Human-Swarm Interactive Music Systems: Design, Algorithms, Technologies, and Evaluation

Pedro Lucas and Kyrre Glette

RITMO Centre for Interdisciplinary Studies in Rhythm, Time, and Motion
Department of Informatics
University of Oslo
Oslo, Norway
pedroplu@uio.no

Abstract. This paper presents considerations for developing Human-Swarm Interactive Music Systems (IMS), based on previous work in the field. We discuss design principles, algorithms, technologies, and evaluation methods for creating user-centred Human-Swarm IMSs using architectural approaches, swarm strategies, and levels of embodiment in implementation. Our contribution aims to establish a framework for future applications and research studies on swarm-based music platforms.

Keywords: Interactive Music Systems, Digital Instruments Design, Swarm Intelligence, Multimodality

1 Introduction

Sending, processing, and response are three stages that form a concise and straightforward model to represent Interactive Music Systems (IMS). However, the different contexts in which an IMS can be developed give rise to several levels of complexity, demanding a critical cross-disciplinary investigation. This expands the model to more concrete representations and design considerations for innovative applications [9].

This paper focuses on a specific instance of an IMS related to a Human-Swarm system. This type of IMS refers to improvisational systems that allow a user to interact with a swarm of artificial agents that are self-organized (working locally without a central controller) and exhibit emergence (interaction between agents in the swarm produces higher-level patterns and structures) [32]. These and other properties are commonly based on the theory of *Swarm Intelligence*, which can be found in nature and has been modelled in computational simulations.

This type of IMS is important in its potential to develop various levels of representation of sonic and/or musical units, ranging from micro sounds for granular synthesis to the embodiment of individual artificial musicians capable of collaborating with human performers to achieve complex music improvisations.

Depending on the levels of representation, modelling an IMS as a Human-Swarm system can have benefits. In the case of music improvisation, musical elements can be highly interactive and uncertain. Therefore, swarm strategies are a good fit for a process that can reproduce such behaviour for real-time music composition [1]. Additionally, embodied representations of artificial musical agents with the role of additional musicians can lead to collaborative and enjoyable human-machine experiences [14].

To advance the development of Human-Swarm systems, we contribute with the proposal of a framework that includes four relevant areas: *design considerations, algorithms, technologies, and evaluation methods*. This proposal is based on previous work on swarm intelligence applied to IMS, theoretical explorations of multi-agent systems that use swarms in music, and musical agents. A thematic analysis approach was used to extract information from these works, having these four areas as central themes. Our contribution is intended to support applications and research studies concerning music platforms that use swarm approaches, as well as provide a foundation upon which creativity can be effectively channelled.

This paper is organized as follows: Section 2 presents a background of Human-Swarm IMSs. Section 3 presents the framework including a detailed discussion focused on the four areas mentioned above. Finally, Section 4 provides conclusions and future directions for Human-Swarm IMSs.

2 Background and Related Work

The interest in musical interaction with artificial swarms began with Blackwell and Bently's work [2], where they proposed the first application of swarm intelligence to music. They related music features to swarm descriptors, such as attraction and repulsion, suggesting that improvised music is a self-organized system that can lead to complex musical structures. This self-organization is carried out by local interactions between individuals and the environment, which can be direct or indirect. Indirect interactions are mainly focused on in some works [4] [1] [32], considering the concept of *Stigmergy*, which is a mechanism that manifests when an individual modifies characteristics of the environment so that other individuals respond to it later.

In most swarm applications, the elements that participate in self-organization interactions can vary in terms of the size of the musical material. Blackwell [1] presented a classification based on perceptual time-scales, which can be seen as musical material elements organized by size. The elements as events are: *micro* (small-scale times like tenths of a millisecond), *mini* (musical notes or sound objects), *meso* (phrases or groups of mini-events), and *macro* (time encompasses form and lasts several minutes or more). This classification is also useful to determine the level of embodiment that the agents from swarms can have regarding their interaction with human performers, which is reflected in the works described below.

The usual strategies that utilize swarm intelligence in music systems are focused on mappings of sonic or musical features over spatial properties in swarms. The musical interaction is given by the swarm dynamics, which commonly has led to interactive solutions in which the agents from the swarm are hidden elements with a low embodied perception. This concept is portrayed by *Swarm Music* [2] [4], which is based on flocking algorithms and a process of capturing, updating, and interpretation so that users can modify the dynamics of the swarm for influencing the musical input. Another relevant work is *Musebots* [6], which explores the concept of Musical Metacreation (MuMe) related to the automation of aspects regarding musical creativity to model a musician more than an instrument, and thus closer to working on music improvisation. A higher embodiment can be achieved through visual feedback and gestures in a 3D environment to display agents, as the work of Unemi and Bisig [30], which shows an interactive installation where the user acts as a conductor for influencing flock's musical activity; moreover, agents can also perceive aspects of musical outputs and operate in a 3D space as virtual sound sources, as shown in [7] and [23]. Physical implementations develop mappings with spatial or sonic properties from entities as robots, as described in works such as [31], [33], and [13]. Other approaches include using quantum physics simulations [16] and physical-virtual environments that portray full embodiment with agents as musicians [14].

Theoretical frameworks that support swarm applications have been explored for Human-Swarm IMS. In this case, we have the concept of *Musical Agents*, which are entities as computer programs that generate music autonomously or in collaboration with human musicians [25]. These entities can be part of Multi-Agent systems, such as the *Virtual Musical Multi-Agent System (VMMAS)* [34] and the *Mobile Musical Agents* project based on the *Andante* project, which deals with musical agents that decide to migrate and react to changes in the environment [29]. Architectures under these theoretical structures have been proposed, such as *MAMA* [19] [18], which is grounded on the theory of communicative acts and enables agents to reason about intentionality, or the *MASOM* architecture [24] that works with *Self-Organizing Maps* based on musical agents, that has been used in works such as REVIVE [26] [27], and Spire Muse [28]. Additionally, an approach that involves improvisation with human interaction was elaborated and presented as a concept called *Live algorithms* [3] for representing analysis, process, and synthesis modules for IMSs in the human-machine domain.

When it comes to Human-Swarm interaction and collaboration, it is essential to consider how agents can work together, which can be achieved through *negotiation behaviours* to satisfy the interests of the individual agents, such as in [10]. Synchronized works, as in the case of those based on pulse-coupled oscillators inspired by fireflies and implemented as fireflies [21] [20], or self-synchronization with percussive robots that achieve equilibrium [13], are also examples of collaboration. Interaction and collaboration can be conceptualized in terms of influence and motion, as seen in the system *Swarm Lake* [12], which also uses a game development approach for its design and considers environmental features to conceptualize a theme in a hypothetical world presented to the user. Moreover, it is possible to have higher levels of control for swarm collaboration considering *swarm dynamics* (e.g. swarm-wide; that is, control over a group more than an individual) instead of *direct control* of sound parameters (e.g. audio volume). Control regarding swarm dynamics is present in most related works and significantly affects the resulting music [32].

In summary, most of the previously cited works that involve systems use a swarm representation to map sonic or music features, which can be based on different musical material sizes. The complexity for some of them rises in a final musical piece that can be achieved in an improvisation musical session together with a human performer, but others can reach a higher level and become actual artificial musicians interacting with each other and with the user, which demands more sophisticated ways to develop and represent agents. We are mainly interested in this last type of system to remark the embodiment of agents in a swarm, but without discarding the possibility of building solutions with more abstract representations for lower levels of embodiment. As the human is part of the system, this work intends to provide means to increase the understating of a swarming process in human-machine music performances.

3 Human-Swarm IMS Framework

The section presents a framework to enhance the creation process of a Human-Swarm IMS. The developer can start to look at the general considerations described below to create a unique solution, then specify the architecture to use and check if the solution complies with the swarm design properties listed later. Moreover, the sound generation can follow mapping strategies according to the nature of the designed swarm, and suitable algorithms can be implemented to support that design. Finally, the technologies to choose would depend on the design and the available resources.

All these considerations are presented and discussed below.

3.1 Design Principles

The design and development of IMSs have been explored in a variety of works for several years [8] [9] [15], emphasizing user interaction, system design, and mapping strategies. In this work, we want to provide a more specific scenario for Human-Swarm IMSs which have used implicitly or explicitly the design approaches explored before. In consequence, we present in this section a set of design principles based on previous work related to Swarm Intelligence applied to IMSs.

3.1.1 General Considerations

The following sections focus on specific considerations regarding architectures, swarm design, and sound mappings. On top of this, other considerations are recommended to develop a Human-Swarm IMS as illustrated commonly in literature, such as:

- **Idiosyncratic Approach:** Design is mostly a personal choice [1], and that is reflected in IMSs that want to achieve specific goals which are recommended to be primarily related to artistic intentions and creative process more than technological-driven motivations. This is also called a practice-driven approach [17]. However, guidance in this process is relevant for a solid structure that supports those personal choices, and this paper intends to suggest such guidance.
- **Representation and Dynamics:** Two significant decisions are required to design a swarming system: *representation* and *dynamics* [1]. The *representation* has to do with inputs and outputs and how they are processed, and *dynamics* is the swarm algorithm that interacts with the representation. These decisions are based on an architectural approach that is explained in Section 3.1.2.

-Novelty: We can achieve novelty through self-organized approaches considering three aspects: music representation, music style definition, and music style evolution [11]. These aspects can be explored in the results obtained from the system. Finding ways to have a fast switch between instances of these aspects helps to fine-tune our musical intentions.

-External Inspiration: The design of a Human-Swarm IMSs can approach several levels of embodiment, which require integrating multiple disciples in complex cases. Thus look at other areas such as game design (e.g. *Swarm Lake* [12]) and human-robot interaction (e.g. *Dr.Squiggles* [13]) can enable several possibilities to enhance the experience.

3.1.2 Architectures

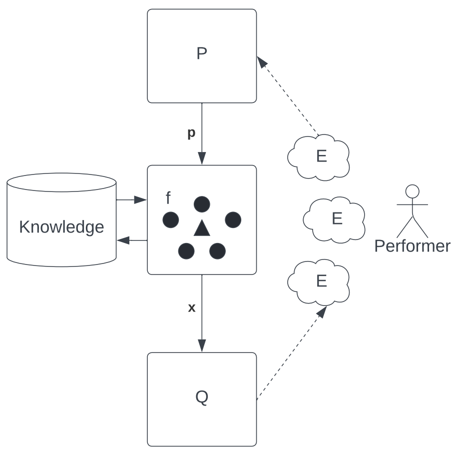

Fig. 1: *PQf+K* Architecture. This structure is based on Blackwell's work [1] with the addition of a knowledge base for swarm dynamics.

Works on Human-Swarm IMSs usually depict specific architectures based on the particular problem they are solving. However, especially in theoretical works, there are proposals where system modularization can lead to a clear design base. We depart from the simplest *sending, processing, and response* stages to model an IMS as in [9], which can be seen as the traditional minimized structure *input, process, output*. This model can be expanded to a complex set of units to describe a system to *capture, update, and interpret* information in the environment in which a human performer is a participant [2] [4], and if we want to see them closer to the human process of improvising music, we can portrait them as *perception, cognition, and musical execution* [34].

For Human-Swarm IMSs, we need structures that encourage coexistence between the human performer and the artificial entities, thus considering the models mentioned above, the concept of *Live Algorithms* developed by Blackwell et al. [3] is a suitable choice of representation. A Live Algorithm is "an autonomous music system capable of human-compatible performance... the Live Algorithm listens, reflects, selects, imagines, and articulates its musical thoughts as sound in a continuous process", hence a Live Algorithm works with collective human-machine musical improvisation. This concept is structurally represented by the *PQf* architecture proposed in [1], having *P* for analysis, *Q* for synthesis, and *f* for patterning supporting the two major decisions mentioned previously: representation (*P, Q*) and dynamics (*f*).

We propose to add an explicit module to this representation called *Knowledge* since there are applications that require a knowledge base for the dynamics depending on the algorithm that is being used as in [34] that applies a fuzzy mechanism, or in [19] that uses a knowledge base for musical agents based on communicative acts. Fig. 1 illustrate this proposal as the *PQf+k* architecture.

The knowledge can itself be modelled with high complexity; however, it is sensible to take into account limitations and the trade-off of using a knowledge base in a real-time setup since IMSs are improvisational systems and potential problems like *latency* can affect the user experience significantly.

The advantage of this modularization is the flexibility to change among strategies so that system properties are adjusted in real-time if needed (e.g. change the knowledge base or swarm algorithm in the middle of the performance); that is why a particular emphasis on this architecture is given for the interfacing between modules.

Another useful approach is using a *Finite State Machine (FSM)* to model an individual agent behaviour or the external influences of the human performer, which can exhibit different states when the performer interferes in the environment [23]. Moreover, FSM can help to minimize the complexity of designing multimodal systems, which is relevant, especially for Human-Swarm IMSs that target higher embodiment [5].

3.1.3 Swarm Design

Commonalities found in previous work referenced in Section 2 related to Human-Swarm IMSs lead us to propose the following design principles:

-Decentralization: Even though most swarm systems are developed over a centralized platform, the nature of a swarm should target decentralization, which implies looking for local communication methods and rules between individuals and the environment to portray independence from global management. Inspiration of decentralized behaviours can be found in animal swarms.

-Emergence: Emergent behaviour allows a swarm to create dynamic and unpredictable musical outcomes. This is also known as *self-organization*, which arises from the collective actions of individuals. The system should allow the emergence of complex and adaptive behaviours from the interactions of individuals, resulting in unique and creative musical compositions that are co-created by the swarm.

-Stigmergy: As mentioned earlier, this mechanism manifests when an individual modifies characteristics of the environment so that other individuals respond to it later. Modelling stigmergy can be useful for indirect control through the environment and limit direct interaction with agents when it is not entirely possible (e.g. interaction with a swarm of physical drones).

-Scalability: The design should support the accommodation of various agents, ranging from small groups to large crowds. This feature is the system's scalability in terms of technical infrastructure and user experience, which ensures that the system handles different swarm sizes and that the interaction remains meaningful and enjoyable, regardless of the group size when the design allows it.

-Stability: For some swarm systems in which agents can fail individually (e.g. each agent can be a physical robot that could potentially withdraw), the musical task should continue with the rest of the participants and the consequences of losing some of them should not impact, at least, the essence of the performance. Consideration of this aspect results in a more stable swarm system.

- **Flexibility and adaptability:** Systems should be flexible and adaptable to different musical styles, genres, and contexts, as the goal is music improvisation. The system should allow for customization and configuration to suit different musical perspectives and should be able to adapt to changes regarding the swarm's size, behaviour, or musical preferences over time.
- **Time-scale of Material:** Depending on the system's focus, the sonic or musical material in terms of duration can be framed as *micro, mini, meso,* or *macro*, as described earlier. The solution's complexity level could rise as time increases since more sophistication is required for higher levels like *macro*, which deals with complex musical structures.
- **Level of Embodiment:** The swarm individuals can be conceived as mere abstract units that contribute to a musical solution, which can be hidden from the user to a certain extent. However, if these individuals are closer to artificial musicians to collaborate with, it is necessary to provide a level of embodiment that transcends into the spatial domain. In that sense, multimodal approaches through 3D environments are helpful, which could require spatial audio solutions and visualization strategies.
- **Environmental Perception and Actuation:** The swarm system should sense the environment to respond accordingly with actions through direct interaction or by stigmergy. Thus it requires defining and designing sensing capabilities according to the level of embodiment and decentralization as well as suitable output mediums. For instance, a robot swarm can be equipped with microphones and speakers for music sensing and actuation in the environment.
- **Level of Control:** Human-Swarm IMSs require a certain level of control from a human operator, in which the designer should define how much of this control is provided from a fully manual operation to a completely autonomous system. Allowing a real-time definition of these levels could increase the diversity of the music material produced by human-machine improvisations. Additionally, controls can act over the swarm dynamics, sound parameters, or higher descriptors as commands.
- **Feedback and Transparency:** To support decision-making during music improvisation, it is essential that the actions performed by the swarm are *transparent* to the user. This can be achieved through adequate feedback from the artificial agents and any human operators involved in the system. Auditory feedback is particularly important in an IMS, but visualization and haptic feedback can also be useful for confirming actions. However, designers need to be careful not to overwhelm the user with too much information and consider whether certain types of feedback might go against the artistic purposes of the system.
- **Accessibility and Inclusivity:** The design can consider an inclusive and accessible system for diverse participants, including individuals with different abilities, backgrounds, and musical skills. If the intention is to cover a wide variety of performers, the design should consider multiple modes of participation and accommodate different levels of physical, cognitive, and musical abilities, ensuring that everyone can participate and contribute to the music-making process.
- **Trust:** Building trust with a non-human agent requires calibration between a person's expectations of the agent and the agent's capabilities. Exploration of trust at different levels might significantly enhance the musical result.

-Room for Failure: We can design a system with a high amount of constraints, but it could restrict potential interesting results that can emerge from the music improvisational process; thus, to encourage the element of surprise in the results, we can leave some room for failures and user exploration in that context.

Several of these principles overlap and belong mostly to the swarming nature of the solution, which mainly deals with spatial properties.

3.1.4 Mapping Strategies

The most common mapping strategies for Human-Swarm IMSs relate sonic or music parameters to spatial properties; for instance, amplitude and pitch from a specific sound sample could be associated with coordinates X and Y of an agent, and music can emerge from the swarming behaviour. These associations can be simple and direct, as the example provided, or use non-linear or probabilistic approaches; it depends on personal choices and the designer's goals.

The previous example considers `swarm-sound/music` mapping; however, the interaction with a user demands establishing `human-swarm` mapping strategies. In that sense, apart from usual ways to feed musical input (e.g. using MIDI controllers), motion capture techniques for gestures, or other sensing solutions, can be used to manipulate swarm parameters to have a `human-swarm-sound/music` mapping; nevertheless, an option of `human-sound/music` mapping can be combined with swarm dynamics depending on the design.

As we deal with swarms, mappings can also focus on the dynamics of collective actions and general descriptors. For instance, as the swarm explores the spatial environment, the centre of mass can be a parameter that influences higher musical features, like the global panning or a general reverb effect, which can also have more complex interaction in terms of the behaviour of every individual, leading to a dense music result.

For certain applications, especially in a physical domain, there could be noises with a significant effect on the sonic result (e.g. motors, propellers, etc., from robot swarms); in that case, we can include these sounds as part of the performance by processing them through mapping strategies that allow their inclusion to the musical result.

Consequently, we can create a rich and engaging musical experience through a mapping design that encourages the participation of all actors and situations while allowing individual expression and creativity from the user, according to adjustable levels of autonomy in the system.

3.2 Algorithms

Based on the works listed in Section 2, we can identify common strategies for handling the *input*, the *processing algorithm*, and the *synthesis of sonic output* or other useful feedback, as described below.

-Input: The audio stream of a music performance is a typical source of input. It can be analyzed using signal processing techniques to extract features for further usage, such as loudness, pitch, and onsets. Musical material can also be collected directly from human performers through common interfaces like musical keyboards or traditional instruments. However, complex control mediums like gestures and image recognition require sophisticated capture strategies. In such cases, machine learning algorithms for real-time data collection can be useful for these tasks by applying classification techniques to identify discrete states and regression strategies for continuous values.

-Process: Common *Swarm intelligence* approaches use flocking strategies based on the *Reynolds's boids algorithm*, in which agents have attraction and repulsion rules concerning neighbours as well as velocity matching. These rules can be structured on reasoning mechanisms that take advantage of descriptive parameters through algorithms such as *fuzzy logic* or *language processing through communicative acts*. Other proposals consider mathematical models that define acceleration or velocities for the agents' position calculated from local individuals and the performer's spatial features.
Additional techniques used in this category include *Particle Swarm Optimization (PSO)*, *Ant Colony Optimization (ACO)*, and *Genetic Algorithms*. However, in some cases, the goal is not to optimize specific parameters but to fulfil musical intentions that take advantage of the algorithm's mechanics. Other strategies, such as *Self-Organizing Maps*, can be used for sound organization and pattern recognition. Music generation through real-time input and pre-loaded knowledge as *Markov Chains*, can lead to interesting results. Synchronization techniques, such as *Pulse-Coupled Oscillators* inspired by the behaviour of fireflies or custom strategies based on the analysis of temporal events in the audio stream, can be applied to rhythm.
Switching between algorithms requires that they share similarities in a swarm. The selection of behaviours determines the overall structure of the swarm, while the weighting of different behaviours affects the current dynamics of the simulation.

-Output: The output depends on the mapping between the swarm's spatial properties and the sonic and musical result. Possible mapping strategies include *additive synthesis*, *granular synthesis*, *control based on agents' proximity*, *procedural patching from swarm dynamics*, *modulation synthesis*, and *sound physical modelling*. The choice of mapping depends on the specific musical goals.

The designer can decide the suitable technique to use, and there is plenty of room for applications and research studies regarding algorithms that can be explored at different levels of embodiment, so the user experience has to be taken into account as a centre point of departure to develop a system that characterizes the nature of the musical interaction between human and machine.

3.3 Technologies

We classify potential technologies to use into three categories according to their level of embodiment, as described below.

-Virtual: In this category, agents exist solely in a virtual environment implemented through software on a central device, such as a computer. Input is received via integrated peripherals, MIDI keyboards, or sophisticated devices such as cameras with image recognition algorithms. Sonic output is played through loudspeakers, ranging from a simple mono configuration to multiple channels for spatial audio. While complexity can increase in terms of input and output devices, processing remains centralized, and agents are virtual objects that can produce music as a hidden process or with a higher representation visualized on a screen.

-Physical-Virtual: This category builds on the previous virtual category, but agents reach a higher level of embodiment by sharing the physical space with the performer and being aware of the real environment. Extended reality technologies, such as mixed reality headsets or augmented reality systems, can support this configuration. For a more immersive experience, it may require additional complexity in terms of motion capture, visualization, and audio playback to portray a virtual 3D world that overlaps the physical space where the performance is happening.

-Physical: In this category, agents exist as actual entities, such as robots, which can interact with the human performer. Design principles for human-robot interaction can be applied, and additional considerations such as trust, safety, and treatment of noises are considered. Each agent requires its own input and output capabilities and capacity for local communication, as this category can be approached as a decentralized system.

As technology advances, we can improve the response time for the interaction, integrate better ways to reach transparency, and potentially extrapolate to the participation of larger groups to the performance (e.g. audience with no musical skills).

3.4 Evaluation Methods

Evaluation methods have been proposed before, such as in the work of O'Modhrain [22] that describes methodologies depending on the stakeholder and recommends clearly understanding of what to apply and to whom depending on the interest of the study. In that sense, Human-Swarm systems are focused on the *performer/composer* and the *designer*. The following types of evaluations can be considered to assess these systems.

-Autoethnography: The designer can evaluate the system by using it and reflecting on the music creation process to improve the design.

-Observation: The system can be used by different users in different settings, such as a controlled environment like a laboratory or a concert. The designer can observe the advantages and limitations in those environments to understand how different users can approach the system.

-System Measurements: The designer can measure sections of interest in the system to discover limitations that can impact the user experience, such as latency or jitter.

-Physical and Physiological measurements: For user studies, data can be captured while participants use the system. Physical data, such as positions in space, can be useful for higher embodiment applications, and physiological measurements can give insights into the user's state while performing. An important consideration is that the measurement methods should not interfere with the performance.

-Surveys: We can evaluate the user's response to the system by applying surveys before to gather expectations and, commonly after, to collect points of interest that help to improve the user experience.

We suggest integrating these methods in alignment with the system and the designer's goals. It is important to prioritize the user's experience and the quality of the music created during the evaluation process.

4 Conclusions

This paper presents previous work on Human-Swarm Interactive Music Systems to distil design principles, algorithms, technologies, and evaluation methods to establish a framework for swarm-based music platforms. We organize this information so that designers can explore novel solutions for performers, and researchers can have additional support to contribute to this field.

We do not intend to provide a strict recipe for Human-Swarm IMSs but a starting guide to propose specific principles that work for particular projects, which can increase and optimize the definition of new approaches for future applications.

For future work, we plan to use this framework to create multiple music platforms and enhance these suggestions through research and data analysis.

References

1. T. Blackwell. Swarming and Music. In E. R. Miranda and J. A. Biles, editors, *Evolutionary Computer Music*, pages 194–217. Springer London, London, 2007.
2. T. Blackwell and P. Bentley. Improvised music with swarms. In *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02*, volume 2, pages 1462–1467, Honolulu, HI, USA, 2002. IEEE.
3. T. Blackwell, O. Bown, and M. Young. Live Algorithms: Towards Autonomous Computer Improvisers. In J. McCormack and M. d'Inverno, editors, *Computers and Creativity*, pages 147–174. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.
4. T. Blackwell and M. Young. Self-organised music. *Organised Sound*, 9(2):123–136, 2004.
5. M.-L. Bourguet. Designing and Prototyping Multimodal Commands. *Human-Computer Interaction (INTERACT'03)*, pages 717–720, 2003.
6. O. Bown, B. Carey, and A. Eigenfeldt. Manifesto for a Musebot Ensemble: A platform for live interactive performance between multiple autonomous musical agents. *International Symposium on Electronic Art 2015*, 2015.
7. P. Codognet and O. Pasquet. Swarm intelligence for generative music. In *ISM 2009 - 11th IEEE International Symposium on Multimedia*, pages 1–8. IEEE, 2009.
8. P. R. Cook. Re-Designing Principles for Computer Music Controllers: A Case Study of SqueezeVox Maggie. *NIME 2009*, 2009.
9. J. Drummond. Understanding Interactive Systems. *Organised Sound*, 14(2):124–133, 2009.
10. A. Eigenfeldt and P. Pasquier. Creative Agents, Curatorial Agents, and Human-Agent Interaction in Coming Together. *Sound and Music Computing Conference (SMC2012)*, 2012.
11. M. Gimenes, E. R. Miranda, and C. Johnson. Musicianship for robots with style. In *NIME '07*, page 197, New York, New York, 2007. ACM Press.
12. M. Kaliakatsos-Papakostas, A. Floros, K. Drossos, K. Koukoudis, M. Kyzalas, and A. Kalantzis. Swarm Lake: A Game Of Swarm Intelligence, Human Interaction And Collaborative Music Composition. *ICMC 2014*, 2014.
13. M. Krzyżaniak. Musical robot swarms, timing, and equilibria. *Journal of New Music Research*, 50(3):279–297, 2021.
14. P. Lucas. A Human-Machine Music Performance System based on Autonomous Agents. Master's thesis, University of Oslo, 2022.
15. 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*, pages 23–40. Springer International Publishing, Cham, 2019.

16. M. Mannone, V. Seidita, and A. Chella. Quantum RoboSound: Auditory Feedback of a Quantum-Driven Robotic Swarm. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 287–292, Napoli, Italy, 2022. IEEE.
17. J. McCormack, P. Hutchings, T. Gifford, M. Yee-King, M. T. Llano, and M. D'inverno. Design Considerations for Real-Time Collaboration with Creative Artificial Intelligence. *Organised Sound*, 25(1):41–52, 2020.
18. D. Murray-Rust. *Musical Acts and Musical Agents: theory, implementation and practice*. PhD thesis, University of Edinburg, 2008.
19. D. Murray-Rust, A. Smaill, and M. Edwards. MAMA: An architecture for interactive musical agents. In *Frontiers in Artificial Intelligence and Applications*, volume 141, pages 36–40, 2006.
20. K. Nymoen, A. Chandra, K. Glette, and J. Torresen. Decentralized harmonic synchronization in mobile music systems. *2014 IEEE 6th International Conference on Awareness Science and Technology, iCAST 2014*, 2014.
21. K. Nymoen, A. Chandra, and J. Torresen. The challenge of decentralised synchronisation in interactive music systems. *Proceedings - IEEE 7th International Conference on Self-Adaptation and Self-Organizing Systems Workshops, SASOW 2013*, pages 95–100, 2013.
22. S. O'Modhrain. A Framework for the Evaluation of Digital Musical Instruments. *Computer Music Journal*, 35(1):28–42, 2011.
23. J. C. Schacher, D. Bisig, and M. T. Neukom. Composing with swarm algorithms - creating interactive audio-visual pieces with flocking behavior. In *International Conference on Mathematics and Computing*, 2011.
24. K. Tatar and P. Pasquier. MASOM: A Musical Agent Architecture based on Self-Organizing Maps, Affective Computing, and Variable Markov Models. *The 5th International Workshop on Musical Metacreation (MuMe 2017)*, 2017.
25. K. Tatar and P. Pasquier. Musical agents: A typology and state of the art towards Musical Metacreation. *Journal of New Music Research*, 48(1):56–105, 2019.
26. K. Tatar, P. Pasquier, and R. Siu. REVIVE: An Audio-visual Performance with Musical and Visual AI Agents. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, volume 2018-April, pages 1–6, New York, NY, USA, 2018. ACM.
27. K. Tatar, P. Pasquier, and R. Siu. Audio-based Musical Artificial Intelligence and Audio-Reactive Visual Agents in Revive. In *Proceedings of the International Computer Music Conference and New York City Electroacoustic Music Festival*, 2019.
28. N. J. W. Thelle and P. Pasquier. Spire Muse: A Virtual Musical Partner for Creative Brainstorming. In *NIME 2021*. PubPub, 2021.
29. L. K. Ueda and F. Kon. Mobile musical agents. In *Companion to the 19th annual ACM SIGPLAN conference on Object-oriented programming systems, languages, and applications - OOPSLA '04*, page 206, New York, USA, 2004. ACM Press.
30. T. Unemi and D. Bisig. Playing Music by Conducting BOID Agents - A Style of Interaction in the Life with A-Life. *Conference on the Simulation and Synthesis of Living Systems*, 2004.
31. Y. Uozumi, M. Takahashi, and R. Kobayashi. A Musical Framework with Swarming Robots. In R. Kronland-Martinet, S. Ystad, and K. Jensen, editors, *Computer Music Modeling and Retrieval. Sense of Sounds*, volume 4969, pages 360–367. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.
32. D. Waghorn. Controlling emergence in an interactive multi-agent musical system. Master's thesis, Monash University, 2016.
33. J. Wong, C. Williams, and V. Mirecki. The Application of Swarm Robotics in Music. *Archives of Worcester Polytechnic Institute*, 2020.
34. R. D. Wulfhorst, L. Nakayama, and R. M. Vicari. A multiagent approach for musical interactive systems. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems - AAMAS '03*, page 584, New York, USA, 2003. ACM Press.