Compositional Considerations in the Design of an Interactive Agent-based Musical Ecosystem

Zeynep Özcan
Independent Artist
ozcanz@icloud.com

Anıl Çamcı
University of Michigan
acamci@umich.edu

ABSTRACT

In this paper, we explore the application of a biologically inspired artificial ecosystem to real-time interactive music composition. Particularly, we examine the uses of a food chain hierarchy and animal behavior to create emergent and complex compositional structures. We discuss the algorithmic creation of musical forms through the temporal layering of agent species that populate a musical ecosystem. We also explain the compositional implications of situating the listener as an interactive agent in this system. We provide examples from our ongoing research project that investigates the compositional affordances of interactive augmented sonic environments and how listeners experience such works.

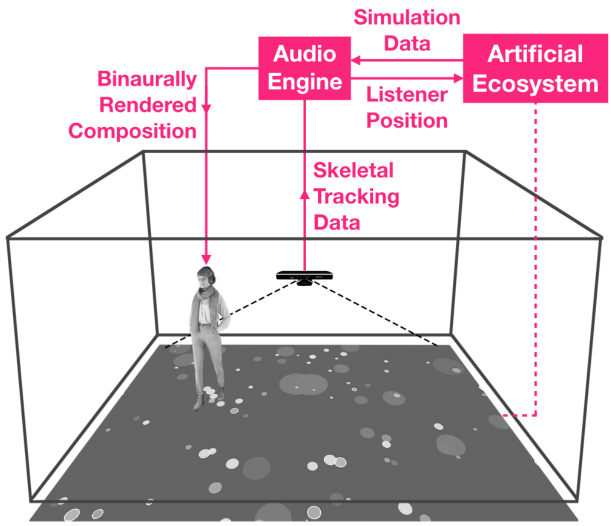

Figure 1. The data flow across the components of *Proprius*, including an artificial ecosystem implemented in Processing, an audio engine which sonifies the various agents that make up this ecosystem, and a position-tracking system that both situates the listener as an interactive agent in the ecosystem and renders a binaural output of the system relative to the listener's position. The visual representation of the simulation is underlain here to illustrate the listener's navigation of the ecosystem.

1. INTRODUCTION

The artistic potential of artificial ecosystems has been extensively used in multimedia arts. Such systems often give way to emergent behavior based on computational models such as swarming simulations [1], L-Systems [2], evolutionary algorithms [3], and ecosystemic models [4]. The organisms in these ecosystems are modeled in the form of software entities called agents, which display life-like intelligent behavior and autonomy. The concept of autonomy is employed by many artists to give self-organizing properties to software-based systems. Autonomous agents often display the ability to observe their environments and make decisions accordingly; they can also interact with other agents with varying degrees of complexity based on rule systems. Inhabiting complex dynamic environments, these agents act towards “a set of goals or tasks for which they are designed” [5]. Similar to those in natural ecosystems, artificial agents are often functions of energy flow and nutrition cycles. Occupying hierarchical structures such as food chains, these agents trade token units (e.g., energy, biomass) in systems that simulate community dynamics [6].

In this paper, we discuss the design of one such artificial ecosystem that generates music in real-time. We describe our ongoing research into autonomous interactive sonic environments where the sonification of animal behavior within an ecological simulation is used as a means to create augmented reality music compositions [7]. In our installation *Proprius*, the composer imparts creative license to not only the agents of a virtual ecosystem but also the listener that interacts with these agents in augmented space [8]. Here, we discuss this system's ability to create musical elements at various structural units ranging from gesture to form: we look at the inherent musical abilities of the agents in terms of their behaviors and their sonification characteristics, which are themselves inspired by natural properties. We then look at how the listener's involvement in this system as an interactive agent affects the work's progression. While doing so, we provide reports and insights gained from previous performances of this work.

2. IMPLEMENTATION

Fig. 1 illustrates the overall implementation of *Proprius*. The artificial ecosystem and the behaviors of the agents that populate it are implemented using the multimedia programming platform Processing [1]. The output of the simulation is visualized in the form of an overhead view of the ecosystem, as seen in Fig. 2. This visualization serves as an interactive score of the piece, allowing the composer to view the progression of the work in terms of the agent pop-

[1] https://processing.org

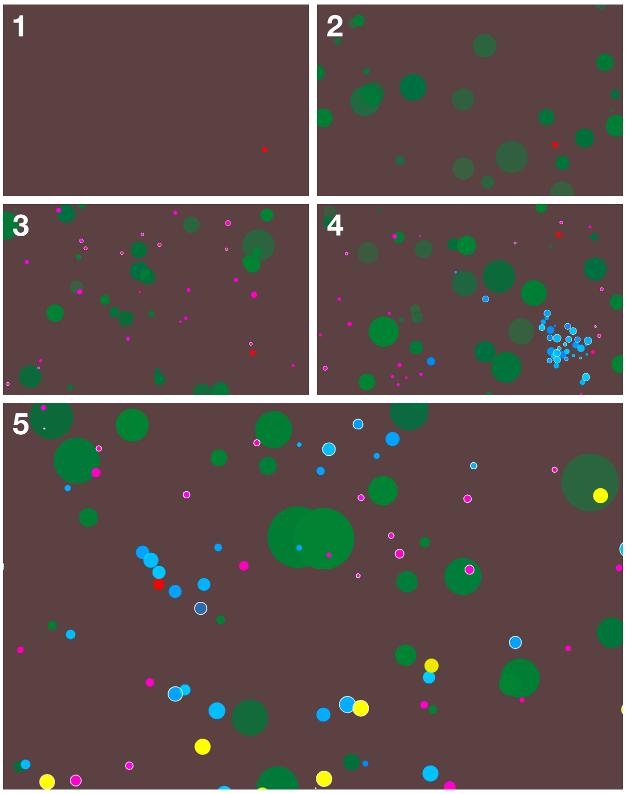

Figure 2. Screenshots from the five scenes of *Proprius*, showing the progression of the piece in 5 movements. The green circles represent the plants, the pink circles represent the insects, the blue circles represent the birds, and the yellow circles represent the big cats. The red square towards the bottom right-hand side in Scene 1 indicates the listener's current position.

ulation. The composer can also monitor the movements of the listener, who experiences the composition in the form of audio augmented reality without any visual cues.

The simulation data generated in Processing is transmitted to Max via OSC for the musical output to be generated in real-time. The audio engine allocates a unique voice for each agent based on its species. To augment the listener's physical surroundings with the musical ecosystem, we track the listener's position and orientation with a Kinect sensor. The tracking data is transmitted to Max via Synapse [9]. This is then fed into the simulation to situate the listener as an interactive agent within the system.

For the spatialization of the individual voices, we use the ICST Ambisonics Externals for Max [10]. The Cartesian coordinates of the agents generated in Processing are converted to polar coordinates relative to the listener before they are fed into the Ambisonic encoder for the corresponding voices to be localized in space. The Ambisonic scene is then decoded to binaural audio using the IRCAM Spat Tools [11]. The resulting binaural rendering of the scene is broadcast to the listener via headphones.

3. MUSICAL ECOSYSTEM

Musical ecosystems are artificial systems that are devised to generate musical output [12, 13]. When composing such a system, the artist is often concerned with the design of the artificial agents and their behaviors, including how they interact with one another and, in some cases, with the audience. In these systems, the autonomous agent population can be structured in various ways. In *Filterscape*, for instance, this structure is based on the recycling of resources in the food chain [12]. In *Living Melodies*, a genetic programming framework governs the evolution of the musical communications among the agents [3].

In *Proprius*, the trophic structure of a food chain is implemented to create hierarchical dynamics across the species. A trophic structure defines the feeding relationship between organisms in an ecosystem. An organism's place in the food chain, which is based on its feeding behaviors and its prey-predator status, determines its trophic level [14]. Driven by their survival instincts, the artificial agents in *Proprius* navigate a virtual space to seek resources. The real-time sonification of these behaviors form the compositional elements of the piece. Once a simulation is initiated, the listener experiences the work by exploring the physical space onto which the ecosystem is superimposed.

3.1 Musical Agents

Humans possess an evolutionary inclination for identifying the sources of sounds in daily life [15]. In electroacoustic music, the degree to which a listener is capable of, or interested in, associating a sound with a particular real-world source can vary depending on the sonic context and the nature of the sound (e.g., environmental sounds vs. the sound of a sine wave). The composer Denis Smalley refers to the listeners' reflex to create a link between a musical work and the sounding world outside as "source bonding" [16]. In *Proprius*, we exploit this reflex to guide the listener through their experience of the work. To achieve this, we implemented sonification algorithms for each species based on the natural vocalization characteristics of the animal it is modeled after. In sonifications studies, this is described as a *model-based* approach, which is based on "the human ability to associate the characteristics of a perceived sound with the source that may have generated it" [17]. Since the listener's experience of *Proprius* relies solely on audio cues, this approach helps the listener distinguish between different species and navigate through the composition both spatially and temporally.

The selection of the species for each trophic level of *Proprius* was based on the unique behavioral characteristics of the species in order to diversify the compositional structures introduced in each movement. These species are:

- **Plants as Producers**: Plants in *Proprius* are passive organisms that lack mobility. Situated at the bottom of the food chain, the plants function as a source of nutrition for other organisms in the ecosystem. Musically, the plants establish the tonal framework of the piece. For the sonification of the plants, we followed the metaphor of a choir by implementing a variant of formant synthesis with the fundamental frequency of the individual voices tuned to the size of the plants. These tunings are quantized to a common scale to create a harmonic structure across individual plants.

- **Insects as Primary Consumers**: Insects in *Proprius* feed on the producers. The sounds of insects are generated by passing filtered noise through amplitude envelopes that are modeled after insect vocalizations. The center frequency of the peak filter applied to the noise is mapped to the size of the insect. This results in chirping sounds at different frequencies, resembling the stridulations of the insects in nature.

- **Birds as Secondary Consumers**: The size of a bird dictates this omnivorous organism's eating behavior: the small birds eat plants while the big birds act as predators that feed on insects. Birds in nature inherit song templates that consist of a rich range of notes [18]. To emulate the spectral richness of bird song, we implemented an FM synthesizer, where the carrier frequency is mapped to the bird's size and the modulator frequency is mapped to the bird's energy. The amplitude envelopes applied to the resulting sounds are modeled after natural bird vocalizations.
- **Big Cats (i.e. Felidae) as Tertiary Consumers**: In *Proprius*, the Felidae are the apex predators, who feed on the primary and the secondary consumers. Big cat vocalizations exhibit a wide range of frequencies starting from glottal pulses below the audio range to dense roars at higher frequencies [19]. To emulate the breathing and vocalization characteristics of these animals, we used amplitude modulations that range from low-frequency to audio-rate oscillations. An overall amplitude envelope also controls the modulation of the fundamental frequency of the vocalization to create the spectral modulations observed in lion roars [20].

3.2 Common Mappings

Although unique synthesis algorithms are used for each species, some parameter mappings are common across all organisms. These are:

- **Size to Fundamental Frequency**: In nature, the size of an animal's vocal tract and the fundamental frequency of its vocalizations are correlated to its overall size. Therefore, small animals tend to produce higher frequency sounds than larger ones [18]. Accordingly, the size of an organism in *Proprius* is inversely proportional to the fundamental frequency of its vocalizations.
- **Health to Dynamic Range**: In *Proprius*, the health of an organism determines the dynamic range of the sounds it can produce. Although a decline in health due to the presence of a disease agent manifests itself as an overall decrease in the loudness of an animal, some animals can also act quietly when performing certain behaviors, which will be discussed in the following section.
- **Age to Vocalization Rate**: The age of an organism is mapped to the rate at which its vocalizations are triggered. Therefore, all organisms exhibit a gradually decreasing (i.e. rallentando) tempo throughout their lifespan.
- **Energy to Spectral complexity**: Energy transfer is the primary function of the food chain in *Proprius*. When an organism feeds on another organism, it absorbs the prey's energy. The organism then uses this energy to perform other behaviors. The spectral complexity of an organism's sound is manipulated by a mapping between its energy level and the threshold of a normalized wavefolding applied to its sound.

4. COMPOSITIONAL FRAMEWORK

4.1 Behavioral strategies

Fig. 3 shows the compositional structure of *Proprius* in terms of the progression of its scenes and the agent behaviors introduced in each scene. These behaviors determine the structural function of the sounds produced by the agents. Behaviors based on prey-predator interactions (e.g., flee and pursue) have shorter spans whereas other behaviors, such as grazing and wandering, last longer. This creates structural variations between the scenes of the piece: for instance, due to a lack of prey-predator relationships prior to Scene 4, the sounds of the organisms in the first 3 movements display more textural qualities. As a result, the work displays a 2-part form as shown in Fig. 3. The introduction of the hunting animals in Scene 4 introduces gestural sounds that mark the transition from Part A to Part B. The temporal contrasts between the simultaneous textural and gestural elements in Part B establish foreground and background relationships. Whereas a hunting event is perceived more as a figure phrase, the concurrent dormant behaviors establish a textural context for these figures.

Throughout the piece, organisms switch from one behavior to another based on their current drive (e.g., feeding, flight, pursuit). The change in an organism's behavior affects its attributes, which, in return, changes the outcome of its sonification. For instance, when an insect's energy falls below a certain threshold, it can transition from the wandering behavior to grazing. While grazing, it might encounter a predator, which will prompt its behavior to switch to fleeing. If it manages to evade this predator, it might rest or return back to grazing. The emergent combinations of these behaviors will create complex musical phrases that can span from seconds to minutes.

Furthermore, some of the behaviors bring about collective actions across multiple agents. In Fig. 2 Scene 4, the birds represented with blue circles appear to have formed a cluster at the bottom right side of the scene. Here, the birds are performing the flocking behavior, which generates a highly dense sonic structure that travels through space.

A transition between behaviors can also create significant dynamic variations. For instance, during the stalking behavior, the predator reduces its amplitude to approach its prey stealthily. This results in a pianissimo phrase that will rapidly transition to fortissimo as the predator switches to the pursuit mode when it gets close enough to its prey.

Animals in nature adapt their vocalizations to their current situations. For instance, they can use rough and low-frequency vocalizations to sound intimidating, or tone-like high-frequency vocalizations to sound non-threatening [21]. Mimicry is a similar adaptive behavior, where an animal emulates another animal's behaviors to go unnoticed. In *Proprius*, the mimicry behavior, whereby a prey copies the actions of another organism to deceive a predator, creates phrases that are reminiscent of the call and response structures in improvised music. Furthermore, this behavior can push an organism's sonification algorithm to new limits. For example, when a bird performs mimicry, it adopts the fundamental frequency of its predator (i.e. the big cat), which is outside of the bird's regular frequency range. When this frequency is applied to the bird's sonification algorithm, new and unexpected timbres emerge.

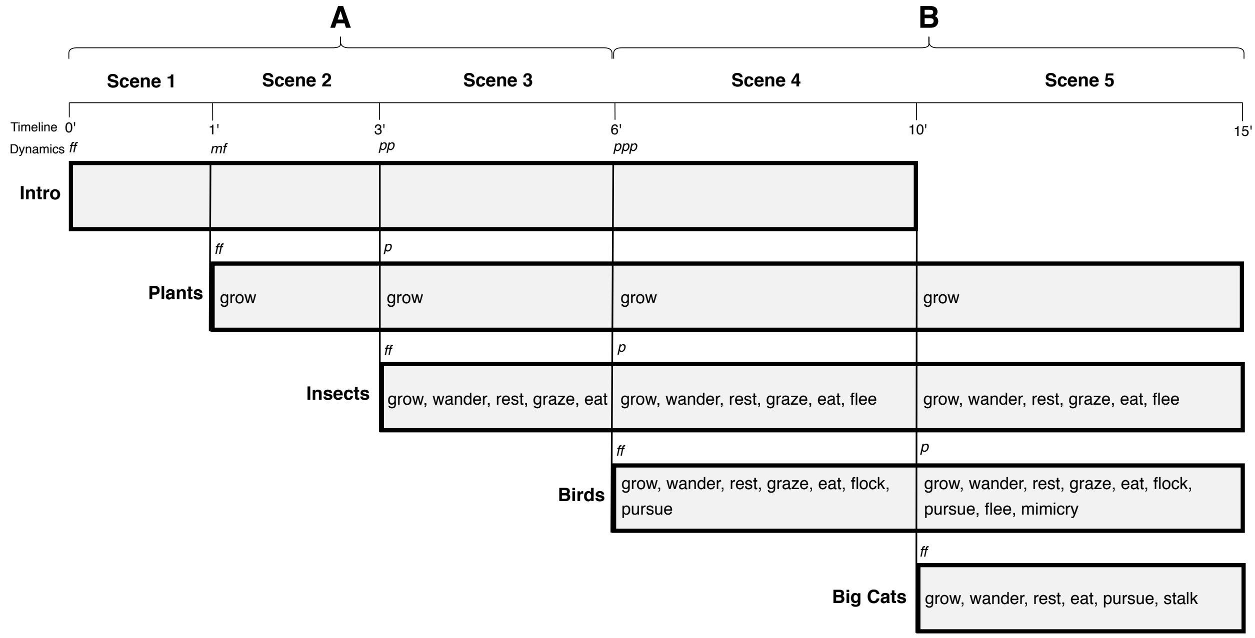

Figure 3. Compositional structure of *Proprius*. In each scene, a new species is introduced into the ecosystem. The behaviors of the individual species are indicated here per scene. Also shown are the approximate durations of each scene, and the overall dynamics for each species.

4.2 Temporal Layering of Scenes

Proprius consists of five musical scenes, four of which represent the trophic levels of the ecosystem as seen in Fig. 3. Instead of transitioning from one movement to another, each new scene is added on top of the existing ones. With each consecutive scene, a new species is introduced into the system, creating a new musical layer in the composition. The timing of when a new species is introduced is the only pre-composed aspect of the work. The gradually increasing durations for consecutive scenes, as shown in Fig. 3, allow the listener to observe the behaviors and the resulting hierarchical relationships introduced into the system with each new species.

While each scene serves as the exposition of a new species, the opening scene, namely Scene 1, functions as an open space devoid of any organisms. In this scene, a continuous ambient noise gives the listener an impression of an open landscape. As the listener navigates this space, their movement sweeps a resonant filter on this noise following a random curve. This allows the listener to explore the navigable space of the composition and get a sense of how the system reacts to their movements without being guided in a particular direction. The noise from the first scene gradually fades out with the introduction of each new species. The noise becomes completely inaudible by the time Scene 5 is introduced.

To articulate the introduction of a new species, the system applies a subtle amount of dynamics compression to the sounds of the existing agents with a compressor side-chained to the sounds of the newly introduced species. This strategy effectively gives the new species a dynamic prominence by slightly decreasing the sounds of the other agents whenever an agent from the new species makes a sound. The threshold of this compressor is gradually increased to remove this effect over time, withdrawing the temporary emphasis given to the new species.

Furthermore, with the introduction of each new species, the distance modeling for the sounds is altered. By increasing the amount of low-pass filtering and reverberation applied, distant organisms are made to be perceived farther away as the piece progresses. This results in a perceived expansion of the augmented audio scene throughout the piece while the navigable space remains the same.

5. INTERACTIVITY

In 1977, the composer Joel Chadabe argued that composing through interactive processes is fundamentally different than traditional methods of composition [22]. The feedback loop between actions and perceptions assume a critical role in such processes [23]. Interactive systems can incorporate the listener into this loop, effectively granting them an artistic license over the progression of the work.

5.1 Composing Interactions

Natural ecosystems have inspired many works of art throughout history. Advances in digital computing in the 20th century have enabled the real-time simulation of such systems, where the audience can influence the system's evolution over time. Today, artificial ecosystems can afford a great variety of user interfaces [24].

The primary mode of interaction in *Proprius* is the navigation of the physical space which the composition is superimposed onto. The listener is introduced into the system as an interactive element both to give them a sense of agency in the composition and to enable musical progressions that are unique to the listener.

The design of a musical ecosystem can be used to compose not only its musical output, but also how the audience interacts with it. The two-part form of *Proprius* seen in Fig. 3 is intended to promote the latter. The textural elements that gradually evolve in the first three scenes of the work encourage the listener to explore the physical space and gain a sense of the system reacts to their presence. The ensuing scenes, which form Part B of the piece, introduce more animated elements with foreground qualities that invite the listener to be more attentive of their surroundings.

Despite such design choices, the listener's personal choices can override the pre-conceived expectations of the composer (i.e. the designer of the system). For instance, in a performance of *Proprius*, one listener indicated that the sounds of the plants were their favorite aspect of the piece. As a result, this listener was very dynamic during Part B in an attempt to avoid predators so as to retain the prominence of the plants in the observed enviroment.

5.2 Listener as a disease agent

Musical ecosystems can employ a variety of methods to integrate the audiences into the progression of a piece. A common technique situates the user as a resource where the organisms tend to flock to the user to gain food or energy [13, 25].

In *Proprius*, we found that the contextualization of the listener as a resource in the system has inhibited the listener's navigation of the space. In this configuration, the listener becomes a point of attraction, which the agents flock to. We observed that this approach created a significant amount of compositional variation around the listener, which effectively mitigated the need for the listener to move in physical space. Since the composition is conceived to emerge from real-time algorithmic and interactive processes, the listener's active embodied involvement is conceived as an integral component of the work.

To achieve this, rather than contextualizing the listener as a point of attraction, we introduce them into the ecosystem as a disease agent. As a result, the health of the artificial agents that are in the vicinity of the listener gradually degrades, prompting the agents to move away from the listener to seek resources that will improve their declining health. This causes the compositional elements to move away from the listener over time, effectively forcing the listener to explore the physical space to search for other musical agents.

Since the plants in *Proprius* are immobile agents, the proximity of the listener to a plant will cause a steady decline in the agent's health, resulting in a gradual fade-out in its auditory output. To further articulate the effect of the listener on the plants, we also mapped their health to their fundamental frequency. Since the plants are the only musical agents in the scene that exhibit harmonic interrelationships, this mapping causes the plants in the vicinity of the listener to be detached from the harmony that they are a part of. A listener has expressed that, upon noticing this deterioration in the harmonic structure of the piece, they tried to maintain their distance to the plants to retain the harmony.

Overall, we observed the use of the listener as a repelling force to create less cacophonous musical outcomes than when we introduced them into the system as a resource. Furthermore, we also noticed an increase in the listeners' tendency to explore the augmented space. Some listeners likened their experience to playing a game. One listener described that they tried to pursue the big cats to curb their population so that they wouldn't prey on the birds. Another listener mentioned having tried to move with the birds, but upon noticing that the birds were actually trying to move away from them, they decided not to disturb them. Such instances can be indicative of the listeners' ability to empathize with the artificial agents. We find this type of connection between the listener and the parts of a musical composition to be an intriguing affordance of interactive musical ecosystems, and we hope to further explore this aspect of *Proprius* in its future iterations.

6. CONCLUSION

In this paper, we described our ongoing research into interactive agent-based musical ecosystems. More specifically, we discussed the compositional strategies utilized in the design of *Proprius*. The creative decisions made in the conception of this work were predominantly inspired by ecological models. Other compositional choices were made to articulate the form of the piece and to highlight the interactions between the listener and the artificial agents. Overall, the organisms in *Proprius* are conceived as parts of a musical ensemble that take up different parts of the frequency spectrum, occupy different sections in the temporal progression of the piece, and eventually serve different musical functions.

Moving forward, we plan to add new species to our ecosystem to expand its musical capabilities. This will not only diversify the timbral qualities of its output, but also enable new emergent interactions that will lead to new musical phrasings. Furthermore, by integrating *Inviso*, our cross-platform authoring tool for immersive sonic environments [26], into our current system, we intend to implement agents that emit sounds with limited directionality, increasing the spatial complexity of the musical output.

We also plan to explore the incorporation of visual elements in our ecosystem. When asked about whether they would have liked *Proprius* to be accompanied by visuals, most listeners have indicated that they would prefer it to remain an exclusively auditory experience. However, with recent advances in virtual and augmented reality technologies, we expect new possibilities in this domain to emerge. Particularly, with the upcoming wave of VR systems that rely on inside out tracking (i.e. those that do not require an external tracking system), it could be possible to superimpose *Proprius* onto much larger spaces. Furthermore, using such systems, visual interfaces and hand-held controllers can be utilized to implement rich user interactions that can have more nuanced or more exaggerated effects on the compositional outcome, offering new immersive musical experiences.

7. REFERENCES

[1] T. Blackwell, "Swarm music: Improvised music with multi-swarms," in *AISB Symposium on Artificial Intelligence and Creativity in Arts and Science*, 2003.

[2] J. McCormack, "Aesthetic evolution of L-systems revisited," in *Workshops on Applications of Evolutionary Computation*. Springer, 2004, pp. 477–488.

[3] P. Dahlstedt and M. G. Nordahl, "Living melodies: Coevolution of sonic communication," *Leonardo*, vol. 34, no. 3, pp. 243–248, 2001.

[4] O. Bown, "Ecosystem models for real-time generative music: A methodology and framework," in *Proceedings of the 35th International Computer Music Conference*, 2009.

[5] P. Maes, "Artificial life meets entertainment: Lifelike autonomous agents," *Communications of the ACM*, vol. 38, no. 11, pp. 108–114, 1995.

[6] R. F. Antunes, F. F. Leymarie, and W. Latham, "Two decades of evolutionary art using computational ecosystems and its potential for virtual worlds," *Journal For Virtual Worlds Research*, vol. 7, no. 3, 2014.

[7] A. Çamcı, Z. Özcan, and D. Pehlevan, "Interactive virtual soundscapes: A research report," in *Proceedings of the 41st International Computer Music Conference (ICMC)*, 2015, pp. 163–169.

[8] Z. Özcan and A. Çamcı, "An Augmented Reality Music Composition Based on the Sonification of Animal Behavior," in *Proceedings of the Audio Engineering Society Conference on Audio for Virtual and Augmented Reality*, 2018.

[9] J. Bellona, "Kinect-Via-: Max/MSP Performance Interface Series for Kinects User Tracking via OSC," *University of Oregon*, 2012.

[10] J. C. Schacher and P. Kocher, "Ambisonics spatialization tools for Max/MSP," in *Proceedings of the 32nd International Computer Music Conference*, 2006, pp. 274–277.

[11] T. Carpentier, M. Noisternig, and O. Warusfel, "Twenty years of Ircam Spat: Looking back, looking forward," in *Proceedings of the 41st International Computer Music Conference*, 2015, pp. 270–277.

[12] A. Eldridge and A. Dorin, "Filterscape: Energy recycling in a creative ecosystem," in *Proceedings of the Workshops on Applications of Evolutionary Computation*. Springer, 2009, pp. 508–517.

[13] J. McCormack, "Eden: An evolutionary sonic ecosystem," in *Proceedings of the European Conference on Artificial Life*. Springer, 2001, pp. 133–142.

[14] M. Allaby, *A dictionary of ecology*. Oxford University Press, 2010.

[15] J. A. Ballas, "Common factors in the identification of an assortment of brief everyday sounds." *Journal of experimental psychology: human perception and performance*, vol. 19, no. 2, p. 250, 1993.

[16] D. Smalley, "Spectromorphology: explaining sound-shapes," *Organised sound*, vol. 2, no. 2, pp. 107–126, 1997.

[17] 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, 2013.

[18] D. McFarland, *Animal behaviour: psychobiology, ethology, and evolution*. John Wiley & Sons, 1993.

[19] A. Farnell, *Designing sound*. MIT Press, 2010.

[20] G. Ananthakrishnan, R. Eklund, G. Peters, and E. Mabiza, "An acoustic analysis of lion roars. II: Vocal tract characteristics," in *Proceedings of the Fonetik 2011*, 2011, pp. 8–10.

[21] J. J. Ohala, "Cross-language use of pitch: an ethological view," *Phonetica*, vol. 40, no. 1, pp. 1–18, 1983.

[22] J. Chadabe, "Some reflections on the nature of the landscape within which computer music systems are designed," *Computer Music Journal*, pp. 5–11, 1977.

[23] J. C. Schacher, "Action and perception in interactive sound installations: An ecological approach." in *Proceedings of the 9th New Interfaces for Musical Expression*, 2009, pp. 286–289.

[24] H. Ji and G. Wakefield, "Endogenous biologically inspired art of complex systems," *IEEE computer graphics and applications*, vol. 36, no. 1, pp. 16–21, 2016.

[25] R. Berry, W. Rungsarityotin, A. Dorin, P. Dahlstedt, and C. Haw, "Unfinished Symphonies-songs of 3 1/2 worlds," in *Proceedings of the Workshop on Artificial Life Models for Musical Applications*, 2001, pp. 51–64.

[26] A. Çamcı, K. Lee, C. J. Roberts, and A. G. Forbes, "INVISO: A Cross-platform User Interface for Creating Virtual Sonic Environments," in *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. ACM, 2017, pp. 507–518.