IEMI: AN IMMERSIVE, EVOLVING MOTION-CONTROLLED INSTALLATION

Ahmet Emin MEMIS (ahmetem@student.imv.uio.no) [1] and
Tejaswinee KELKAR (tejaswinee.kelkar@imv.uio.no) (0000-0002-2324-6850) [2]

[1,2] *Department of Musicology*, **University of Oslo**, Norway

ABSTRACT

Interactive Evolutionary Algorithms (IEA) are an extension of Evolutionary Algorithms, where the optimization process is centered around a persons subjective evaluation. IEAs add another layer of user interaction to transform the evolutionary process from being solely dependent on objective selection to incorporating real-time subjective evaluations. This project presents an interactive installation that allows users to engage with a sound bank by moving in space, where the mapping is updated using an evolutionary algorithm. Interactive feedback from the user is captured through button press. In this paper we present a two-stage iterative evaluation of the system. First, a pilot user study was conducted which questions the system's usability and motion capture quality, as well as making qualitative measurements. The results suggest that the system provides a new and engaging way for users to interact with their space. The system created in this paper is source open and available in a GitHub repository.

1. INTRODUCTION

Creating immersive installations is a challenge with very specific constraints. The need for making sound installations in museums and exhibitions interactive on the one hand has to be balanced with the subtle nature of the soundscapes and their sonic qualities. Overt interactive patterns exhaust the user and the soundscape, whereas repetitive sound patterns cause listener fatigue. In this paper, we explore interactive soundscapes with the help of Interactive Evolutionary Algorithms (IEA) and motion capture.

Evolutionary Algorithms are a subfield of Artificial Intelligence that uses mechanisms inspired by biological evolution. They are designed to solve problems by iteratively improving a population of potential solutions according to defined criteria [1]. Interactive Evolutionary Algorithms (IEA) are an extension of Evolutionary Algorithms, where the optimization process is centered around a person's subjective evaluation [2]. IEAs add another layer of user interaction to transform the evolutionary process from being solely dependent on objective selection to incorporating real-time subjective evaluations. IEAs make evolution processes user-centric, answering to various needs of design processes. However, this approach necessitates the user to continuously be responsive to the interface, especially in live performances. This has previously been solved by tightly sampled questionnaires, preference feedback using input tools and more. Tight sampling of user feedback creates fatigue in responding, and fails to provide seamless interaction. The metacognition of remembering to give feedback is a barrier to experiencing flow and immersion in a musical context.

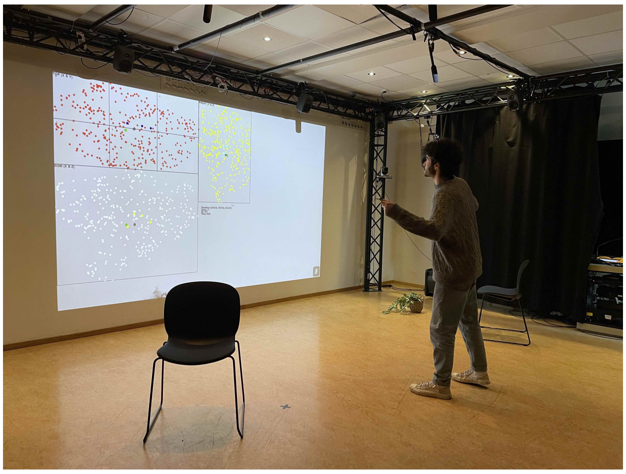

Figure 1. A user interacting with the IEMI

We present IEMI: a movement-based evolutionary algorithmic composer — a system that works with optical infra-red motion capture data as the movement feedback, where the sound is redistributed and mapped to the capture space in which the user can interact with the sound by moving their body. We present two user studies to analyze the interactive potentials of this family of algorithms in creative sound design. We make a case for movement-related feedback to shape the behavior of the system in complex and interactive ways. We see a potential for this work to contribute to the field of interaction design in sound installation.

The source code for this project is available in a GitHub repository [1] .

Recent advancements in motion capture and audio processing technologies have led to the emergence of new forms of musical instruments that enable users to engage with their bodies in novel ways [3–5]. This project aims to implement an installation that allows users to navigate sound-banks in real-time using their body movements. The controls of the installation are based on position in the

[1] https://github.com/aememis/IEMI

space, quantity of motion, and parameters that are governed by an evolutionary algorithm. These parameters are mapped to various effects including delay, tremolo, reverb and frequency filtering. We use an Infra Red (IR) motion capture technology, hereafter IR mocap, to track the body movements of the user and processing of the voice using the extracted motion values. A reflective marker was placed on the body, and its positions was live streamed to be converted to motion amounts in the Python environment, then to Pure Data software where the audio processing operations were performed based on pre-calculated values. The vocal input is received through a wireless headset microphone and directed to Pure Data [2] .

This project contributes to the field of human-computer interaction by exploring the potential of using motion capture and IEA technologies to create more embodied musical experiences. By enhancing the user's sense of embodiment, the vocal instrument offers new possibilities for vocal expression and creativity. The following sections of this paper provide a more detailed explanation of the methodology, results, and a discussion of the implications and potential future directions of this work.

2. RELATED WORK

With new developments in hardware, software, and human-computer interaction there is a growing interest in designing interactive systems that engage users in novel ways and enhance their cognitive and emotional experiences. Research on theoretical and experimental methods of extending artistic outcome of the body instrument has been conducted in the field of embodied music cognition. Tanaka and Donnarumma [6] explored the notion of the body as a musical instrument and argued that the body has the potential to be seen not only as a tool for producing sound but as an instrument in its own right, with its own unique affordances and sonic qualities. Jensenius and Wanderley's [7] explorations on extending the body instrument into the realm of movement-based performance art aligns closely with this work's principles. Through the integration of sensor technology and computer programming, Jensenius discusses that the body can be transformed into an instrument that generates new sounds and visual displays based on the performer's movements. Additionally, Mainsbridge [8] offers valuable insights into the use of motion-controlled performance systems to create live performances that reflect the unique style of each musician.

In this work, we take inspiration for designing interaction from several works in which position, posture and movement are mapped. One particularly influential model is OtoKin [9], which introduces a collaborative performance space abstracted in 3 dimensions and divided into specific zones for sound controls. Additionally, Fornari's application of evolutionary algorithms for sonic landscape design aligns closely with our strategy to build the auditory environment in response to user interaction [10]. They used these algorithms to generate birdsongs to build an artificial sonic landscape. Furthermore, they discuss the feasibility of integrating interactivity for fitness in their proposed system. To explore the integration of adaptivity within interactive music systems, Erdem et al.'s CAVI [11] provided valuable insights. Their work discussed the development of an interactive system that is as much an adaptive composition system as an interactive musical instrument. In this context, they presented the CAVI, a co-adaptive instrument that generates audiovisual responses dynamically evolving in response to user interaction.

2.1 Research Objectives

Drawing inspiration from this related work, we aim to create a motion-based interactive system for exploring evolutionary algorithms. Our main research objective is to learn

> does using motion capture for user feedback create an immersive sound experience to interact with an evolutionary algorithm.

To this end, we created a prototype for the algorithm, and studied interaction behavior with this algorithm. We are interested in how this user interaction is able to create a concept of control of sound parameters. Since the feedback is both motion based, and unary, we are interested in how the gradually changing slow interaction is experienced by people interacting with such a system.

3. DESIGN AND IMPLEMENTATION

The design focus for IEMI is to allow the user to control music creation through their body movements and spatial interaction, while experiencing the auditory results to provide real-time feedback that contributes to the evolution. Our system comprised an interactive space equipped with motion capture and spatial audio technology, alongside a computational framework that employs evolutionary algorithms.

Both prototypes utilized parameters steered by the evolutionary algorithms, projected into three-dimensional space for user interaction. This setup allowed users to navigate and explore with the parameter space through physical movement. The auditory outcome dynamically reflects the user's position within the environment, based on a proximity-based selection of parameters where the nearest data point in the space takes precedence.

Our interaction design focus was creating an interactive space in which the user can freely move and sound-walk to explore and adapt the space based on their musical preference. To achieve this, we integrated motion capture to track the user's motion and a feedback mechanism to gather real-time feedback from the user. The user navigated the interactive space using the marker, with the system processing a continuous stream of the marker's position as input. For giving subjective feedback, we used a computer mouse where a click represented a *like*.

In both prototypes, our sound design aimed to create an engaging auditory experience that showcased the adaptive behaviour of the sound installation. To achieve this, the

[2] `https://puredata.info/`

initial prototype was designed to manipulate effect parameters on vocals and continuous audio samples, while the second prototype introduced the dynamic selection of audio samples from a comprehensive sound bank. Figure 2 illustrates the main components of the system.

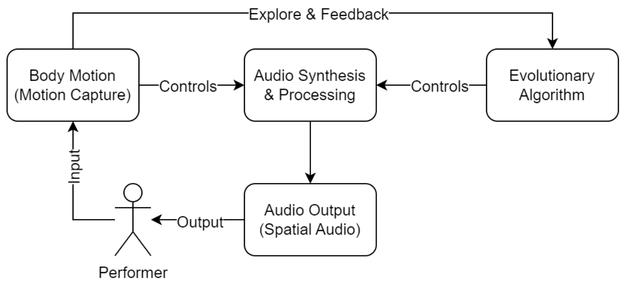

Figure 2. Main components of the IEMI system.

3.1 Equipment

To capture the user's body motion, we used Optitrack's infrared motion capture system with 8 cameras[3]. The voice input was captured using a wireless headset microphone. The output was played through a rectangular array of twelve speakers which were suspended from the ceiling. Finally, the visual content was projected to a wall of the room, which can be seen in Figure 4.

3.2 Motion Capture

To track the user's location, we used a reflective marker controlled by the user. The real-time positional data of this marker was continuously streamed from a stationary computer to Python environment using NatNet[4], an SDK capable of cross-platform motion capture data streaming. Within this environment, we obtained the marker's exact position in three-dimensional coordinates.

The motion intensity was quantified by processing the position data stream, specifically by calculating the Root Mean Square (RMS) over a sliding window of the last 100 samples. This provided a dynamic measure of user's activity within the space. For the design of P1, which required segmentation of the room floor into distinct zones, the interaction space was divided into six virtual zones, and the specific zone occupied by the user was continuously identified in real-time.

3.3 Data Representation and Preprocessing

We designed the individuals in the population to represent certain characteristics of the auditory feedback. These included effect control parameters for P1, while being features representing the audio samples for P2. These parameters are stored as an array of floating point numbers, scaled between 0 and 1.

To facilitate room-scale interaction with the population through motion, dimensionality of the population was reduced and projected into a three-dimensional space. We employed Uniform Manifold Approximation and Projection (UMAP) algorithm for this purpose, which is known for its ability to preserve more of the global structure compared to other well-known dimensionality reduction techniques [12]. UMAP also served as an effective clustering tool by grouping similar data points and assigning distinct characteristic sounds to different areas in the room. This eliminated the need for users to search randomly. During this process, direct link to the original population was preserved. This approach ensured that interactions with the UMAP representation could directly influence the original population. The UMAP representation was scaled to match the dimensions of the physical room that the installation took place.

As the user navigated the space with the marker, the nearest individual was determined through ongoing calculation of Euclidean proximity in three-dimensional space. The individual nearest to the user's current position was deemed the active selection. The active individual took precedence and controlled the ongoing sonic outcome. Moreover, the feedback received from the user was directly associated with the corresponding individual in the population. These projection approaches enabled effective interaction with high-dimensional populations through a scaled-down projection. Figure 3 shows the continuous transition between the UMAP projection and the population.

Incorporating these calculations, a real-time visualization of the motion capture interaction was displayed to the users, aiming for increased engagement and understanding of the system's response to their movements. Figure 4 shows a screenshot of this visualization during the interaction. It illustrates the spatial layout of the UMAP projection within the interaction zone across three planes. Additionally, it shows the six virtual zones, the user's position, 5 nearest neighbours, closest neighbour, and additional contextual information like current position in coordinates and elapsed time.

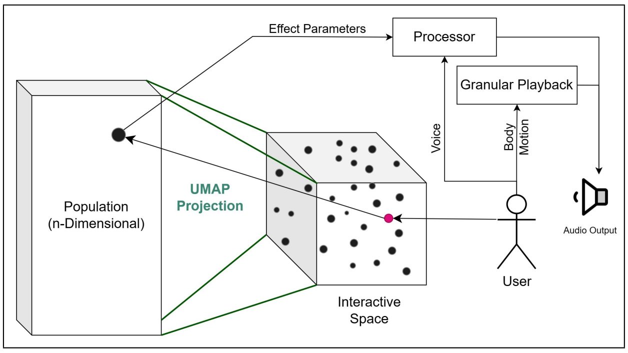

Figure 3. The system selects the nearest point within the 3D space based on user's position, identifies the corresponding individual in the population, and activates it to control the audio effects. Simultaneously, user's body motion controls the granular playback.

[3] https://optitrack.com/
[4] https://optitrack.com/software/natnet-sdk/

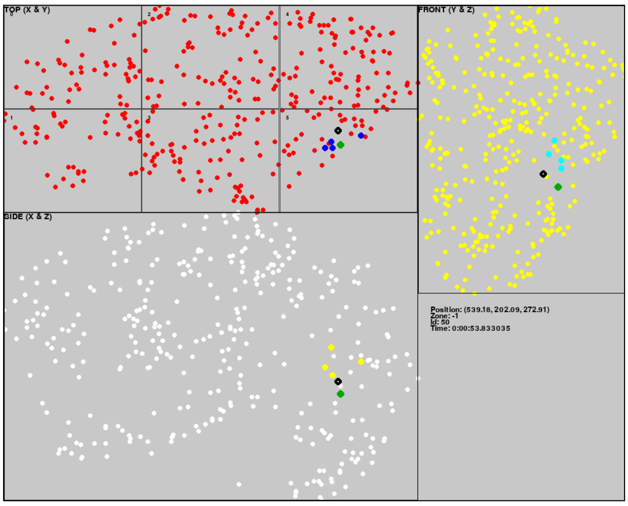

Figure 4. A screenshot of the visualization shown to the user, highlighting the UMAP projection's distribution within the interaction space. A black dot represents the marker's location and a green dot marks the nearest individual, and six virtual zones are shown within the TOP view.

3.4 First Prototype - P1

The first prototype of IEMI was designed to provide an engaging auditory experience using user's vocals, and audio samples. In the core of the prototype, the user sings while controlling sound synthesis and various audio effects are applied to their voice and the audio samples. P1 allowed the user to create the auditory outcome through their voice and body movements—either directly influencing the sound or indirectly, via the evolutionary algorithm, to guide the system's adaptive responses.

The main sonic elements of prototype P1 were the vocal input and a set of audio samples. The vocal input is captured using a headset microphone. Meanwhile, the audio samples, including both ambient sounds and vocal elements, were played back in grains.

The evolutionary algorithm parameters controlled the audio effects with a complex mapping. Initially, the installation began with a randomly generated population. This population then underwent evolution, driven by real-time feedback from the user. Firstly, similar to [9], we introduced a "sound-walk" concept where the floor of the room was divided into six zones, allowing the user to use these zones as a control method. As the user stepped into a zone, playback of pre-assigned audio samples was triggered. These samples comprised of open source recordings of a steel hang instrument, speech, whispering, and layered acapella singing. The design intent behind this selection was to craft an immersive soundscape containing several spectral, and envelope characteristics. As the primary focus of this prototype was on vocals, we sourced mainly vocal material to maintain a consistent theme. The zones also controlled other parameters in the sound generation, such as carrier frequency of drone and grain lengths.

The spatial audio distribution was linked to the user's orientation with respect to the x-axis, with variable rules

User Interaction	Control parameter
Quantity of Motion	Volume
Virtual zone	Triggering audio samples
Position	Spatial distribution
Height	Grain speed
Population Parameters	Delay, Tremolo, Reverb and Frequency Filtering

Table 1. Mapping of user interaction to control parameters.

applied to different samples enrich the sonic experience. This approach applied direct or modified correlation to the user's orientation to sound distribution, introducing a nuanced responses to the sound distribution. Moreover, moving away from the center resulted in a more focused sound output from a narrower range of speakers, whereas moving towards the center produced a broader sound.

This prototype was designed to respond dynamically to a user's physical activity by establishing a direct relationship between the quantity of motion and the auditory feedback. Specifically, standing still would result in increased volume, while more movement would lead to reduced volume. This design choice aimed to encourage the user to pause and immerse themselves with the auditory stream. Moreover, it refined the feedback mechanism, ensuring that interactions were accurately attributed to the most recently engaged individual within the evolutionary parameter space. Finally, to further enrich the overall experience, we integrated continuous drone sounds and other ambient elements. Table 1 shows the mapping between motion and audio processing.

Necessary calculations were made either in Python, or Pure Data, a visual programming language tailored for audio processing. The parameters calculated within Python environment are transmitted to Pure Data for the sound synthesis and further processing.

3.5 Second Prototype - P2

The second iteration focused on exploring a sound collection to create an immersive and adaptive dynamic auditory landscape. To achieve this, we used a large dataset of audio samples, and an evolutionary algorithm that processes audio features extracted from these samples. The virtual space was designed to evolve by gradually selecting audio samples that aligned more closely with user preferences.

We used the *Freesound Dataset 50 k* [13], a large dataset of short audio clips. From this dataset, we randomly selected 4000 audio clips, each with durations ranging from 0.25 to 4 seconds.

To prepare the data for integration into the installation, we focused on capturing their characteristics through calculating a set of features. This was achieved by analyzing the clips within sliding windows, and computing the mean values of several key audio features. These features were chosen strategically to represent the clips' dynamic range, timbral quality, and texture. These features included Root Mean Square for intensity, Spectral Bandwidth and Centroid for timbral characteristics, Spectral Flux for sound texture changes, and Mel-Frequency Cepstral Coefficients for unique sound signatures, along with Spectral Flatness to measure tonal versus noisy sounds. The length of each clip was also recorded to understand the temporal aspect.

The evolutionary process was driven by the extracted audio features to ensure that the population of audio samples could evolve in response to user feedback. Each individual within the space represented a unique set of audio features. As the user interacted with the environment, the system continuously identified and selected the audio sample whose audio features most closely matched those of the individual currently being engaged by the user. To identify the most similar audio sample, we calculated audio features' Euclidean distance. This approach allowed the system to adjust to user preferences by introducing new audio samples from the dataset that best matched their interests. Figure 5 illustrates this process.

3.6 Algorithm Design

Evolutionary behavior in IEAs is mainly guided by the principles of genetic algorithms, which mimic biological evolution through the creation and optimization of solutions via selection, mutation, and crossover. The fitness of the audio samples is determined by user feedback, which makes the process interactive and allows for adaptation in response to user preferences.

For both prototypes, the algorithm initiated with a population size of 350. In the case of P1, this consisted of randomly generated parameters, whereas for P2, it involved a random selection of audio features from the dataset.

The performer is granted the freedom to navigate the 3-dimensional UMAP representation of the population and engage with and react to the auditory results. Navigating such a complex space effectively poses a significant challenge and makes it nearly impossible to construct a mental model of their interactions [14]. Therefore, similar to the strategy suggested by [14], the system was designed to be explored in shrinking steps. In this iterative approach, evaluation process progresses with only a portion of the population being assessed. Subsequently, the system evolves by generating individuals that resemble the preferred solutions. This search method, although it may deviate from the Darwinian evolutionary approaches, facilitates faster convergence, while maintaining the control over the direction of the process.

For each individual in the population, the performer's feedback is unary, recorded as "liked", or left "not liked or encountered" (e.g., 1 or 0). Once the portion of the rated individuals exceeds a certain threshold, this triggers the selection process. While deciding on this threshold, one key factor was ensuring that the number of individuals resulting from crossover and mutation processes closely matches the size of the initial population. The selection process selects the rated individuals for producing offspring, while the unrated ones are eliminated in the next generation.

Following the selection, the individuals progress with the variation operators, specifically crossover and mutation. In the crossover phase, for each pair of individuals, a randomly chosen segment from one partner is exchanged with the corresponding segment from the other partner. The offspring produced by this exchange is passed to mutation phase, where the randomly chosen parameters within individuals are altered by a random factor. The probability of mutation was set to 0.5. To enhance diversity, the mutated individuals are added as new entries for subsequent generations. Figure 6 shows the flow of the interactive genetic algorithm.

4. EVALUATION

We developed IEMI in two iterations. The first prototype, P1 was evaluated with the help of a purely qualitative pilot study. The interview questions pertained to immersion, interaction and perception of control, and free communication about the general experience of the system.

The evaluation sessions involving participants were conducted in accordance with the ethical guidelines by the Norwegian Agency for Shared Services in Education and Research. Participants consented to participation anonymously and their interactions with the system were not recorded in any form. Additionally, their responses to interviews and questionnaires were collected without any personal data.

4.1 Pilot Study

The first iteration of the algorithm was tested in a pilot study. The aim of this study was to obtain a more free feedback from the participants about the interaction, and learn the patterns of interaction that are interesting for people. There were 5 participants in the study, 3 women and 2 men with ages ranging from 28 to 61. The study participants were all musicologists from the University of Oslo, with 10+ years of training in music. In 10-minute long interviews, we asked the participants about various kinds of feedback regarding both the motion as well as vocal interaction. Not everyone was interested in testing vocalization, and many commented on the nature of the task of vocalizing to not fit the purposes of a sound installation. Since vocal datasets were included in this first iteration, one participant also remarked that they were unable to distinguish at times if the vocal sounds they heard were previously produced by them.

Each of the 5 participants in this study remarked on thinking that the sound-bank was what was evolving, and not the effects. We realized that the soundbank like aspects of the design take precedence in the soundscape due to their salience. Effects-based sound design makes it harder to distinguish between changes and evolution. This creates an expectation in the participants to *want to hear* changes in soundbanks. Thus, we concluded that a system directly playing sound samples, rather than synthesizer-generated sound, would offer users clearer interaction and enhance the evolutionary changes over time. Following the results of this study, it was decided that P2 - the second prototype, would be based on a soundbank-based design.

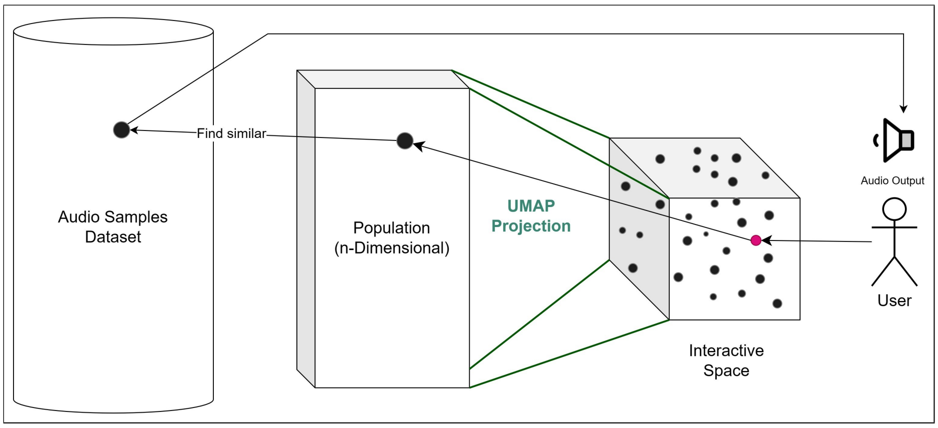

Figure 5. The system selects the nearest point within the 3D space based on user's position, determines the corresponding individual in the population, and plays the most similar audio sample from the dataset. (P2)

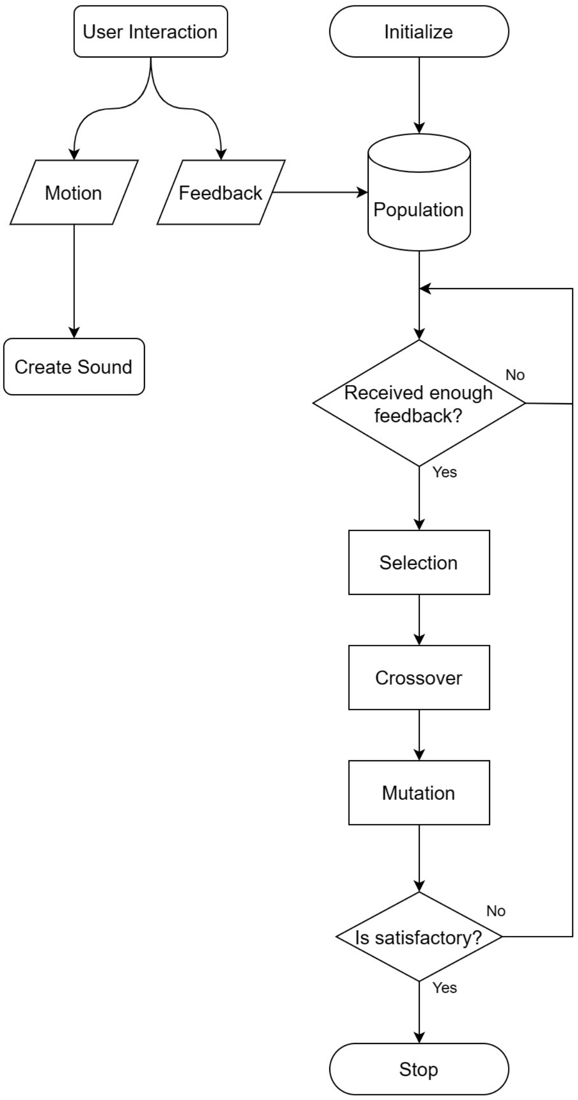

Figure 6. Flowchart of the IEMI Interactive Genetic Algorithm

4.2 Follow-Up Study

The follow up study consisted of 11 participants, with an average age of 30. The follow-up study was divided in two parts. In the first part, the participants interacted with the first prototype P1, while in the second they interacted with the second prototype P2. We were interested in capturing the differences in their experiences based on the first iteration of the experiment.

In this group, the range of years of musical training was more varied, with 45% of the participants having had less than 3 years of musical training and 45% having had over 10 years of musical training. Coding their feedback was based on the following categories of responses, each aligned with specific questions asked after the sessions:

- Degree of immersion — Did you feel immersed in playing with the installation?
- Degree of liking the interaction — Did you like interacting with the sounds?
- Degree of experience of control — Did you feel like you had control over the changes in sound?
- Degree of understanding how it works — Did you understand how the sound and movement worked together?
- Degree of experiencing change — Did you feel like the system changed over time?
- Degree of experiencing control over change — Did you feel like you had control over the change in the system?
- Degree of liking changes to the soundscape — Did you feel like the system changed in a way that you liked?

One participant remarked:

> *The spatial interaction was interesting, the second experiment was more immersive than the first one. Moving my arms allowed me to control the sound to a great degree and that was an engaging creative experience.*

Controlling sounds in space can become an inviting and exploratory task, where not restricting oneself to the interfaces of buttons, screens, or tools frees people to move as they like. Another participant here remarked:

> *I felt like the system could be played in a variety of ways which greatly influenced the overall experience in interesting ways.*

Participants remarked that the system became significantly more interesting once they realized that standing still could be used as a way of controlling the system.

4.3 Results

We focus on user feedback relating to qualitative and experiential aspects from the users. Out of the 16 participants involved in both P1 and P2 testing, more than half reported that they had fun using the system, and enjoyed navigating the soundscape through movement.

In the Follow-Up user study, we found that there was a greater degree or invitation to interact with the P2 system than P1. There is a large impact of the sound banks on the results of the surveys in these tests. As the surveys stand at the moment, it is impossible to separate the effects of sound design from the preference-related aspects of particular sounds that people like and dislike within the sound installation. The differences in the choices of soundbanks made a huge difference to the interaction of the system. As such, operationalizing this system with any soundbank is a possibility. Navigating point-clouds in 3D space with a target marker with visual as well as auditory feedback has potential in creating engagement and immersion.

5. DISCUSSION

In this paper, we described a prototype instrument for interacting with evolutionary algorithms controlling sound. We demonstrated two prototypes with a pilot user study and a final user study, indicating some factors that influence perception of control through unary feedback, and how people relate to this type of algorithm design. However, the two stand-out factors from these studies were in the "like-only" interaction paradigm, or unary feedback, as well as the perception of control.

5.1 Unary feedback for Engagement

How to obtain user feedback from participants in sound installations is a question that is even harder to answer at scale. This type of a question may also have different answers in different performance contexts. For example, in the context of a museum, engagement means moving as little as possible, whereas in the context of a club, maximal motion indicates maximal engagement. In the development of these tools, we considered these to be unary feedback, where negative feedback by the user is hard to capture. The system design was therefore based on a one-way feedback mechanism. We note that this type of feedback is quite common in popular algorithmic interactions on social media, where there has been no way to indicate displeasure and disengagement with content, and non-engagement can be interpreted both as indifference as well as disliking something. Some developers refer to this type of feedback as "Unary Implicit", where the data is similar to binary data, but in the real world the data is better represented as integers, since the 0 value of such a dataset is overloaded with both disliking, or never encountering something.

5.2 Perception of Control

When moving in space in order to control sound is established as a baseline condition within a task such as this, users can sometimes perceive controlling the sound when there is no control or change. Aspects of this are explored by various authors writing about agency and control in NIMEs [15–17], but we think this instrument prototype can be used to explore this further. Most participants in a study like this remark wanting to understand the system, and their interest in solving the puzzle of sonic interaction. The relationship between moving less in order to create a louder sound was perhaps the most elusive one for creating an illusion of control.

The results of the experiments suggested that the installation was capable of providing an intuitive musical experience for users with varying levels of musical background and age. They also provided valuable feedback on the system, highlighting areas for improvement such as a looper functionality or having a more musical case as the core of the sound design such as a drum sequencer.

In the qualitative interviews, study participants commented on their perception of control of the system. Since IEA-based systems take some time to evolve, the effects of the interaction are not immediately available for participants. In that way, change is both gradual as well as subtle. Participants liked the subtlety of this change, and got used to the time scale at which feedback occurred. However, this is both an advantage as well as a limitation while designing sight-specific works that change gradually.

5.3 Limitations and Future Work

Both the evaluation of P1 as well as P2 is with a small sample size focusing on qualitative and experiential aspects of the system. Overall, the results of the user study provide valuable insights into the potential of motion capture technology for embodied musical performance. The feedback from participants highlights areas for improvement and future development of the system. This study is a proof of concept study for incorporating interaction in evolutionary algorithms based interactive music systems.

As future work, we would like to test this system with a statistically significant user sample. The nature of the system is such that one could replace soundbanks to navigate through datasets of their own choice in an active listening scenario.

In general, an IEAs are hard to benchmark due to the freely changing nature of the system. To benchmark this, we plan to conduct systematic study using the same path and different algorithmic conditions to understand how systematic change can be measured in IEMI, and other IEAs.

Acknowledgments

We thank all the participants who generously shared their valuable feedback and gave us their time to interacting with this soundscape.

6. REFERENCES

[1] A. E. Eiben and J. E. Smith, *Evolutionary Computing: The Origins*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 13–24. [Online]. Available: https://doi.org/10.1007/978-3-662-44874-8_2

[2] H. Takagi, "Interactive evolutionary computation: fusion of the capabilities of ec optimization and human evaluation," *Proceedings of the IEEE*, vol. 89, no. 9, pp. 1275–1296, 2001.

[3] A. O. Effenberg, U. Fehse, G. Schmitz, B. Krueger, and H. Mechling, "Movement Sonification: Effects on Motor Learning beyond Rhythmic Adjustments," *Frontiers in Neuroscience*, vol. 10, p. 219, 2016.

[4] G. Dubus and R. Bresin, "A Systematic Review of Mapping Strategies for the Sonification of Physical Quantities," *PLOS ONE*, vol. 8, no. 12, p. e82491, Dec. 2013, publisher: Public Library of Science. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0082491

[5] A. Aska and M. Ritter, "Approaches to Real Time Ambisonic Spatialization and Sound Diffusion using Motion Capture," Jan. 2016.

[6] A. Tanaka and M. Donnarumma, "The Body as Musical Instrument," in *The Oxford Handbook of Music and the Body*, Y. Kim and S. L. Gilman, Eds. Oxford University Press, Oct. 2019, p. 0. [Online]. Available: https://doi.org/10.1093/oxfordhb/9780190636234.013.2

[7] A. Jensenius, M. Wanderley, R. Godøy, and M. Leman, "Musical Gestures: concepts and methods in research," *Musical Gestures: Sound, Movement, and Meaning*, Jan. 2009.

[8] M. Mainsbridge, *Body as Instrument: Performing with Gestural Systems in Live Electronic Music*. Bloomsbury Publishing USA, Feb. 2022, google-Books-ID: wGxWEAAAQBAJ.

[9] P. Dahlstedt and A. Skånberg, "Otokin: Mapping for sound space exploration through dance improvisation," vol. Paper i proceeding 2019, 08 2019.

[10] J. Fornari, "An evolutionary algorithm to create artificial soundscapes of birdsongs," *International Journal of Arts and Technology*, vol. 9, p. 39, 03 2016.

[11] C. Erdem, B. Wallace, and A. R. Jensenius, "CAVI: A Coadaptive Audiovisual Instrument–Composition," in *NIME 2022*. The University of Auckland, New Zealand: PubPub, Jun. 2022. [Online]. Available: https://nime.pubpub.org/pub/cavi

[12] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," *arXiv preprint arXiv:1802.03426*, 2018.

[13] E. Fonseca, X. Favory, J. Pons, F. Font, and X. Serra, "Fsd50k: An open dataset of human-labeled sound events," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 829–852, 2022.

[14] P. Dahlstedt, "Evolution in Creative Sound Design," in *Evolutionary Computer Music*, E. R. Miranda and J. A. Biles, Eds. London: Springer London, 2007, pp. 79–99. [Online]. Available: https://doi.org/10.1007/978-1-84628-600-1_4

[15] G.-M. Schmid, "Evaluation of musical instruments from the musician's perspective," Ph.D. dissertation, University of Basel Basel, Switzerland, 2015.

[16] W. Brent, "Perceived control and mimesis in digital musical instrument performance," 2012.

[17] C. Erdem, Q. Lan, and A. Jensenius, "Exploring relationships between effort, motion, and sound in new musical instruments," *Human Technology*, vol. 16, pp. 310–347, Nov. 2020.