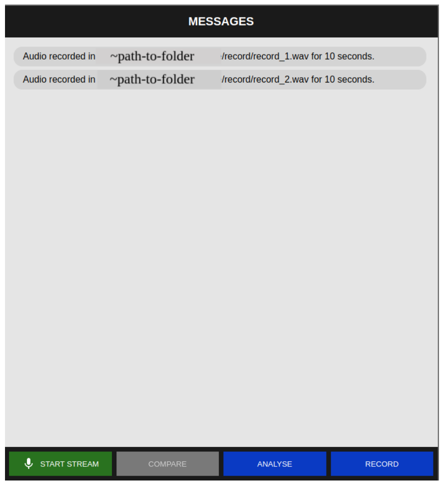

Figure 1: System's user interface which makes interaction between user and co-creative artificial agent possible.

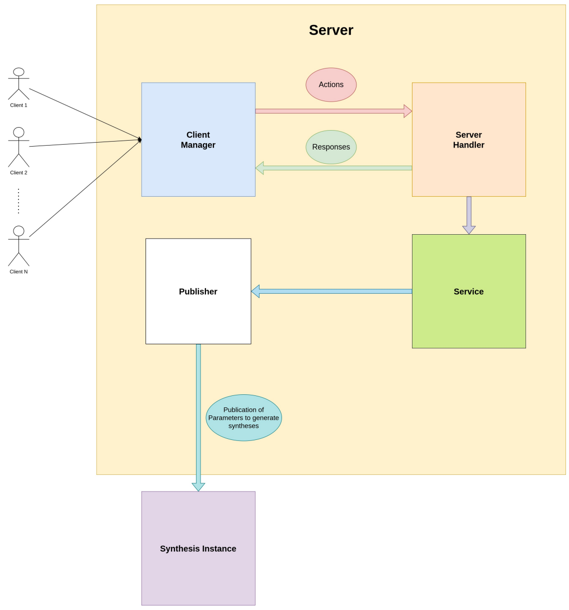

Figure 2: Architecture schema of co-creative system.

its reception by the agent. A more detailed description of these actions and the routines performed both by the users and the server side can be seen in the Table 1.

In order to establish communication between the client and the server, in this case between the user and the co-creative agent, the Websocket protocol was employed. This protocol selection was chosen because, unlike the HTTP protocol – which terminates its connections upon receiving a response – Websocket maintains two-way communication, ensuring continuous connectivity until termination is initiated by one of the parties [22].

Moreover, the decision to use the Websocket protocol stems from its ability to efficiently handle a considerable volume of messages within a defined time frame [22], consequently increasing the system's capacity to process the continuous flow of information that a performance can produce.

Table 1: Actions and routines make in front-end and back-end

Actions	Front-end routine	Back-end routine
record	Send audio data to back-end process it	Process the audio data and stores it in a wave file, returning a warning message to the user that audio has been recorded
analyse	Send a message to the back-end requesting analysis of recorded audio	Extract parameters from the signal and save it in a file, returning a message to the front that the audios have been analyzed
open	Send this message to start sending real-time audio signal to the back	Start monitoring audio data in real time
compare	Send this message to the back to start comparing the audio files with the captured signal in real time	Process captured audio data in real time and compares it with data pre-recorded by the record command
stop	Sends message to stop comparison	Stop comparing data captured in real time and pre-recorded audio
close	Send message to close real-time audio capture	Close real-time audio capture

By choosing this type of protocol, therefore, our focus was on ensuring means for external events to the system, in this case the musical performance, to determine the program's execution flow, as its different formats, in this case its different symbolic interactions, made the system listens to events that happen over time and react to them as they occurred.

This allowed us to structure an architecture with greater flexibility and responsiveness, in addition to seeking better performance and a consistent structure capable of supporting the volume of data to be processed.

In this way, the server implementation was defined based on four layers, namely:

- **Client Manager:** Responsible for managing user connections, as well as sending messages to a specific user, or to other connected users;
- **Server Handler:** Responsible for handling server routines and managing messages received from clients and send to be processed by service layer;
- **Service:** Responsible for carrying out actions requested by users;
- **Publisher:** Responsible for publishing messages or data to be executed or processed by third-party systems.

To complement the scope of the system, as shown in Figure 2, the server sends information to third-party systems to synthesize sounds using different processes. In our case, we choose Supercollider as an instance to produce FM, Granular, and Wavetable synthesis.

It is worth mentioning that this communication is carried out using the Open Sound Control (OSC) protocol, which makes it possible to use the system with other types of sound synthesis programs, such as Max/MSP or PureData.

This sending of data is carried out by calculating the similarity between the sound captured in real time and other pre-recorded sounds, which will be better highlighted in the next subsection.

4.2 Digital signal processing and computing similarity

As the structured architecture proposed and explained in subsection 4.1 aims at a reactive approach to the various symbolic interactions that occur between human and artificial agents throughout the performance, both DSP analysis and similarity calculations should follow this same principle, as they should not interrupt or block the data flow to be processed.

Thus, the starting point for performing calculations on both pre-recorded and real-time audio signals s is given by the root mean square of them, as shown in (1):

$$f(t) = \sqrt{\frac{1}{K} \cdot \sum_{k=t\cdot K}^{(t+1)\cdot K-1} s(k)^2}, \qquad (1)$$

where K is the number of samples we have in a frame and t is the frame where root mean square (RMS) will be computed.

After that, and as shown in (2), the gradient is computed using second order accurate central differences in the interior points and either first order accurate one-sides (forward or backwards) differences at the boundaries:

$$\nabla f(t_i) = \frac{f(t_{i+1}) - f(t_{i-1})}{2h} + \mathcal{O}\left(h^2\right), \qquad (2)$$

where t_i is the frame, h is a non-homogeneous step size and $\mathcal{O}\left(h^2\right)$ represents the order of error proportional to h^2.

To avoid huge distortions, and to extract meaningful results, this gradient is normalized using (3):

$$X_{\text{norm}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}, \qquad (3)$$

where X is the array resulted from gradient's calculation.

The data resulted in this gradient computation is analyzed in order to find the peaks of the signal, which takes into account a one-dimensional data series defining a local maximum as a point x such that:

$$x_{i-1} < x_i \text{ and } x_i > x_{i+1}. \quad (4)$$

To avoid inappropriate local maxima, we can optionally select conditions such as height, limit, distance, prominence and width that are adjusted according to processes and data obtained in each performance.

These values are then stored in files that can be used throughout the performance to calculate the similarity with the signals captured in real time.

To calculate the similarity between these signals, the cosine of the angle produced by their arrays is computed, as shown in (5):

$$\text{similarity}(\vec{u}, \vec{v}) = \frac{\vec{u} \cdot \vec{v}}{||\vec{u}||||\vec{v}||}, \quad (5)$$

where $\vec{u}$ and $\vec{v}$ are, respectively, the vectors of input and pre-recorded signals peaks computation, $\cdot$ is the dot product and $||||$ is the Euclidean norm.

The closer the value of this angle's cosine is to zero, given a minimum threshold also defined from the data obtained in each performance, the more similarity is detected by the artificial agent.

This similarity, therefore, allows this artificial agent to make decisions throughout the performance and contribute co-creatively with the performers in the production of emergent processes.

Also, given the indeterminacy of each result obtained in the various performances, musicians have to adapt to the adverse conditions that each one may generate, even more so when thinking about a geographically distributed performance context.

In other words, this allows us to raise questions about what conditions musicians are able to adapt and to adjust their anticipation according to the confirmations or surprises that occur, the ability to enjoy artistic freedom, and the amount of interest inspired by the productions of the machines based on reciprocal intentionality.

4.3 Digital Signal Processing Performance and Discussion

We carried out some tests of JANIS system's digital signal processing calculation – to measure its feasibility in real-time performances. In this setting, a classic guitar standalone solo performance was evaluated.

Figure 3 shows the computation of the peaks obtained from the RMS gradient of the guitar signal, applied to different attack and sound production modes.

To detect the peaks, tests were applied with a threshold varying from 0.3 to 0.5 according to the variance that each attack mode could produce.

Furthermore, we decided to analyze the spectrogram of each signal, as shown in Figure 4, in order to identify which scenarios detection could be inefficient given the complexity of the sound spectrum formed.

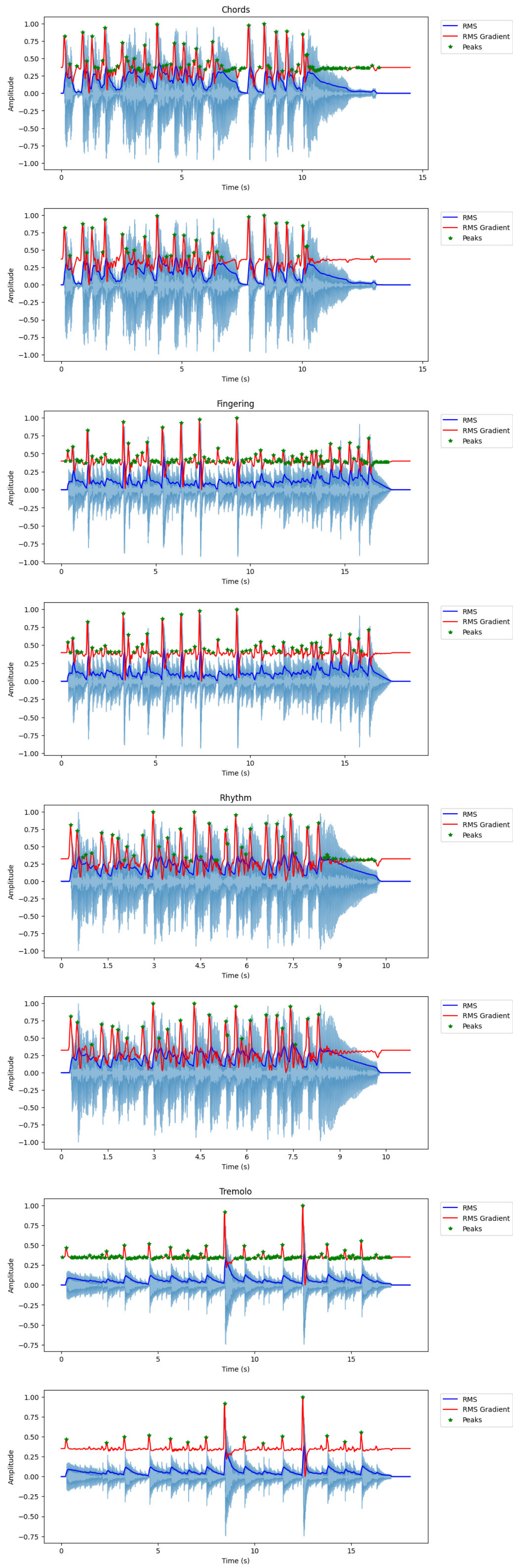

Figure 3: Peak finding results applied in different attack and sound production modes.

From the graphs presented, it is possible to evaluate that sounds that are more defined in terms of their spectrum, such as those presented by *fingering* and *tremolo*, have less inaccuracy in detecting their peaks, whether they are at points of greater magnitude or high variance.

For less defined sounds, such as *chords* and *rhythm*, depending on the defined threshold values, we see some inaccuracy in finding peak points in the RMS gradient.

This shows that, depending on the parameterization chosen for each performance, the actions of the artificial agent can vary significantly according to what it is “hearing”. This means that the musician also needs to be attentive and present, being able to produce meaningful interactions based on the cognitive stimuli that the artificial agent may provides.

Also, by projecting the similarity of the input signal based on the greater or lesser clarity in the segmentation of pre-recorded signals, we can control the synthesis parameters to generate distinct sounds from a given configuration. For example, if the input signal has high energy and the segmentation is clearer, we can work on the harmonicity, rhythm or even the amplitude of the output signal to be clearer, periodic or harmonic. If the signal has less energy and the segmentation is less clear, these same characteristics can result in sounds that are less harmonic and have a more dynamic rhythm. However, these characteristics can be explored based on the different artistic intentions that artists want to produce.

From the perspective of distributed sound performance, the synchrony and complexity of the resulting synthesis should not be the starting point, but rather the different configurations that allow agents to listen to and interact with each other.

The network latency is neither significant nor it is the cause of engagement, putting into play the synergy that must be managed by both the human and the artificial agent to produce novel and engaging musical expressions.

From the discussion here, it is possible to understand that *JANIS* has a great potential to allow multiple performance configurations that can produce co-creative process between both human and artificial agents over a NMP context and can also be part of several distributed performance configurations.

5. CONCLUSIONS

In this paper, we focused on discuss the possibilities to apply the concept of co-creativity and symbolic interactions in the context of Networked Music Performance in order to promote the involvement and engagement of geographically distributed musicians during the performance.

The use of artificial co-creative agents can be effective in promoting interaction between human agents, since their active interaction provides the creation of projection spaces that can be interesting for both to share their actions and efforts. This happens because both human and artificial agents are in a situation of adaptive interdependence and incorporating memories that promote a rich and dynamic creative environment for their interactions. This phenomenon has been widely explored in recent years in

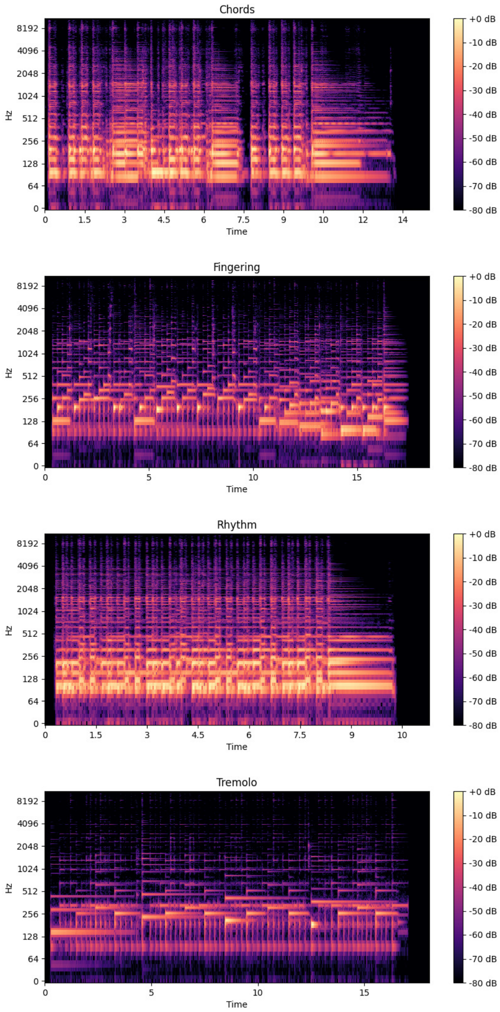

Figure 4: Spectrogram analysis of this guitar attack and sound production modes.

the context of human-computer co-creativity, and could be used to promote meaningful experiences in NMP contexts.

Furthermore, by proposing the prototype of the *JANIS* system, which, based on the tests carried out, has great potential to generate multiple performance configurations, our focus converges on studying cognitive and artistic strategies to promote engagement and adaptation of musicians in geographically distributed performances.

Finally, we intend to apply more tests to this system to understand its possibilities and limitations in addition to developing its communication with other digital musical instruments and its application within artistic research involving musicians and dancers, which motivated the development reported here.

Acknowledgments

This work is funded by the FCT - Foundation for Science and Technology, I.P./MCTES through national funds (PIDDAC), within the scope of CISUC R&D Unit - UIDB/00326/2020 or project code UIDP/00326/2020*

6. REFERENCES

[1] A. S. D. Almeida, R. Vieira, R. S. Oliveira, and F. L. Schiavoni, “Desafios da Internet das Coisas Musicais,” in *Anais do XVIII Simpósio Brasileiro de Computação Musical (SBCM 2021)*. Brasil: Sociedade Brasileira de Computação - SBC, Oct. 2021, pp. 249–252.

[2] K. Nishida, A. Yuguchi, K. Jo, P. Modler, and M. Noisternig, “Border: A Live Performance Based on Web AR and a Gesture-Controlled Virtual Instrument,” Jun. 2019, publisher: Zenodo. [Online]. Available: https://zenodo.org/record/3672860

[3] B. Matuszewski and F. Bevilacqua, “Toward a Web of Audio Things,” in *Sound and Music Computing Conference*, Limassol, Cyprus, Jul. 2018. [Online]. Available: https://hal.science/hal-01874968

[4] L. Turchet, C. Fischione, G. Essl, D. Keller, and M. Barthet, “Internet of Musical Things: Vision and Challenges,” *IEEE Access*, vol. 6, pp. 61 994–62 017, 2018, conference Name: IEEE Access.

[5] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti, “An Overview on Networked Music Performance Technologies,” *IEEE Access*, vol. 4, pp. 8823–8843, 2016, conference Name: IEEE Access.

[6] G. Assayag, “Human-Machine Co-Creativity,” in *in Bernard Lubat, Gérard Assayag, Marc Chemillier. Artisticiel / Cyber-Improvisations. Phonofaune, 2021, Dialogiques d’Uzeste. ⟨hal-03542917⟩*, May 2021. [Online]. Available: https://hal.science/hal-03542986

[7] O. Harold, “The Concept Of Creativity in Art,” *The British Journal of Aesthetics*, vol. 19, no. 3, pp. 224–231, 07 1979. [Online]. Available: https://doi.org/10.1093/bjaesthetics/19.3.224

[8] M. A. Boden, Ed., *Dimensions of Creativity*. The MIT Press, Jun. 1994.

[9] R. K. Sawyer and D. Henriksen, *Explaining Creativity: The Science of Human Innovation*. Oxford University Press, Jan. 2024.

[10] M. A. Boden, *The Creative Mind*, 2nd ed. 11 New Fetter Lane, London EC4P 4EE: Routledge, Feb. 2004.

[11] A. Williams, “Behind the Idea: Meta-Models for Creativity,” in *Intersecting Art and Technology in Practice: Techne/Technique/Technology*. Taylor & Francis, Dec. 2016, pp. 182–197.

[12] G. Assayag, “Sex Machine: IA CoCreative et position du sujet,” Jan. 2024. [Online]. Available: https://www.youtube.com/watch?v=LWT99riOWeQ

[13] M. T. Llano, M. d’Inverno, M. Yee-King, J. McCormack, A. Ilsar, A. Pease, and S. Colton, “Explainable Computational Creativity,” 2022.

[14] A. Carôt, P. Rebelo, and A. Renaud, “Networked music performance: State of the art,” in *Audio Engineering Society Conference: 30th International Conference: Intelligent Audio Environments*, Mar 2007.

[15] A. Carôt, C. Hoene, H. Busse, and C. Kuhr, “Results of the fast-music project—five contributions to the domain of distributed music,” *IEEE Access*, vol. 8, pp. 47 925–47 951, 2020.

[16] R. Vieira and F. L. Schiavoni, “Managing an IoMusT environment and devices,” in *11th Workshop on Ubiquitous Music (UbiMus 2021)*, ser. Proceedings of the 11th Workshop on Ubiquitous Music (UbiMus 2021). Matosinhos, Portugal: g-ubimus, Sep. 2021, pp. 32–44. [Online]. Available: https://hal.science/hal-03396836

[17] R. Vieira, D. C. Muchaluat-Saade, and F. L. Schiavoni, “Sunflower: An Interactive Artistic Environment based on IoMusT Concepts,” in *Proceedings of the 2022 ACM International Conference on Interactive Media Experiences*, ser. IMX ’22. New York, NY, USA: Association for Computing Machinery, Jun. 2022, pp. 245–248. [Online]. Available: https://doi.org/10.1145/3505284.3532818

[18] J. Malloch, J. Garcia, M. M. Wanderley, W. E. Mackay, M. Beaudouin-Lafon, and S. Huot, “A Design Workbench for Interactive Music Systems,” in *New Directions in Music and Human-Computer Interaction*, ser. Springer Series on Cultural Computing, S. Holland, T. Mudd, K. Wilkie-McKenna, A. McPherson, and M. M. Wanderley, Eds. Cham: Springer International Publishing, 2019, pp. 23–40. [Online]. Available: https://doi.org/10.1007/978-3-319-92069-6_2

[19] F. L. Schiavoni and M. G. d. Queiroz, “Medusa: um ambiente musical distribuído,” Master’s thesis, Universidade de São Paulo, 2013. [Online]. Available: http://www.teses.usp.br/teses/disponiveis/45/45134/tde-08092014-172703

[20] J. Manzolli and E. H. Constantino, “grainBirds: Evolutionary granular synthesisers distributed in Fog Computing,” in *Anais do XVIII Simpósio Brasileiro de Computação Musical (SBCM 2021)*. Brasil: Sociedade Brasileira de Computação - SBC, Oct. 2021, pp. 95–102. [Online]. Available: https://sol.sbc.org.br/index.php/sbcm/article/view/19432

[21] A. Linson and E. F. Clarke, “Distributed cognition, ecological theory, and group improvisation,” in *Distributed Creativity: Collaboration and Improvisation in Contemporary Music*. Oxford University Press, Dec. 2017, vol. 2, pp. 52–69.

[22] G. Gonçalves, P. Bastos, and D. Oliveira, “O uso de websockets no desenvolvimento de sistemas baseados em uma arquitetura front-end com api,” in *Anais da I Escola Regional de Sistemas de Informação do Rio de Janeiro*. Porto Alegre, RS, Brasil: SBC, 2014, pp. 57–63.