Audio Engineering Society

Conference Paper

Presented at the Conference on
Audio for Virtual and Augmented Reality
2018 August 20 – 22, Redmond, WA, USA

An Augmented Reality Music Composition Based on the Sonification of Animal Behavior

Zeynep Özcan[1] and Anıl Çamcı[2]

[1] *Istanbul Technical University*
[2] *University of Michigan*

Correspondence should be addressed to Zeynep Özcan (ozcanz@icloud.com)

ABSTRACT

In this paper, we discuss the immersive sonification of an artificial ecosystem in the form of an interactive augmented reality music composition, named *Proprius*. We first offer an overview of existing work that utilizes ecological models in compositional and sonification contexts. We then describe the behavioral and ethological models utilized in *Proprius*. We evaluate the musical characteristics of animal behaviors, and discuss our approach to sonifying them in the context of an interactive augmented reality composition. We provide details of our system in terms of how it implements ecological simulation, immersive audio, and embodied interaction.

1 Introduction

Biological systems has been used by many artists and researchers in creative applications that rely on the modeling of animal behavior in virtual environments. These applications, also referred to as artificial or computational ecosystems, take the form take the form audio [1], visual [2], and multimodal [3] art installations. Furthermore, with advances in digital computing, it became possible to simulate complex ecological systems in real-time. This has enabled the implementation of interactive works, where the user input affects the evolution of a system.

Proprius is an autonomous interactive sound environment, where the sonification of animal behavior within an ecological simulation is used as a means to create an interactive augmented reality music composition. It is part of our ongoing research on interactive virtual sonic environemnts [4, 5]. The artificial ecosystem designed in Processing simulates the animal behaviors. The simulation data are then fed into Max, which renders the immersive audio scene in real time. The listener's position in the scene is determined with a Kinect sensor; as the listener explores the exhibition space, a binaural audio scene augments their physical environment.

Behavioral ecology provides an evolutionary and ecological framework of animal behavior by studying the ways in which animals adapt their behavior to maximize chances of survival and reproduction [6]. In an artificial ecosystem, autonomous animal-like agents can exhibit a variety of behaviors: they can interact with other agents, observe their environment, and make decisions according to rules of varying complexity. The organization of agents within an ecosystem are modeled within ecological pyramids, which constitute a framework for the energy flow and nutrition cycles. An organism's place in the food chain, which is based on their feeding behavior, determines their trophic level. We explore the systematic relationship between behavioral ecology and data sonification as a means to create complex compositional structures that rely on the interactions between the agents of our ecological model. This approach allows us to utilize animal behaviors and how these behaviors permeate through a food chain to establish compositional structures at micro and macro levels.

In this paper, we first offer an overview of existing studies on the use of ecological models in music composition and sound art. We then describe the animal behaviors modeled in *Proprius*, and compare the musical qualities that these behaviors exhibit. We delineate our approach to sonifying these behaviors, and structuring the resulting sonifications in the context of an augmented reality composition. Furthermore, we offer detailed descriptions of the tools and techniques utilized in the implementation of our system.

2 Related Work

Numerous artists and researchers have explored the use of ecological models in algorithmic art and music composition. In one of the earliest examples of an interactive artificial ecosystem, Christa Sommerer and Laurent Mignonneau's *A-Volve* allows visitors to design virtual creatures, and insert these into the system in ways that interfere with the relationship between preys and predators [7, 2]. The visitors interact with the creatures projected into a glass pool by inserting their hands into the water. A similar tactile interaction in the context of an artificial ecosystem is used in *Archipelago* where the participants manipulate artificial life by touching and moulding a mini landscape built out of sand which the simulation is projected onto [8].

The majority of the artistic works that involve ecological simulations utilize either the visual or the audiovisual domain [9]. There are, however, works that rely solely on the auditory domain as well. For instance in *Living Melodies*, the autonomous agents of a sonic ecosystem sing to attract other agents. The chorus of the mating calls between these agents result in a musical composition, where only the sonic structures which the agents find to be musically pleasing according to their genetic code become audible by the listener[1].

In the artist Jon McCormack's audiovisual installation *Eden*, the user's position determines where in the artificial ecosystem resources will be generated. As a result, instead of singing to each other as in *Living Melodies*, the agents of this ecosystem sing to the audience to attract their attention so as to increase the resources available to them. Over time, the singing pattern of an agent evolves according to audience preferences in order to increase the agent's chances of survival [10].

Recent projects that involve artificial ecosystems make use of emerging immersive technologies. In their *Artificial Nature* series, the artists Haru Ji and Graham Wakefield utilize modern AR and VR systems to create audiovisual nature-based aesthetic experiences in the form of interactive artworks [11]. In different iterations of this series, the artists experiment with various interaction models and hardware (e.g., head-mounted displays, Kinect sensors). In one such iteration named *Time of Doubles*, the members of the audience are presented as sources of energy, and projected as "doubles" to the visual output of the virtual environment [12]. Similarly in *Proprius*, we utilize modern depth tracking systems to determine a listener's position and orientation both to render the binaural sound field according to the listener's navigation of the exhibition space, and to introduce the listener into the ecosystem as an external agent.

As seen in *Eden* and *Time of Doubles*, artificial ecosystems that enable user interaction often situate the users as a resource in the form of food or energy for the agents of the ecosystem. In Alan Dorin's audiovisual installation *Pandemic*, however, the audience assumes the role of a disease agent: the color of a visitor's clothing acts as a transmissible disease that infects the artificial agents of the ecosystem. According to the artist, the process of disease transmission becomes an aesthetic object [13]. In *Proprius*, the listeners similarly acts as a disease-like external factor that lowers the health of nearby organisms.

3 Proprius

Proprius is an interactive augmented reality composition based on the sonification of organisms in an artificial ecosystem. The consecutive layers of a food chain are used as movements that dictate the temporal progression of the music. The listener explores a physical space augmented with *Proprius* to experience the compositional unfolding of the ecosystem. The listener's presence in the system affects nearby organisms; the listener is thereby situated as an interactive agent that contributes to the progression of the work.

3.1 Modeling of the Ecosystem

In real-life ecosystems, organisms depend on energy to survive. Based on how these organisms obtain energy, they can be categorized under two main groups: producers (i.e. autotrophs) and consumers (i.e. heterotrophs) [14]. Producers obtain energy from the sun while consumers eat other organisms to gain energy. The organization of such agents within an ecosystem are often modeled within ecological pyramids, which delineate a framework for the energy flow and nutrition cycles within the system. An ecological pyramid of energy can be seen in Fig. 1.

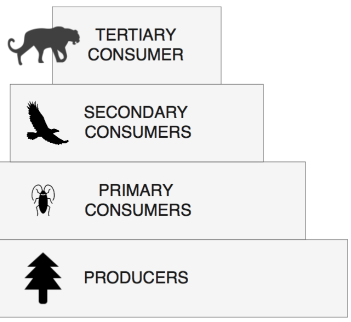

Fig. 1: An ecological pyramid. The energy flows from the bottom level (i.e. producers) to the higher level.

A *trophic structure* defines the feeding relationship between organisms in an ecosystem. The trophic level of an organism indicates its place in the food chain [15]. In *Proprius*, organisms are modeled according to these trophic levels. The 1st trophic level of a food chain contains *primary producers* (i.e. green plants). At the 2nd trophic level, *primary consumers* (i.e. herbivores) gain energy by feeding on *primary producers*. At the 3rd trophic level, *secondary consumers* (i.e. omnivores) feed on organisms at the 1st and 2nd trophic levels. Organisms at the 3rd level can be either preys or predators depending on their feeding behavior. *Tertiary consumers* (i.e. carnivores) at the 4th trophic level are meat eaters. These are natural predators that feed on *primary* and *secondary consumers* [15].

To model the trophic levels of the ecosystem in *Proprius*, we selected one species from each level. These selections were made based on the unique behavioral characteristics of each organism type, which facilitated the diversification of the compositional structures introduced at each trophic level. The selected groups are plants as primary producers, insects as primary consumers, birds as secondary consumers, and big cats as tertiary consumers.

3.2 Modeling of Animal Behavior

Animals in an ecosystem constantly make decisions about when to eat, when to rest, when to reproduce, and when to flee from predators. They decide upon such behaviors according to stimuli that are either internal (e.g., hunger) or external (e.g., danger). Autonomous animal-like agents can exhibit a variety of behaviors in artificial ecosystems: they can interact with other agents, observe their environment, and make decisions according to rules of varying complexity.

In *Proprius*, animals choose their behaviors according to their current drives (e.g., feeding, flight), with the underlying goals of survival and reproduction. If an animal is not hungry, it will wander. However, when it is hungry, it will prefer feeding-driven activities (e.g., graze and pursue). Although it will try to find resources that will maximize the energy gain, if it happens to be chased by a predator, it will prioritize survival, and flee. The behaviors and attributes of organisms at each trophic level of *Proprius* is shown in Table 1.

Proprius is designed with Processing, a java-based multimedia programming platform. [1] The ecosystem is implemented using within the object-oriented design paradigm: the species at a trophic level is implemented as a class that describes the behaviors and attributes of an organism. Arrays of organisms are then instantiated from the classes. Behaviors that simulate animal movement (i.e. flee, wander, pursue, and flock) are designed after Daniel Shiffman's implementation of Reynold's model of steering behavior [16].

[1] https://processing.org

Table 1: Attributes and behaviors of the organisms in each trophic level in *Proprius*.

	Attributes	Behaviors
Producer	health, energy, age, size, position	grow
Primary Consumer	health, energy, age, size, position, sex, vision	flee, wander, rest, graze, eat, reproduce, dodge, mimicry
Secondary Consumer	health, energy, age, size, position, sex, vision	flee, wander, rest, graze, eat, reproduce, pursue, flock
Tertiary Consumer	health, energy, age, size, position, sex, vision	wander, rest, eat, reproduce, pursue, stalk

3.3 Compositional Structure

Proprius consists of five musical scenes, four of which represents the trophic levels of the ecosystem seen in Fig. 1. In each scene, a new layer of sound is introduced into the composition. As the composition progresses, new layers are superimposed onto the navigable area of the system as seen in Fig. 2.

The opening scene, namely Scene 1, is an open space devoid of life. This scene allows the listener to explore the physical space, and get a sense of how the system reacts to their movements. This layer opens with an ambient noise. The listener filters this noise as they walk through the exhibition space. The ambient noise from Scene 1 gradually fades out throughout the composition as new trophic levels are introduced. Scene 2 introduces of producers (i.e. plants).

Primary consumers are introduced into the system in the form of insects in Scene 3. In this scene, interactions between agents begin to emerge. Secondary consumers are introduced into the system as birds in Scene 4, where prey/predator relationships begin to occur between consumers. This introduces new behaviors such as flight and pursuit into the system. Tertiary consumers are introduced into the system as big cats in Scene 5. These are the apex predators, which are not hunted by the other, and therefore don't perform the flight behavior. With the addition of the tertiary consumers, the artificial ecosystem starts to function as a whole.

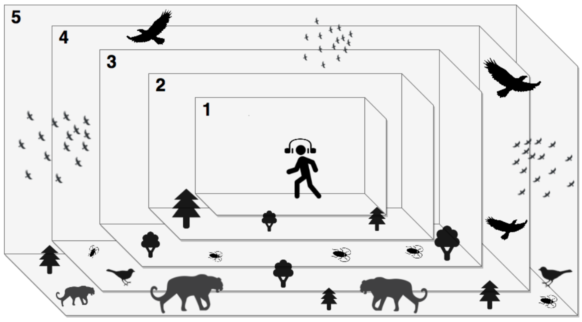

Fig. 2: A visual representation of the trophic levels (i.e. musical scenes) in *Proprius*.

4 Sonification of Organisms in Proprius

Data sonification is utilized in a variety of fields ranging from science to sound art. According to Thomas Herman, data sonification must be systematic and reproducible [17]. A systematic relationship between behavioral ecology and data sonification is capable of yielding complex compositional structures as a result of the interactions between the agents of an ecological model. In *Proprius*, we adopt a *model-based* approach to sonification that is "grounded in the human ability to associate a perceived sound and its characteristics with the source that generated it" [18]. Informed by the sonic characteristics of the organisms that are being sonified, we use attributes and behaviors of individual agents, seen in Table 1, as parameters to synthesize sounds in real-time. Our general approach to sonification in terms of the external and internal factors that affect the sound of an organism can be seen in Fig. 3.

The sonification in *Proprius* is implemented with Max, a visual language for multimedia programming. [2] An object-oriented approach similar to that of the simulation side of the project is adopted for the sonification engine. Based on the behavioral characteristics and the natural voices of each organism, we selected core synthesis methods for each organism type. This approach both facilitated the articulation of individual species, and diversified the timbral qualities of the composition.

4.1 Attribute Related

To sonify the attributes of organisms, we created parameter mappings that are common across all species.

[2]https://cycling74.com

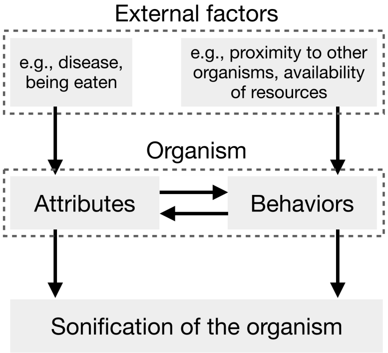

Fig. 3: The sonification model in *Proprius*.

The health of an organism is mapped to the amplitude of its sonification. The energy parameter is mapped to spectral complexity, which is inversely proportional to the amplitude threshold of the wavefolding applied to a sound. As a result, energy translates into an overall spectral richness. The size of an object determines the fundamental frequency of the sonification. This mapping is motivated by the inverse proportionality between an animal's size and the pitch range of its vocalizations. Finally, the age of an organism is mapped to the rate at which its vocalizations are triggered (i.e. older an organism, slower the re-triggering rate). Besides these mappings, each species are allocated a unique synthesis method and amplitude envelopes in accordance with the model-based approach of our sonification.

4.2 Behavior Related

Behaviors determine the extent of sound events. Behaviors based on prey/predator interactions (e.g., flee and pursue) have shorter duration whereas other behaviors such as grazing and wandering extend for longer periods. This creates structural variations across the scenes of the piece: for instance, due to a lack of prey/predator relationships until Scene 4, the sounds thus far display more textural qualities. The introduction of hunting animals in Scene 4 brings about more gestural sounds.

To avoid a cacophony of simultaneous sounds events, behaviors are used to articulate foreground and background relationships. For instance, when an animal is executing a hunting-related behavior, a side-chain compressor subtly subdues other concurrent sound events to direct the listener's attention to that behavior.

4.3 Real-time synthesis methods

The ambient noise in In Scene 1 is passed through a resonant peak filter. As the listener moves through the exhibition space, their position data moves the center frequency of this filter on a smooth random curve. This is intended to provide the listener with a sense of how the system reacts to their movements, and where the boundaries of the navigable region of the piece are.

For the plants In Scene 2, the metaphor of a chorus is used based on the stationary nature of these organisms. To achieve this, we used a variant of formant synthesis with the fundamental frequency of the individual synthesizers tuned to the size of the corresponding plant. The fundamental frequencies are quantized to a common scale to create harmonic structures across all plants.

The sounds of the insects in Scene 3 are generated with filtered noise passed through amplitude envelopes that are modeled after natural insect sounds. The center frequency of peak filter on the noise is mapped to the size of the insect. The bird sounds in Scene 4 are generated with FM synthesis, the carrier frequency of which is mapped to the size of the bird. The modulator frequency is dependent on the bird's amplitude envelope which are modeled after natural bird sounds.

The sound of the big cats in Scene 4 are generated with a filtered noise and a sine wave oscillator are added together. The frequency of the oscillator is mapped to the size of the big cat. The amplitudes of noise and oscillator are modulated with a low frequency oscillator. The frequency of a sound moves up and down to create a roaring effect.

5 Augmented Reality Audio Implementation

Unlike a traditional music composition, *Proprius* is experienced by navigating a physical space, where the spatial animation of the artificial agents serves as a means of immersion. The sonification of the simulated ecosystem is generated in the form of a binaurally rendered Ambisonic sound field that augments the immediate reality of the listener. We track the position of the listener in the exhibition space to spatialize the sounds of the organisms relative to the listener. Furthermore, we use this data to introduce the listener into the ecosystem as an interactive agent. As a result, a

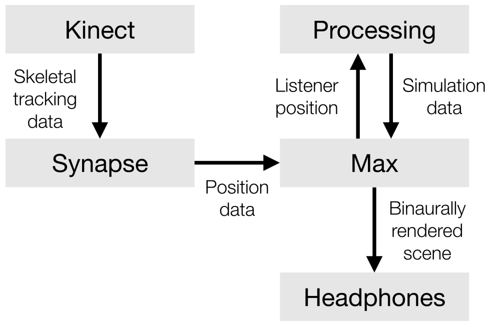

Fig. 4: Data flow across the various hardware and software components of *Proprius*.

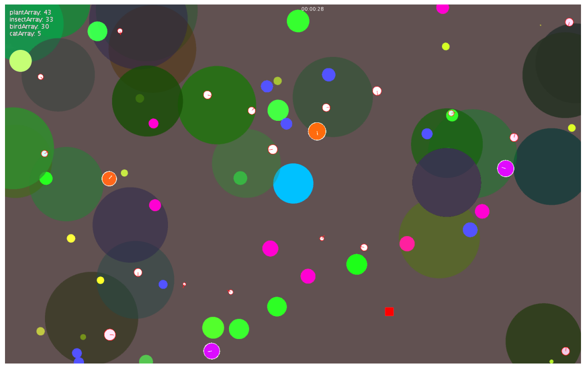

Fig. 5: Processing screen that provides the composer with a real-time visual representation of the simulation. Visual objects of different colors and shapes represent the various agents in the ecosystem. The red square towards the bottom right-hand side of the screen indicates the listener's position.

feedback loop between the user and the ecosystem is formed.

For listener tracking, we use the Microsoft Kinect sensor, which can perform a motion-capture technique called *skeletal tracking*. This technique estimates the 3D positions of skeletal joints when a human form is recognized in depth data. We pass the skeletal tracking data to Max via *Synapse*, which outputs Open Sound Control (OSC) messages [19]. This data is then passed to Processing via *OscP5*. The simulation data is passed back to Max, where the synthesis and spatialization of sounds occur. See Fig. 4 for a diagram explaining the data flow across various hardware and software components of *Proprius*.

The visual system seen in Fig. 5 allows the composer to monitor the location of the listener, and the progression of the work in terms of the behaviors of the organisms that make up the ecosystem. The listener, on the other hand, experiences *Proprius* system aurally without such visual cues.

5.1 Spatialization of Organisms

To spatialize the organisms in *Proprius*, we use Ambisonics, which facilitates the design and manipulation of complex sound fields, such as those generated by our system. To integrate Ambisonic processing into *Proprius*, we used the *Ambisonics Externals for MaxMSP* developed by Jan Schacher and Philippe Kocher at the Institute for Computer Music and Sound Technology (ICST) in Zurich [20].

The ICST Ambisonics Tools can render an arbitrary number of moving sound sources around a stationary listener. The sounds and their position data are encoded using the *ambiencode~* object. A corresponding decoder, which stores information about the speaker configuration, outputs an Ambisonic signal at the desired order.

Since the augmented reality approach in *Proprius* requires the listener to move in physical space, the Ambisonic encoding based around a stationary listener needs to be updated relative to the momentary listener position. To achieve this, the position data pertaining to the individual organisms in the simulation is normalized to the listener position before they are fed into the encoder, as seen in Fig. 6.

Since an augmented reality is dependent to the perspective of the user, the implementation of an audio augmented reality often necessitates the use of headphones to render a version of the sound field that is exclusive the user. Binaural navigation of a higher-order Ambisonic sound field allows the user to explore a spatially accurate rendition of an acoustic space [21]. To adapt the Ambisonic output of our system to binaural audio, we use a virtual speaker approach using IRCAM's SPAT tools. The *Spat.virtualspeakers~* down-mixes the multi-channel Ambisonic stream into a binaural stereo track "while preserving the spatial image of the original sound scene" [22]. The resulting binaural signal is output to the headphones equipped by the listener, who is navigating the physical space.

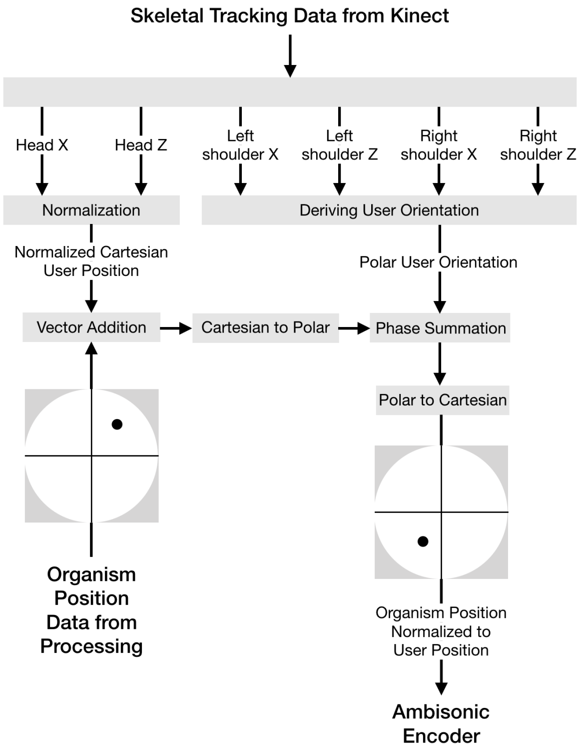

Fig. 6: Processing of position data in *Proprius*. Skeletal data obtained from Kinect via Synapse is combined with the simulation data from Processing to determine the position of each organism relative to the listener. This position is then fed into the Ambisonic encoder to situate the object in the sound field.

Head-tracking is deemed essential to achieve plausible binaural experiences. Particularly in an augmented reality audio system, the user's orientation is a fundamental aspect of the immersive audio processing. However, head-tracking requires specialized hardware that can complicate the hardware configuration of a system. To facilitate the implementation of our system in various exhibition situations, we use the skeletal tracking data from the Kinect to estimate the listener's orientation without using a head-tracking system. As seen in Fig. 7, we use the left and right shoulder joint data to derive an orientation vector for the listener. While this approach does not account for situations where the listener bends their neck, it offers an approximation for front-back spatialization in a flexible hardware configuration.

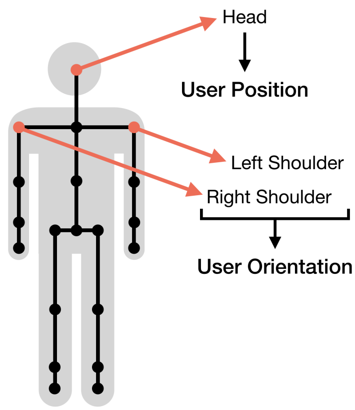

Fig. 7: Use of the skeletal tracking data obtained from Kinect via Synapse. The head position is used to determine the listener's location. Left and right shoulders are cross-referenced to derive an orientation vector for the listener.

6 Future Work and Conclusion

With recent advances in immersive media technologies, such as head-mounted displays and position tracking systems, we will explore our options to substitute the now-obsolete Kinect sensor with a modern head-tracking technology, which will potentially improve the binaural rendering of our sound field. We are currently experimenting with the use of the Vive headset with promising results although the occlusion of the listener's view is an undesirable side effect withing the augmented reality framework of *Proprius*.

We plan to implement new natural interaction techniques using gesture recognition to allow the user to interact with the ecosystem in a more detailed manner. This will also require the integration of new behaviors for the animals to react to these interactions, which will enrich emergent qualities of our system.

In this paper, we described *Proprius*, an autonomous interactive sound environment, where the sonification of animal behavior within an ecological simulation is used to create an interactive augmented reality music composition. With *Proprius*, we explore the artistic potential of biologically inspired models for electronic music composition.

We expect that with recent advances in immersive media technologies, augmented reality compositions such as *Proprius* will gain further prominence. We believe that our work offers a strong contribution as a case study that blends music composition with ecological simulation, human-computer interaction, and augmented reality audio.

References

[1] Dahlstedt, P. and Nordahl, M. G., "Living melodies: Coevolution of sonic communication," *Leonardo*, 34(3), pp. 243–248, 2001.

[2] Sommerer, C. and Mignonneau, L., "A-Volve-an evolutionary artificial life environment," *Artificial Life VC Langton and C. Shimohara Eds., MIT*, pp. 167–175, 1997.

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

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

[5] Çamcı, A., Lee, K., Roberts, C. J., and Forbes, A. G., "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*, pp. 507–518, ACM, 2017.

[6] Krebs, J. R. and Davies, N. B., *Behavioural ecology: an evolutionary approach*, John Wiley & Sons, 2009.

[7] Sommerer, C. and Mignonneau, L., "Art as a living system: interactive computer artworks," *Leonardo*, 32(3), pp. 165–173, 1999.

[8] Ji, H. and Wakefield, G., "Endogenous Biologically Inspired Art of Complex Systems," *IEEE computer graphics and applications*, 36(1), pp. 16–21, 2016.

[9] Antunes, R. F., Leymarie, F. F., and Latham, W., "Two decades of evolutionary art using computational ecosystems and its potential for virtual worlds," *Journal For Virtual Worlds Research*, 7(3), 2014.

[10] McCormack, J., "Evolving for the Audience," *International Journal of Design Computing*, 4, p. 14, 2002.

[11] Ji, H. H., *Artificial natures: Creating nature-like aesthetic experiences through immersive artificial life worlds*, University of California, Santa Barbara, 2012.

[12] Ji, H. and Wakefield, G., "Virtual world-making in an interactive art installation: Time of doubles," *Virtual Worlds: Artificial Ecosystems and Digital Art Exploration. Science ebooks*, 2012.

[13] Dorin, A., "Pandemic," `http://users.monash.edu/~aland/pandemic.html`, 2012.

[14] Barrows, E. M., *Animal behavior desk reference: a dictionary of animal behavior, ecology, and evolution*, CRC press, 2011.

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

[16] Shiffman, D., *The Nature of Code: Simulating Natural Systems with Processing*, Daniel Shiffman, 2012.

[17] Hermann, T., "Taxonomy and definitions for sonification and auditory display," in *Proceedings of the 14th International Conference on Auditory Display (ICAD 2008)*, 2008.

[18] Dubus, G. and Bresin, R., "A systematic review of mapping strategies for the sonification of physical quantities," *PloS one*, 8(12), p. e82491, 2013.

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

[20] Schacher, J. C. and Kocher, P., "Ambisonics spatialization tools for max/msp," *Omni*, 500(1), 2006.

[21] Tylka, J. G. and Choueiri, E., "Comparison of techniques for binaural navigation of higher-order ambisonic soundfields," in *Audio Engineering Society Convention 139*, Audio Engineering Society, 2015.

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