AI & SOCIETY (2020) 35:87–101
https://doi.org/10.1007/s00146-018-0862-4

ORIGINAL ARTICLE

Bird Song Diamond in Deep Space 8k

John Brumley[1] · Charles Taylor[2] · Reiji Suzuki[3] · Takashi Ikegami[4] · Victoria Vesna[5] · Hiroo Iwata[6]

Received: 15 February 2018 / Accepted: 24 August 2018 / Published online: 3 September 2018
© Springer-Verlag London Ltd., part of Springer Nature 2018

Abstract
The Bird Song Diamond (BSD) project is a series of multifaceted and multidisciplinary installations with the aim of bringing contemporary research on bird communication to a large public audience. Using art and technology to create immersive experiences, BSD allows large audiences to embody bird communication rather than passively observe. In particular, BSD Mimic, a system for mimicking bird song, asks participants to grapple with both audition and vocalization of birdsong. The use of interactive installations for public outreach provides unique experiences to a diverse audience, while providing direct feedback for artists and researchers interested in the success of such outreach. By following an iterative design process, both artists and researchers have been able to evaluate the effectiveness of each installation for promoting audience engagement with the subject matter. The execution and evaluation of each iteration of BSD is described throughout the paper. In addition, the process of interdisciplinary collaboration in our project has led to a more defined role of the artist as a facilitator of specialists. BSD Mimic has also led to further questions about the nature of audience collaboration for an engaged experience.

Keywords Birdsong · Collaboration · Iterative design · Language · Interspecies communication · Site/habitat specificity

1 Introduction

Bird communication extends beyond the handful of functional categories that we generally attribute to them: mating, territory, and warnings (Dooling 1982). Recent techniques in collection and analysis have allowed us to dissect bird songs to reveal grammar and syntactic structures that were previously unobservable (Hedley 2016). Bird Song Diamond (BSD) is an ongoing project that seeks to disseminate this scientific research on birdsong (Arriaga et al. 2014), as well as emerging techniques in collection using sensory networks (Yu et al. 2016) and analysis based on artificial life simulations (Ikegami et al. 2017). The project has developed over the past 5 years as a public outreach extension of a larger research initiative entitled "Acoustic Sensor Arrays for Understanding Bird Communication." Our solution to reaching large audiences while providing active involvement with the subject matter has been the development of immersive and interactive art installations. Employing insights from the above research and integrating concerns regarding the collision and fusion of anthropogenic activity with biological and

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00146-018-0862-4) contains supplementary material, which is available to authorized users.

✉ John Brumley
s1630526@u.tsukuba.ac.jp

Charles Taylor
taylor@biology.ucla.edu

Reiji Suzuki
reiji@is.nagoya-u.ac.jp

Takashi Ikegami
ikeg@sacral.c.u-tokyo.ac.jp

Victoria Vesna
vv@ucla.edu

Hiroo Iwata
iwata@kz.tsukuba.ac.jp

[1] Program in Empowerment Informatics, EMP Office, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan

[2] Professor Emeritus, Department of Ecology and Evolutionary Biology, UCLA, Los Angeles, USA

[3] Graduate School of Informatics, Nagoya University, Nagoya, Japan

[4] The Graduate School of Arts and Sciences, University of Tokyo, Tokyo, Japan

[5] Department of Design Media Arts, UCLA, Los Angeles, USA

[6] Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba, Japan

geological systems (Chertow 2008; Crist 2013; Harraway 2016; Malm 2018; Stengers and Goffey 2015; Wainwright and Mann 2018; Wark 2016), BSD allows participants to approach new understandings of bird communication and perception in a more holistic and contemporary sense by actively participating in these systems from the standpoint of a bird. Promoting public awareness of scientific research in this participatory and immersive setting allows intuitive understanding of abstract ideas, and continues in the tradition of aesthetically informed scientific outreach (Hein 1990; Tuchman 1971). Through multisensory stimuli and participation, participants' awareness is focused to contain the audio and visual elements which are of primary interest to birds (Graham 2011). Therefore, both audio and visual contents are selected based on proximity of concern for a bird to limit, or channel, the participant's viewpoint to that of a bird's perception.

The past has seen numerous depictions of birds and bird song in art. In music (Head 1997; Kraft 2013), in painting (Bugler 2012), sculpture (Brancusi 1928; Calder 1971; Kelley 1978; Sosolimited 2016) and installation (Simon 2014; Boursier-Mougenot 1999), birds have generally been represented in artwork as separate entities from the viewer. More recently, interactive and immersive artwork which engages the viewer as a participant within the work has dissolved the distance between the depiction of a bird and the becoming a bird (Kac et al. 1996; Legrady et al. 2013; Milk et al. 2012). The BSD project builds on methods used in interactive and immersive artworks to further close the experiential distance between participant and subject.

The difficulty in subjectively understanding the qualities of bird perception is managed through human participation within environments modulated to the concerns of birds (Gibson 1986; Nagel 1974). The necessary technology involved depends on the sensation that BSD is presenting to the participant. However, for participants to experience the complex dynamics focused on by the associated research, comparably complex methods for audio and visual production and immersion must be used.

Mirroring artificial life techniques of simulated natural systems for exploring the grammar of birdsong, the creation of complex multi-modal immersive installations brings these concepts to the public by having audience members actively participate as components of an emergent system (Ehnes 2010). Involvement within the installation, while surrounded by bird-centric stimuli, enables an embodied understanding of both the larger social processes of birds and individual bird perception (Massumi 2002; Thrift 2007).

To create such multisensory-media-saturated environments, the technical systems used to drive the experience have had to evolve with each iteration, adapting to the unique physical demands and cultural context of a space. For example, creating immersion in an outdoor space like Times Square (Sect. 3.1.1) requires strategies and technologies that augment and integrate the surrounding environment into the experience, while controlled, indoor settings (Sects. 3.1.2; 3.2.2) have more freedom but also demand more effort to construct the entire experience. Striving to recreate bird-like perception involves the development of not only software simulations of bird behavior, but of hardware which can effectively reproduce these perceptions, often physically replicating behavioral characteristics. The trajectory from the early BSD installations to the most recent has seen a steady increase in scale, software complexity, and sculptural elements. As each iteration has gradually added more components to our overall inventory of materials, the later installations have been able to adapt more readily. In addition, over the series of BSD installations, multiple collaborators with different specializations have shaped the technical and iterative design of the project to dovetail with his or her existing research. From audio feedback systems to sensor networks to machine-learning classifiers to stereoscopic projection and robotic speaker systems, the combinations of novel forms of mediated experience have allowed BSD to explore various methods for constructing embodied experiences.

This paper is interested in the development of an element of the BSD project known as Bird Song Diamond Mimic (BSD Mimic). Following the first installation in 2014, every installation has contained the system, either as a peripheral element or as the primary focus of the interaction.

The most recent installation at the Ars Electronica Center in Linz, Austria, during the 2017 Ars Electronica Festival, featured the most developed and integrated BSD Mimic system to date. Taking advantage of gained insights from the previous versions, the interaction design was more streamlined and the underlying software more sophisticated. Audience members were invited to participate by attempting to mimic the birdsong of different species of birds drawn from Europe and Japan. Presented within the unique Deep Space 8k theatre, immersive large-scale projected visuals with ten channels of audio supported the thematic aspects of the system while providing visual feedback from the BSD Mimic system.

The installation in Deep Space 8k integrated three diverse research interests for the first time in an art installation: large-scale swarm simulation (Mototake and Ikegami 2015), controllable clusters of fixed-position directional speakers, and a birdsong classification system for annotating many localized sound sources using a microphone array (Chiba et al. 2017). While the first two systems were used for establishing the audiovisual environment, the third system provided a means for evaluating similarities between human and bird vocalizations. By comparison with the other two systems, BSD Mimic required far greater interdisciplinary collaboration between both the researchers studying bird song classification and the artists developing the installation.

Even the existing software used in the research was integrated with the installation software, so that audience members were able to directly interact with the algorithms used in classifying birdsong. This BSD Mimic system became the primary mode of interaction with the audio and visual elements of the installation, while also creating a performative platform for the participants. BSD Mimic also allowed for post-performance analysis of how well individual vocalizations approached a target birdsong; further testing the limit of the birdsong classification software beyond its normal use case. The aggregate results of the performances were stored to evaluate overall user performance and quantitatively reveal the success or failure of the interaction.

A further outcome from receiving quantitative information regarding the BSD mimic system was that comparisons between collaborative interaction and individual interaction could be made. Having more concrete evidence regarding the effectiveness of different modes of interaction can assist with future installation planning. Iterations of BSD Mimic had made use of either solo or collaborative interaction design, and while observation had led to the hypothesis that collaborative interaction improved participant scores and overall success of the BSD Mimic experience, it was now possible to more formally evaluate data collected from the exhibition and post-exhibition experiments.

Following an overview of the background research underlying the BSD project (Sect. 2), the paper will cover the development of the BSD Mimic system divided into individual (Sect. 3.1) and collaborative (Sect. 3.2) interactions. Section 3.2.2 and its subsections will give a more detailed overview of the Deep Space 8k system and how BSD Mimic connects to the larger installation. Section 3.2.2.3 gives an evaluation of the Deep Space system before concluding with Sect. 4.

2 Research history

Bird Song Diamond (BSD) is a project that we have been pursuing since 2012 and integrates the research of the involved scientists and engineers with the practices of artists and designers. We have taken scientific research performed in the laboratory and adapted it to function in an installation setting with interactivity. The chronological order of past events and installations is shown in Table 1.

2.1 Bioacoustics and linguistics

The origins of this project with research into the grammar and meaning of bird songs have brought together computational linguists, electrical engineers, and artificial life specialists. For example, novel methods of data collection, soundscape partitioning, and automated annotation have been developed (Suzuki et al. 2017) (Kojima et al. 2017). More sophisticated analysis of bird song structure (Sasahara et al. 2012) was developed, and, subsequently, further insight into the syntactic and grammatical capacity of various species (Taylor et al. 2017). Work on soundscape capture and analysis has led directly to the development of the BSD Mimic system (see Sect. 3.1.3).

Table 1 Chronology of major BSD exhibitions with name and location

Year	Exhibition
2012	Project begins with the initial stages of planning
2014	Sculpture Garden UCLA, CA
2015	Art|Sci Gallery, California NanoSystems Institute (CNSI) UCLA, CA
	New York Electronic Arts Festival (NYEAF) Governor's Island, NY
	Times Square New York, NY
	Nanolab, CNSI UCLA, CA
2016	Large Space Tsukuba, Japan
	Artificial Life and Robotics 21 (AROB) Beppu, Japan
	Waterbodies Los Angeles, CA and Harvestworks, NY
2017	Deep Space 8k, Ars Electronica Festival Linz, Austria
	Speculum Artium Trbovlje, Slovenia

2.2 Artificial life

The sound and visual development of the project is also directly related to natural systems and how such systems react in the presence of humans (anthropogenic products). Work has been done to provide a more detailed explanation of the sonic aspects of landscape ecology, resulting in a "soundscape ecology" (Pijanowski et al. 2011) that takes into account the collective biological, geological, and anthropogenic sounds (the biophony, geophony, and anthrophony, respectively), which comprise the entire soundscape (Krause 1987). Our research has contributed to the simulation of natural processes and environments to evaluate emergent behavior and real-world effects (Ikegami et al. 2012) (Sumitani et al. 2017).

Using BSD as a platform to create such artificial systems in real-world situations, as interactive installations, provides both unique audience engagement as well as opportunities for researchers to witness and experiment with artificial life interactions with the real world. The first BSD installation in

2014 took place at the sculpture garden on the UCLA campus and employed artificial life techniques. This audio-only iteration, incorporated into the trunk and branches of a large Coral tree, acted as an autonomous sensor network (ASN), in which interconnected sensors took in environmental information (light and humidity), processed incoming data based on a chemical reaction model, and then modified its own sonic output based on these changes (Maruyama et al. 2013). The multi-channel audio was output through 20 parametric speakers, which are specialized speakers that produce highly directional "beams" of sound[1] (Shi and Gan 2010), giving the impression of many individual sound sources emanating from the branches of the tree. Due to the unique qualities of the speakers, the sonic effect was that of artificial birds inhabiting the branches of the tree. Throughout the installation, the artificial sounds of the speakers and natural sounds of birds and wind and people created a complex soundscape underneath the branches of the Coral tree. As the day progressed and the environment shifted, the ASN in turn reacted and shifted its sounds.

2.3 Artist involvement

Artist involvement was essential to the realization of the primary goal of the BSD project: public outreach. We have been able to successfully present BSD in festivals and public spaces thanks to the contributions of artists, allowing thousands of people to experience the project. Installation work, and especially interactive installation, requires historical and experiential knowledge to place the content of the research into contemporary cultural discourse. The artists developed the narrative structure and interaction design for each installation, and evaluated these aspects in preparing for successive installations. In addition, the artists often participated in the installation by aiding audience members, explaining the system, and troubleshooting issues. Artists also consistently worked and communicated with researchers, leading to collaboration throughout the entire development process. They participated with scientists in the laboratory and out in the field, as well as invited the researchers to contribute to the development of the installations. As the director of the installation, the artist must take the role as a facilitator of specialists, working to bring the expertise of the individual researchers into the installation. The artist must also find coherence and connections between the aesthetics of the installation and the research, so that the final outcome is unified.

3 Mimic and installation development

BSD has been presented in evolving forms across multiple locations since the first installment in 2014. Discussion of each iteration will be divided based on the type of interaction used in the Mimic system starting with an overview of individual followed by collaborative experiences. Details about other aspects of the installations will be covered to provide context for how BSD Mimic related to the larger BSD project. The Deep Space 8k installation will occupy a large portion of the collaborative experience section due to its relevance in the iterative design process and the importance that BSD Mimic played in the system.

Because the presentation of installations will not be in chronological order, a brief timeline of the installations is provided here for reference. As described above, the first BSD installation was presented at the UCLA sculpture garden in 2014. The following year, BSD was shown in Governor's Island and Times Square in New York. In 2016, BSD traveled to Japan to show at the LargeSpace immersive display in Tsukuba. The BSD Mimic system was also brought to the 21st Artificial Life and Robotics (AROB) symposium in Beppu, Japan, within the same month. September 2017 was the installation in Deep Space 8k at the Ars Electronica Festival in Linz, Austria. Following the festival, the BSD Mimic system traveled to Trbovlje, Slovenia for the Speculum Artium festival. Numerous other exhibits, containing elements of these landmark presentations, were conducted throughout this period (see Table 1 for a summary).

3.1 Individual experiences

The BSD Mimic system individual configuration consisted of a single microphone, headphones, and optional display (Fig. 1). A participant could listen to instructions in the headphones prompting them to attempt to mimic a given bird song. After the participant's attempt, the software evaluated the audio information, resulting in a score and modification to a visual element on the screen. Audio-related elements had been contained within a single piece of software built using the SuperCollider language and real-time rendering system, which is an application and programming language designed for audio synthesis and composition (Wilson et al. 2011). All analyses and comparisons between bird and user vocalizations were done in real time by storing the fast Fourier transform (FFT) information,[2] at a rate of 40 times per

[1] Parametric speakers use ultrasonic frequencies to minimize the natural spread of waves over distance. Audible sound is used to modulate an ultrasonic wave which is emitted from the parametric speaker, while a second, unmodulated, ultrasonic carrier wave is also sent from the same speaker. When the two waveforms collide with an object, they demodulate to produce audible sound that is the difference between the modulating and carrier waves (Woodford 2018).

[2] FFT is a signal processing technique in which a complex waveform, in this case sound, is broken down into its sinusoidal components. Using this technique, one can estimate the most prominent frequency

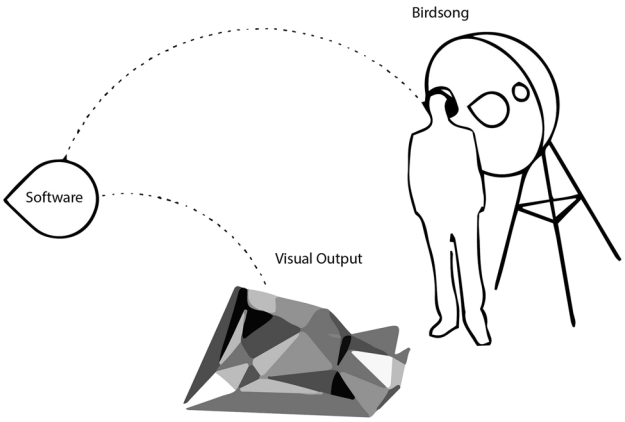

Fig. 1 Single-user configuration for the BSD Mimic system. One hemisphere station with headphones and Diamond display for visual feedback

second, of a saved recording of bird song and live recording of a user. From there, the pitches and volume levels of the bird song and recording were compared, resulting in a value. This value was placed between empirically derived minimum and maximum values, mapping the value to a percentage which was then presented to the user.

3.1.1 Times Square

While presenting BSD on Governor's island, which included a collaborative version of the mimic system, a temporary mimic performance was presented in Times Square. Participants stood on a platform to elevate themselves above the crowds while wearing noise canceling headphones to further separate themselves from the sounds of the crowded street (Fig. 2). The sudden shift to quiet, natural sounds of birdsong from the artificial soundscape of Times Square highlights the lack of integration of nature within urban landscapes. Attached to the microphone stand, a tablet computer was used to display results in the form of a growing diamond generated using the Processing programming language (Reas and Fry 2006) and controlled from the Supercollider audio program using the Open Sound Control (OSC) communication protocol (Schmeder et al. 2010). The diamond would increase in size based on how accurately the participant mimicked the bird song. The elevation of the participant and audible separation from the chaotic surroundings of Times Square proved to be a successful effect in focusing attention on bird song. At the same time, the busy nature of the

Footnote 2 (continued)

of a sound, usually the pitch, by finding the frequency of the component wave with the highest amplitude.

Fig. 2 Times Square mimic system

Fig. 3 Photograph of the LargeSpace without projections (Empowerment Informatics 2014)

location and the "spotlight" nature of being on a pedestal made it difficult to find participants even with large numbers of people in the immediate area.

3.1.2 Large Space

In the winter of 2016, BSD was presented at the University of Tsukuba through an immersive display called LargeSpace (Fig. 4). This display extends the paradigm of the Cave Automatic Virtual Environment (CAVE) system as a one-to-many format display (Cruz-Neira et al. 1993) by dramatically increasing the footprint of the display area. The LargeSpace measures 25 m × 15 m, with a height of 7.7 m enabling large groups of people to enter the display. Stereoscopic projections, created by 12 projectors around the upper perimeter of the space, cover the walls and floor

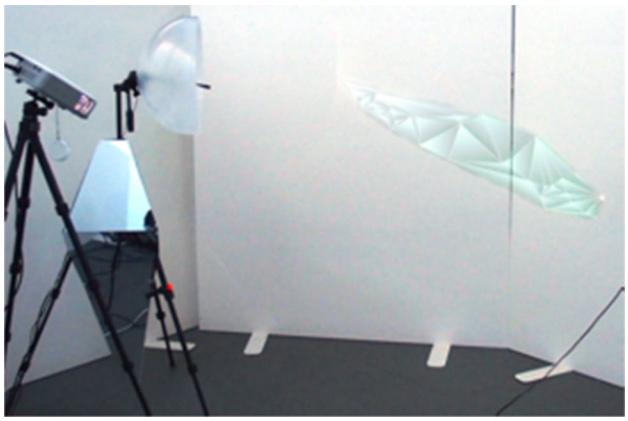

Fig. 4 LargeSpace BSD mimic installation. Smaller hemisphere with the early prototype diamond stand. External projection

while also accounting for the curvature of the projection surface (Fig. 3). In addition to the display, the LargeSpace also allows for tracking a high number of participants using a 20 unit OptiTrack system. LargeSpace also has a wire-driven motion base to provide a single participant the ability to move vertically in the space (Takatori et al. 2016).

Prior iterations of BSD took a more traditional interaction design approach to installations in that the works were constantly running and able to be engaged by audience members at any time. The LargeSpace version of BSD, due to limitations on access to the space and the necessity for technicians to be involved at all times, shifted to be more performance oriented. With scheduled performance days and times, audiences would arrive together and all experience the project simultaneously. Audience members all wore active shutter glasses for viewing the stereo imagery as well as custom-designed bird wings embedded with tracking system markers to determine each person's location in LargeSpace. Each performance lasted 15 min and was followed by a short discussion about the LargeSpace, the performance elements, and the underlying research.

The performance was composed of four distinct phases which presented a narrative portraying birds and bird perspectives in both natural and artificial forms. A single audience member, wearing an origami-inspired bird costume and attached to the LargeSpace motion base, would move through the space to indicate shifts from one phase to the next. Unique sounds, making use of the directional parametric speakers from the previous installations, were created from a mixture of natural bird songs and composed artificial sounds (Maruyama et al. 2014). In addition, interactive flocking simulations were used in different phases to juxtapose natural elements, represented by paper cranes, and artificial elements, represented by drones. The narrative culminated in a real drone flying into the space and influencing the flocking simulation with the use of a tracking marker on the body of the drone, and the origami bird audience member descending back to the floor to initiate a growing diamond (video link).

BSD Mimic was set up outside of the LargeSpace projection room (Fig. 4), and audience members waiting to experience the LargeSpace performance could test their abilities to mimic bird song. The instructions for BSD Mimic were translated to Japanese, and the collection of bird songs was changed to native Japanese birds to reference the local soundscapes that would be familiar to those living in Japan. Adjusting the system to use native bird songs tested a person's existing familiarity with their surroundings, and additionally sought to have participants reflect on their future engagement with environmental sound. To make this version of BSD Mimic more portable, the pedestal to change the height of the participant was removed. A growing diamond shape was again used for visual feedback, but was projected on a nearby wall rather than displayed on a screen.

While many users were able to participate with the installation, it was clear that the scale and the spectacle of BSD in the LargeSpace drew much attention away from the BSD Mimic system. Even while not participating in the LargeSpace performance, audience members frequently watched the performance from outside the display area rather than attempting BSD Mimic itself.

3.1.3 Beppu

Participating in the AROB conference during the same month as the LargeSpace performance allowed for a second installation of BSD Mimic as an entirely standalone, individual interaction experience (Chacin et al. 2016). This installation tested the portability of the system as the entire piece had to be carried by the exhibitors in standard sized luggage. The system was shown outside of the conference room in a slightly darkened, yet open, area of the convention center. In the individual configuration, the audio of the piece was entirely contained within the participant's headphones with the audience able to hear only the participant's mimic attempts. This version tested the use of a semi-transparent, white plastic hemisphere as a container for the microphone which could double as a projection surface for the visual feedback of the growing diamond (Fig. 5). In addition, this was the first version of the BSD Mimic system in which the stand holding the dome and microphone was covered in a mirror-finish acrylic diamond shape. This shape added a physical analogue to the projected diamond, and also doubled as a storage space for the computer running the system.

As a logistical experiment, the difficulty in transporting the hemisphere, which was smaller than hemispheres used in prior installations, proved that if portability was a primary goal for the system, then the current hemisphere material would not be sufficient. This led to later tests using different

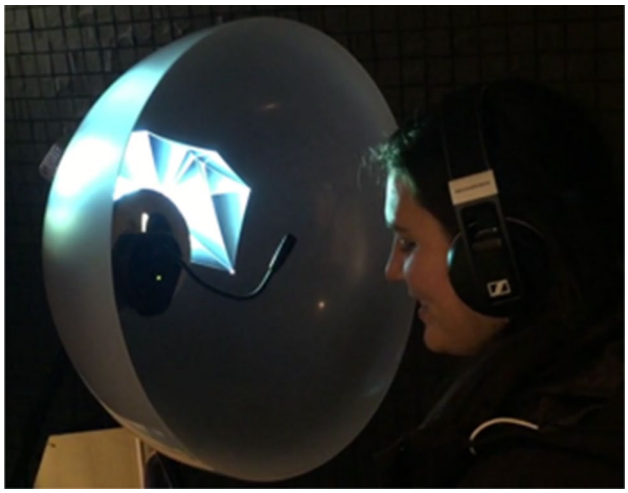

Fig. 5 BSD Mimic installation at Beppu. Similar to LargeSpace installation, but more compact. Diamond projected directly on the hemisphere

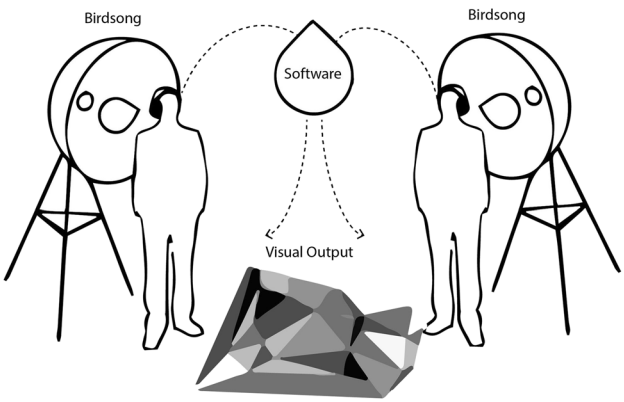

Fig. 6 Two-user BSD Mimic configuration. Two hemisphere stations' both users contribute to the shape of a single-diamond visual

materials and collapsible constructions while still attempting to maintain the same hemispherical shape. Furthermore, observation of the participants showed that many were hesitant to use the system. While many showed interest, the number of participants remained low. The context of the work in a scientific academic conference with a few demonstrations may have contributed to this hesitancy, though social dynamics and feeling of becoming the focus of attention was also evident.

3.2 Collaborative experiences

Adding a second BSD Mimic station, microphone with stand and hemisphere, allowed the system to be shared by a second user (Fig. 6). Whereas the individual system was mostly about internal reflection and pulling a person out of an anthropological context, the collaborative system encouraged participants to share the experience with each other. Each participant could listen to both the target bird song as well as listen to the other participant's attempts at mimicking the bird song. The growing diamond became a visual record of the abilities of each user, adding an additional competitive feeling to the interaction.

3.2.1 Governor's Island

As part of the 2015 New York Electronic Arts festival, BSD was to be installed on Governor's Island in the magazine of a former US Army post in the eighteenth and nineteenth century, now national park, known as Fort Jay. The installation was in a partially subterranean set of rooms with a distinctive "steampunk" aura. One section of the installation consisted of parametric speakers, used in the 2014 installation (Sect. 2.2), mounted around the ceiling of the central room with a projection aimed at the wall opposite to the entrance. The unique layout and acoustics of the underground space allowed for a sharp contrast between the cold, military surroundings, and the sounds of local bird song coming from the speakers. Based on the amount of movement within the space, the projection would cycle between different vantage points shot from a bird's perspective ranging from ground level to far above.

In two long, reverberant rooms attached to the central room, the BSD Mimic system was installed. This was the first publicly shown iteration of the mimic system. Participants received instructions from speakers placed in their respective rooms before singing into the microphone. The two stations on opposite sides of the main room were set up, so that each participant could see the other from across the central room. Both participants could hear the other as they attempted to mimic each bird song creating a collaborative and slightly competitive dynamic. To indicate the state of the competition, the diamond was projected on the floor of the central room and would grow in the direction of the participant who received a better score. When the diamond reached the edge of the projection area, the diamond would be replaced by a new starting diamond in the center of the room.

Most attendees came in groups of two or more, usually on family outings or small day trips to the island. This being the case, there were rarely problems in finding multiple users to participate in the mimic system. The unique sound properties of the mimic rooms added another incentive for participants to vocalize, as the strong echo was pleasant to hear. While the separation of the rooms provided a direct view of the second participant, the visual feedback on the floor was not easy to view making the connection for the participants unclear at times.

3.2.2 Deep Space 8k: major redesign installation integration

Presenting BSD in the Deep Space 8k at the Ars Electronica Center in Linz during the 2017 Ars Electronica Festival offered the opportunity to expose a very large and diverse audience to the project. The Ars Electronica Festival is one of the premier festivals in the world to showcase projects that engage in art and technology. The average attendance to the festival for the five festivals leading up to 2017 was 82,500 people with a wide demographic of festival attendees. The Deep Space 8k is a large-scale immersive display in a similar category to that of the LargeSpace that was recently upgraded from the original version of the system to handle both 8k resolution[3] on the wall and floor surfaces (Kuka et al. 2009). With similar limitations as the LargeSpace regarding length and frequency of performance, it was initially decided to expand upon the systems from the prior LargeSpace performance, while adapting to the differences in display, sound, and tracking capabilities of the Deep Space 8k system.

One of the main goals for this iteration of the project was to fully integrate the BSD Mimic system with two other aspects from the previous installations: the large-scale visualization of swarms of birds with audience interaction; multi-channel directional audio produced by parametric speakers. Incorporating these visual and audio aspects was further supported through the creation of a robotic positioning system allowing the directional speakers to follow the positions of the swarms of birds. By combining the three systems into a cohesive experience, it was expected that the increase in stimuli would further immerse the participants into the sensory experience of a bird. It was important to consider the connection between all the aspects of the installation to construct an experience in which the audience feels responsible for the changes in the installation. Interactivity, in this respect, is meant to give a sense of ownership to the work, leading to longer engagement with the source material. Providing further cohesion amongst all elements, the BSD Mimic microphone stands were sculpturally enhanced to reflect the surrounding visual elements. Placed directly into the projection area of the Deep Space 8k, the stands became a focal point, in both visual and interaction, for the performance (Fig. 7).

Much of the BSD Mimic system had to be re-evaluated and rebuilt to integrate it as a component along with the visual and audio aspects of BSD in Deep Space 8k. Formerly, the BSD Mimic system had been used as a standalone system with the primary goal of having a portable and easy to implement counterpart to the large scale and highly involved flagship BSD installations. Having some variability in potential implementations, the core of the system consisted of one or more microphones mounted to a hemisphere(s) and connected to a computer. Output from the computer was usually achieved through headphones attached to the computer which facilitated the audio-based interaction. Depending on the setting, visual feedback could also be created on a screen or projection surface.

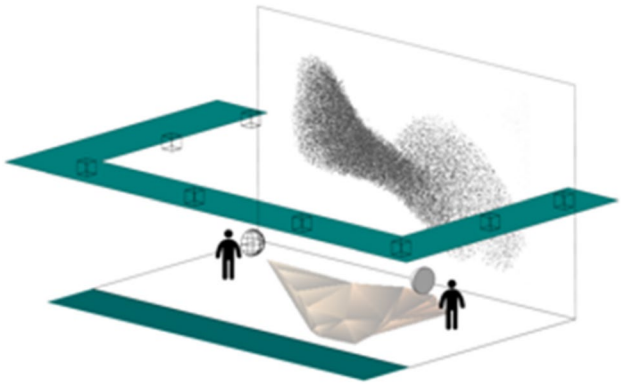

Fig. 7 Diagram of the Deep Space 8k installation. Wall projection features boid simulation. Floor projection features diamond. There were eight parametric speakers around top and two BSD mimic stations on floor

While relatively simple in description, in practice certain parts of the system, especially the sculptural components (hemispheres and stands) which housed the microphones did not follow the original design strategy of portability. On the software side, the program was relatively straightforward to set up, and made use of a graphical user interface to simplify the specification of input and output ports, make minor adjustments to the difficulty level, and adjust levels for noise reduction. However, even when the system had been set up and calibrated for a specific location, the output of the system, consisting of a percentage representing a user's accuracy at reproducing bird vocalizations, fluctuated in unpredictable ways with dramatic variations in percentages that often created more confusion and less time that each user spent with the system. In developing an updated version of the BSD Mimic system for the Deep Space 8k performance, both the physical and software shortcomings were considered while also working to integrate the mimic interaction with the larger performance.

Attending to the issue or portability, the hemisphere used to house the microphone was considered. Previously, transparent plastic hemispheres were either shipped to the installation location or purchased on site when transport was impractical. As these hemispheres were molded plastic with 80–100 cm diameters, their size alone proved too large to fit into any conventional suitcase. The only possibility for

[3] 8k resolution being 8192×4320 pixels. The system also uses a refresh rate of 120 Hz (Sick-Leitner 2015).

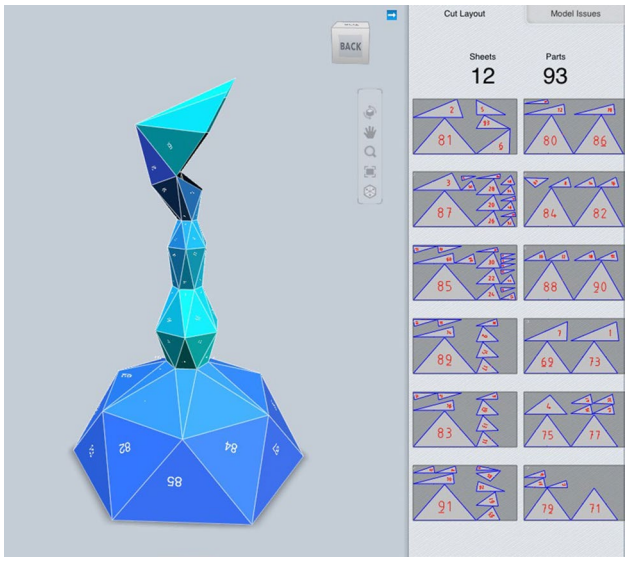

Fig. 8 Screenshot of digital fabrication process with the stand cover for Deep Space 8k using Autodesk Fusion 360 software. Left side is the 3D model of the stand. Right side includes the split triangles arranged on sheets to be cut with a laser cutter

maintaining such a diameter while reducing size and weight was to consider lighter materials that could be folded into more compact forms. However, after consulting with experts in foldable structures and evaluating a range of materials from plastic sheets to paper, it was decided that the aesthetic qualities of solid, transparent, plastic domes outweighed concerns with portability. Rather than continuing with the task of more portable structures, the former stand and hemisphere configuration were maintained.

For the hemisphere structures of the BSD Mimic system to match with the visual projections planned for the Deep Space 8k, a faceted shell, which borrowed visually from the diamond structures projected on the floor of the space (Fig. 7), was to cover the functional stand structure. These shells expanded on the earlier reflective shells used in the Beppu installation described above (Sect. 3.1.3). An artist created a 3D model of the structure, which was split into individual panels using the Autodesk computer-aided design (CAD) software Fusion 360 (Autodesk 2018) and exported as a series of scalable vector graphics (SVG) files for use with a laser cutter (Fig. 8). After cutting sheets of mirror-finish acrylic plastic into each individual face, the shell was constructed into segments and then attached around the outside of the metal support structure. Because the stands could not be in Deep Space 8k before or after the performance, the metal stand was attached to a custom-designed wheeled platform (see Supplementary Materials for images of fabrication). Wireless microphones were used to simplify the movement between spaces and to prevent the visual distraction and potential trip hazard of audio cables traversing the floor between the microphones and audio interface. The entire process was repeated for a second hemisphere. During the performance, the hemispheres were placed over the two diamond structures with platforms to allow users to reach the microphones. Using a mirror finish on the outer shell of the stands, broken reflections from both the projected visuals as well as the audience and participants within Deep Space 8k could be seen across the surfaces.

The software side of the BSD Mimic system was rebuilt to replace the former version for improved accuracy and consistency for a range of users in varying environments. These improvements required the separation of the software into two separate components:

1. the underlying program for analyzing and producing predictions;
2. the front end which served as the user interface for interacting with the installation and, additionally, acted to relay data (audio recordings, scores, triggers) between the first component and the software running the audio and visual aspects of the installation.

As mentioned before, the original BSD Mimic system, while easy to install and maintain, yielded inconsistent predictions and was difficult to protect against small changes in environmental noise, microphone quality, and variations in user vocalizations (vocal range, loudness, breath, and production means: whistling, humming, and singing). To make improvements would require a more sophisticated approach, so work was done to adapt techniques used in the scientific research on soundscape capture and analysis to be used for interspecies vocalization comparisons between humans and birds.

The resulting software made use of machine-learning techniques to create a trained model of birdsong based on hundreds of field recordings for 12 different species of birds. Using this model, a recording from a participant could be fed into the system to produce a feature space distance between the human vocalization and the nearest bird vocalization. This distance represents the "closeness" in which the participant's song came to the song that they were attempting to mimic.[4] As with the previous versions of the system, this value is then mapped to a percentage, from 0 to 100, to provide the user with a more familiar evaluation accuracy to the target birdsong.

The second, frontend component of the redesigned BSD Mimic system was built as an encapsulated object in the visual programming language Max (previously known as Max/MSP) (Cycling'74 2018). This Max object managed

[4] See the Deep Space 8k Mimic Scatterplot in Supplementary Materials for the overall plot of participant mimic attempts against the existing birdsong.

the sequential aspect of the mimic interaction. After recording a participant's vocalization to a specified directory, the object would request an evaluation from the backend by sending an OSC message specifying the target bird song and which of the two participants was being evaluated. Upon receiving a reply message containing the mapped percentage value, additional OSC messages were sent to the software managing the visuals (see Sect. 3.2.2.1) instructing the visual elements to shift accordingly. The percentage was relayed to the participant with an additional message and then the sequence waited for the next participant to approach the microphone. Because the interaction medium was audio-based and Max was already used to drive all of the sounds during the performance, including the mimic audio component as a Max object within the larger sound control software reduced the operational complexity of having to juggle an additional program.

Instructions were given verbally through pre-recorded text-to-speech (TTS) synthesis using stock Apple System Voices in two different languages, "Karen" for English and "Otoya" for Japanese, to accommodate international users. For an event in Slovenia after the Deep Space 8k performance, an additional voice needed to be recorded. Because there were no available TTS voices in the Slovenian language that matched the quality of the other two languages, verbal instructions were recorded by a human. The inclusion of languages and bird song from different parts of the world reflects both the backgrounds of the researchers and artists working on the project as well as the locations in which the project was previously shown.

3.2.2.1 Visual elements The Deep Space 8k, as its name implies, offers up to 8k (8192×4320 at 120 Hz) resolution projections onto connected wall and floor surfaces each measuring 16 m×9 m. Despite the decrease in physical size from the LargeSpace, the Deep Space 8k's projection surfaces were much higher resolution than any prior installation and required a redesign of the visual system. Furthermore, the layout of the Deep Space 8k meant that the 360-degree immersion offered by the LargeSpace was no longer available and that the visuals would need to be structured to keep audience attention facing forward. The performance time was lengthened to a full 30 min, which forced the decision of whether to lengthen the existing LargeSpace phases or to develop a different strategy for this installation. Ultimately, it was decided to take aspects from the prior installation, but not to have them as fixed-length sections. This iteration would focus on the flocking swarms (Ikegami et al. 2017) which would evolve over the course of the performance based on the interactions from participants engaging with the mimic system. The floor projection screen would be partially occluded by audience members, so, rather than displaying complex visuals, the diamond shapes from the Governor's Island installation would be shown as an added representation of the mimic interaction taking place during the performance.

Due to the interactive nature of the performance, it was not possible to pre-record the visual objects and play them back as video. Every object which would respond to audience interaction needed to be generated in real time and modified based on information coming from the BSD Mimic system. The LargeSpace visuals had been written using Open Graphics Library (OpenGL)[5] and were tailored to only work with the unique demands of the size and shape of the space. It was found that making changes to the LargeSpace visuals, even ones which seemed trivial, was plagued by slow turnaround times and limited by the number of people familiar with the system. As an additional problem, the time available for testing in Deep Space 8k was extremely restricted, and the team could not easily visit the system to test and evaluate the qualities of the visuals. Within these constraints, the system needed to work well in a live setting and be able to accommodate easy and fast alterations in layout, speed, and behavior of the various objects being displayed. Based on these limitations, and from the list of compatible platforms with Deep Space 8k, it was decided to use the Unity game engine to drive all of the visual outputs of the installation (Unity3D 2018).

Unity enabled the team to prototype scenes quickly and make modifications directly to those scenes in the editor. Therefore, when the team was able to visit Deep Space 8k prior to the performance, we did not experience significant changes between playback in the Unity Editor and the program builds running on the Deep Space 8k servers. The downside to this system, and one that contributed to delayed development, was that each export needed to contain two builds—for the floor projection and the wall projection. This meant that, for an object to appear to move from the floor to the wall, it was necessary to synchronize the transforms (position, scale, and rotation) of a duplicate of that object. The transform information would pass from the Wall program locally, via UDP,[6] to the Floor program. For efficiency, we set only certain objects to listen and receive these transform updates, that is only the objects which might be shown on both the wall and floor. Ultimately, as a design choice, we completely separated the wall and floor projections and no longer needed to synchronize any of the objects.

The first major visual element was a simulation of two massive swarms of origami cranes and drones projected on the main wall of Deep Space 8k. The simulations, based on Reynolds's boids model (Reynolds 1987), expanded the

[5] OpenGL is an application programming interface (API) used to render computer imagery (OpenGL 2018).

[6] User Datagram Protocol (UDP), a communication protocol which is part of the Internet protocol suite (along with TCP and IP). OSC mentioned before is built on top of UDP (Postel 1980).

Fig. 9 Rendered video of massive boid simulation

prior size limitations by implementing the model to take advantage of GPGPU[7] architecture (Mototake and Ikegami 2015). This allowed for swarm sizes of up to 500,000 individual boids as opposed to only a few 100 using the traditional implementations. The processing overhead for producing the simulation was not feasible in real time at such a high resolution. Therefore, the large-scale simulation was rendered as a stereoscopic video and a small-scale interactive simulation would be generated in real time (Fig. 9).

Video playback in Unity was achieved using the AVPro Video plugin (Renderheads 2018). While Unity offers a video player component, it was not capable at the time to play back video as stereoscopic at the required resolution. Using AVPro, we could achieve stereo 4k video playback in addition to the other elements within the scene. The video was encoded using FFmpeg (FFmpeg 2017) in H.265 (High Efficiency Video Encoding or HEVC)[8] at a variable bitrate to reduce file size and improve streaming playback.

The real-time simulation was adapted from the boid implementation used in the LargeSpace performance to make use of Unity "prefab" instances. Improvements were made regarding flexibility for the size and position of the simulated area. In addition, by interpolating the forward vector of each individual boid toward its respective velocity vector, boid movement became more natural looking and closer to that of the movement in the prerecorded simulation. Two separate simulations, one for origami cranes and the other for origami drones, were initialized at the beginning of the performance. The number of boids in each simulation would increase or decrease based on interaction from the Mimic system. Placing the real-time simulation directly in front of the stereoscopic video would give the appearance of a more immediate, sharper resolution, swarm and a distant large-scale swarm.

The floor projection consisted of an updated version of the diamond shapes previously used as visuals for the mimic system. The original diamonds used in past BSD Mimic systems were two-dimensional shapes generated by applying a Delaunay triangulation (Carlbom 1994) to a cluster of points corresponding to each diamond. Increasing or decreasing the size of the diamonds was done by adding a new point on the periphery of the diamond or subtracting an existing point from the outer edge of the diamond. With each change in the number of points, a new triangulation was applied to maintain a closed shape. By extracting the individual triangles generated in the triangulation and applying slightly different brightness values to each shape, it could give the appearance of a faceted diamond.

However, the two-dimensional diamond shapes did not easily translate to the Unity software, so it was necessary to consider an alternative method to produce a three-dimensional diamond. At first, the point-to-triangulation technique described above was tested, but did not create results that were satisfactory. Rather than attempting to procedurally generate geometry, the existing 3d models of diamonds could be used as the initial unmodified diamonds. This avoided many of the problems that procedurally generating triangulated geometry posed, especially related to shading, point duplication, and properly generating normal vectors for each vertex. The existing points of the ready-made geometry could be moved toward or away from the center of its bounding box to shrink or grow over the course of the performance. For the performance, two diamonds were placed at the base of each mimic stand, and were associated with the respective participant's mimic scores. The movement of the diamonds, and their final shapes at the conclusion of the performance, provided a cumulative indication for the audience and participants of the overall success at the BSD Mimic interaction.

3.2.2.2 Audio elements The audio elements for Deep Space 8k were adapted from those used during the LargeSpace installation. Translating the sound design and hardware from the LargeSpace involved scaling down the number of sound sources as well as drastically improving the flexibility of the hardware (Fig. 10). The reduction in size of the Deep Space 8k meant that the number of directional speakers could be reduced while still covering the majority

[7] General Purpose Computing on Graphics Processing Units (GPGPU) uses multiple graphics cards for parallel processing of intensive calculations (Mung and Mann 2004).

[8] A successor to the h.264 video compression standard, HEVC, was necessary to reduce file size for 8k resolution video (Sullivan et al. 2012).

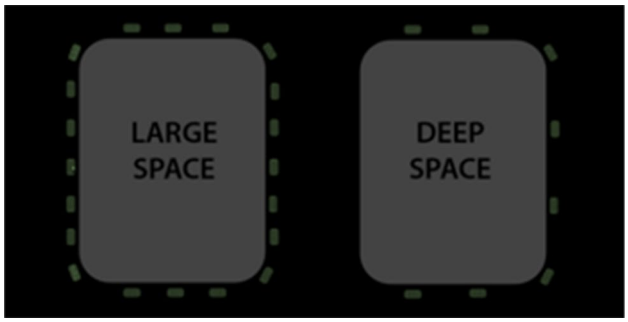

Fig. 10 Reduction of parametric speakers from LargeSpace to Deep Space 8k. For Deep Space 8k, the projection wall is the side with no speakers

of the space. In addition, the acoustic properties of the Deep Space 8k, especially the projection surfaces, enabled the sound to reflect about the room to a much greater extent. A further difference was the added ability to control the direction of the speakers by mounting each individual speaker to a dedicated pan-tilt mechanism, allowing custom static or dynamic speaker configurations. This enabled fewer speakers to match the sonic capabilities of many speakers in a fixed configuration. With these factors in mind, the number of directional speakers was reduced from 20 to 8. Deep Space 8k also provided a powerful 5.1 audio system that was used with stereo audio to produce a more dynamic range of audio.

The directional speakers used for Deep Space 8k were parametric speakers sold as kits by the TriState company (TriState). These same speakers were used in prior iterations of BSD for creating the impression of many individual sound sources without having to use the large number of speakers required in other spatialized sound techniques, e.g., wave field synthesis (Berkhout 1988). While the frequency response of the parametric speakers is limited to 0.4–5 kHz, the speakers are able to reproduce bird song very effectively as the vocalizations used in the installation fell within this range.

Parametric speakers allow for audio to be "beamed" in a particular direction using modulated ultrasonic frequencies that form audible frequencies when coming in contact with a solid object (Yoneyama et al. 1983; Pompei 1999). Aiming a speaker at a wall or other object can create the effect of that object emitting sound while aiming at a listener's ear will cause that sound to appear as if it was right there. In the first installation of BSD, at the UCLA sculpture garden, the parametric speakers were aimed at the leaves of the tree where it was installed, which gave the impression of sound coming from different areas of the canopy overhead. Taking advantage of the occasional breeze, the changing reflection angle of the leaves added further complexity to the sound on the ground (Norihiro et al. 2014).

The sound emitted from the parametric speakers was a generative composition created in the visual programming language Max. The program assigned each of the eight speakers a different bird song, and over time changed out different songs or silenced the speaker. The repetitions of the calls changed depending on the selected song.

Additional sound was played over the Deep Space 8k speaker system. The composition was built from accumulated sounds of unmanned aerial vehicles (UAVs or drones), in particular were consumer drones which bear the names of birds (Parrot, Global Hawk, Grey Eagle, etc.). As the performance progressed, the intensity and amplitude of the drones increased until overpowering all other sounds in the space until finally dissipating at the conclusion (Lyons et al. 2017).

3.2.2.3 Mechanical pan-tilt system Recreating the effect of multiple shifting sound sources that added to the sonic effect of the outdoor installations for the indoor Deep Space 8k installation required a mechanical system to give programmatic control of the direction that each speaker faced at any given time. Each speaker was attached to a servo-controlled pan-tilt mechanism, which, in turn, was attached to the ceiling of Deep Space 8k. This allowed the parametric speakers to point at specified locations within Deep Space or to move based on the simulations of natural processes.[9]

The individual speaker mechanisms were all connected to a microcontroller that facilitated servo movement, which, in turn, was connected to a computer running a custom Unity plugin for providing a visual interface of the pan-tilt configuration (Lynxmotion 2017). By using a virtual recreation of the Deep Space 8k layout, the plugin let users place virtual speakers into the space corresponding with the locations of the physical speakers. The user could then "connect" the virtual speakers to the physical speakers by assigning the respective numerical value indicated on the physical network box. Once the virtual and physical speakers were connected, users could control the physical speakers in a number of different ways. While possible to manually change the rotation of the speakers, it is much easier to use the plugin's targeting system to instruct the speaker where to aim. Any virtual object can be assigned as a speaker's target and that speaker will always attempt to aim at that object; however, the physical speaker will only move as far as its range of motion allows, the limits of which can be assigned in the software. A target can be manually moved, animated, attached to physics simulations, or connected to another object's motion.[10]

[9] See Supplementary Materials for more detailed explanation of the pan-tilt system.

[10] Compatibility with Kinect v2's HD face detection algorithm was also implemented to enable speakers to follow the tracked face of a participant.

For the installation in Deep Space 8k, the speakers were assigned to individual boids in the small-scale live simulation. The directional speakers would then move about the space with the simulated paper cranes and drones to recreate the constantly shifting audio that featured in the previous outdoor installations. In addition, by being attached to a simulation that the audience could influence, additional aspects of the audio, location and movement, became part of the entire simulation of the event. A system for mimicking natural processes should consider the ways that all the aspects of the system can support a particular behavior. While the previous installations used only fixed-position directional audio, it was the addition of a dynamic position system that connected the audio with the visual in a very direct way.

3.2.2.4 System evaluation The Deep Space 8k during the Ars Electronica Festival had to accommodate a wide variety of performances—from children's programs, to live performers, and a number of different physical elements including the BSD Mimic stands. Due to these constraints, BSD could only occupy the performance space for a total of 40 min per day, however, with time required for setup and removal of equipment between 20 and 30 min remained for the performance. While installed at the LargeSpace in Tsukuba, BSD had been adapted from a constantly running installation to fit into a 15-min performance window, so, initially, the LargeSpace performance was used as a template for the Deep Space 8k. As the work developed, the integration of the BSD Mimic system required rethinking about how to combine the more passive audio and visual elements from the LargeSpace with the interaction and spectacle of participants mimicking bird vocalizations.

The BSD Mimic system was the focal point of the performance, and while the audio and visuals maintained their immersive quality, they acted as a backdrop for the mimic performance. Improving from the previous iterations of BSD Mimic, including the boids, as part of the visual feedback added an additional aspect to connect the audience members with the immersive projection space. While the diamonds on the floor were present, audience members viewing the performance often blocked the projections making them not as useful for participants. Because many audience members were in attendance and each performance was monitored by team members, there were no problems in finding people to participate and further instruction or guidance could be administered in the event of any confusion.

Sound interference became an issue at times for the participants. The proximity of the two mimic stations made for a more collaborative feeling, but participants could not always hear the other person as the ambient sound increased in volume and the overall activity in the space intensified. Recording the scores of 211 mimic attempts over 4 days of performances, the overall mean accuracy was 68.16%.

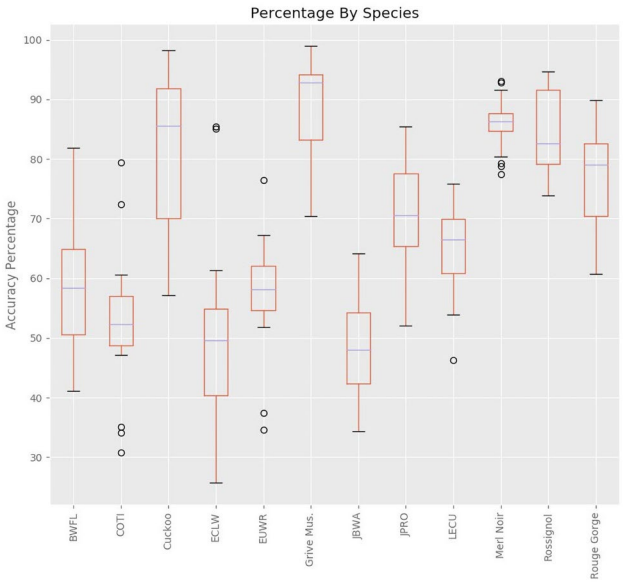

Fig. 11 Box plot of the mimic percentages separated by species of bird. Unabbreviated species names can be found under additional information

Separating the attempts based on species of the bird yielded a wider range of results, with certain species such as the Song Thrush receiving consistently high-accuracy percentages and others, Japanese Bush Warbler and Eastern Crowned Leaf Warbler, receiving much lower scores (Fig. 11).[11] A variety of factors could account for the variation between species of birds including audience familiarity with only certain species of birds, acoustic distance of a bird song from natural human vocalization, or simply differences in quality of bird song recordings.

4 Conclusion

The BSD project is a multifaceted and multidisciplinary installation with the aim of bringing contemporary research on bird communication to a large public audience. With this goal in mind, BSD has exhibited in a variety of international venues and reached thousands of people from USA to Japan to Europe. Developing the project has been an inherently iterative endeavor with each presentation of BSD testing new possibilities for transmitting the research while updating the existing aspects of the project.

The BSD Mimic component of the project has had the most variegated exhibition history of any aspect of BSD, with only a handful of the numerous public showings being

[11] The raw data from the performance can be accessed here: https://bitbucket.org/johnbrumley/bsd-datasets.

described in this paper. The iterative design process led to artistic insights regarding the beneficial nature of collaborative interaction for promoting audience engagement with the subject matter. Technical improvement of the BSD Mimic software, based on observations and user feedback, required researchers to consider alternative applications for their research technology. The demands of software for art installations called for more robust and flexible systems which could be implemented with ease by artists at the installation site. Furthermore, the consideration of context for each exhibition, incorporating language, cultural iconography, and local bird song, solidified audience connections at each installation location. Rather than discarding these site-specific elements at each new location, allowing them to travel with the installation highlighted the exhibition history of the work. Finally, having staged BSD Mimic in some settings with individual user interaction and others with collaborative user interaction motivated a closer inspection of the role that collaboration played for users of the BSD Mimic system. As BSD Mimic is a unique interaction platform, it will be necessary to investigate more collaborative artworks in both unregulated installation settings and controlled experiments to make a more general statement regarding this topic.

Acknowledgements Research was supported by the National Science Foundation (Grant ID: 1125423). Additional support provided by the Program for Empowerment Informatics at the University of Tsukuba and The University of California, Los Angeles. Martin Cody provided many of the recordings used throughout BSD. For people who have been involved with the BSD project over its various incarnations: Naoaki, Chiba, Jun Mitani, Yan Zhao, Joel Ong, Max Kazemzadeh, Itsuki Doi, Norihiro Maruyama, Hikaru Takatori, Aisen Chacin, Masa Jazbec, Takeshi Oozu, Mary Tsang, Carol Parkinson, Linda Weintraub, and many others.

References

Arriaga JG, Sanchez H, Hedley R, Vallejo EE, Taylor CE (2014) Using Song to identify Cassin's vireo individuals. A comparative study of pattern recognition algorithms. In: Martínez-Trinidad JF, Carrasco-Ochoa JA, Olvera-Lopez JA, Salas-Rodríguez J, Suen CY (eds) Pattern recognition. Springer International Publishing, New York, pp 291–300

Berkhout AJ (1988) A holographic approach to acoustic control. J Audio Eng Soc 36:977–995

Boursier-Mougenot C (1999) From Here to Ear [Zebra Finches, Electric Guitars]. Traveling Installation

Brancusi C (1928) Bird in Space. [Bronze, 54 × 8 1/2 × 6 1/2″ (137.2 × 21.6 × 16.5 cm)]. Museum of Modern Art, New York

Bugler C (2012) The bird in art. Merrell, London

Calder A (1971) Eagle [Steel, Painted]. Seattle Art Museum

Carlbom I (1994) Modeling and visualization of empirical data. In: Rogers DF, Earnshaw RA (eds) State of the art in computer graphics: aspects of visualization. Springer New York, New York, pp 19–65. https://doi.org/10.1007/978-1-4612-4306-9_3

Chacin AC, Jazbec M, Oka M, Doi I (2016) Bird Song Diamond: call and response and phase transition work. In: The twenty-first international symposium on artificial life and robotics 2016 (AROB 21st 2016), Beppu, Japan

Chertow MR (2008) The IPAT equation and its variants: changing views of technology and environmental impact. In: Mitchell RB (ed) SAGE library of international relations: international environmental politics, vol 4. SAGE Publications Ltd., London, pp 87–87. https://doi.org/10.4135/9781446262108.n5

Chiba N, Sumitani S, Matsubayashi R, Suzuki R, Arita T, Nakadai K, Okuno HG (2017) An improvement of HARKBird: a wild bird song observation and analysis tool based on an open-source robot audition software HARK. In: Proceedings of the 35th annual conference of the Robotics Society of Japan, RSJ2017ACA3-03

Crist E (2013) On the poverty of our nomenclature. Environ Hum 3:129–147. https://doi.org/10.1215/22011919-3611266

Cruz-Neira C, Sandin DJ, DeFanti TA (1993) Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In: Proceedings of the 20th annual conference on computer graphics and interactive techniques (SIGGRAPH '93). ACM, New York, NY, USA, pp 135–142. https://doi.org/10.1145/166117.166134

Cycling'74 (2018) Max software tools for media. https://cycling74.com/products/max/. Accessed 10 Feb 2018

Dooling RJ (1982) Auditory perception in birds. In: Kroodsma DE, Miller EH (eds) Acoustic communication in birds, vol 1. Academic, New York, pp 95–130

Ehnes J (2010) An audio visual projection system for virtual room inhabitants. In: 20th international conference on artificial reality and telexistence Proceedings, p 118

Empowerment Informatics (エンパワーメント情報学) (2014) http://www.emp.tsukuba.ac.jp/english/environment/research.php. Accessed 10 Feb 2018

FFmpeg (2017) FFmpeg. https://www.ffmpeg.org/. Accessed 10 Feb 2018

Gibson JJ (1986) The ecological approach to visual perception. Lawrence Erlbaum Associates, Hillsdale

Graham M (2011) Through birds' eyes: insights into avian sensory ecology. J Ornithol. https://doi.org/10.1007/s10336-011-0771-5

Harraway D (2016) Tentacular thinking: Anthropocene, Capitalocene, Chthulucene. E-Flux J. http://www.e-flux.com/journal/75/67125/tentacular-thinking-anthropocene-capitalocene-chthulucene/. Accessed 9 Feb 2018

Head M (1997) Birdsong and the origins of music. J R Mus Assoc 122(1):1–23. https://doi.org/10.1093/jrma/122.1.1

Hedley R (2016) Composition and sequential organization of song repertoires in Cassin's vireo (*Vireo cassinii*). J Ornithol 157:13–22. https://doi.org/10.1007/s10336-381015-1238-x

Hein HS (1990) The exploratorium: the museum as laboratory. Smithsonian Institution Press, Washington

Ikegami T, Oka M, Maruyama N, Matsumoto A, Watanabe Y (2012) Sensing the sound web. In: Art gallery at the 5th ACM SIGGRAPH conference and exhibition on computer graphics and interactive techniques in Asia, exhibited

Ikegami T, Mototake Y-I, Kobori S, Oka M, Hashimoto Y (2017) Life as an emergent phenomenon: studies from a large-scale boid simulation and web data. Philos Trans Ser A Math Phys Eng Sci. https://doi.org/10.1098/rsta.2016.0351

Kac E, Bennett E, Connell B, Peragine J, Bynaker C, Lindsay M (1996) Rara Avis. http://www.ekac.org/raraavis.html. Accessed 10 Feb 2018

Kelley M (1978) Birdhouses [Wood, Paint]. Mike Kelley Foundation

Kojima R, Sugiyama O, Hoshiba K, Nakadai K, Suzuki R, Taylor CE (2017) Bird Song scene analysis using a spatial-cue-based probabilistic model (special issue on robot audition technologies). J Robot Mechatron 29:236–246

Kraft D (2013) Birdsong in the music of Olivier Messiaen. Arosa Press, London

Krause B (1987) Bioacoustics, habitat ambience in ecological balance. Whole Earth Rev 57:14–18
Kuka D et al (2009) DEEP SPACE: high resolution VR platform for multi-user interactive narratives. In: Iurgel IA, Zagalo N, Petta P (eds) Interactive storytelling. ICIDS 2009, lecture notes in computer science, vol 5915. Springer, Berlin
Legrady G, Pinter M, Bazo D (2013) Swarm Vision [3 custom designed rails each with Sony PTZ camera, custom software animation, Apple MacPro, 2 projectors (Panasonic PT-DZ6710U or equivalent) or 2 HD large screens, dimensions variable]. https://www.mat.ucsb.edu/g.legrady/glWeb/Projects/sv/swarmvision.html. Accessed 12 Feb 2018
Lynxmotion (2017) SSC-32U USB Servo Controller Board user guide. http://www.lynxmotion.com/images/data/lynxmotion_ssc-32u_usb_user_guide.pdf. Accessed 10 Feb 2018
Lyons M, Brandis K, Callaghan C, McCann J, Mills C, Ryall S, Kingsford R (2017) Bird interactions with drones, from individuals to large colonies. bioRxiv. https://doi.org/10.1101/109926
Malm A (2018) The progress of this storm: on society and nature in a warming world. Verso, London
Maruyama N, Oka M, Ikegami T (2013) Creating space-time affordances via an autonomous sensor network—semantic scholar. In: 2013 IEEE symposium on artificial life (ALife), pp 67–73
Maruyama N, Doi I, Masumori A, Oka M, Ikegami T, Vesna V, Taylor C (2014) Evolution of artificial soundscape in a natural environment. In: Exploiting synergies between biology and artificial life technologies: tools, possibilities, and examples at ALIFE, p 14
Massumi B (2002) Parables for the virtual: movement, affect, sensation. Duke University Press, Durham
Milk C, Tricklebank B, George J, Meyers A, Chasalow B (2012) The Treachery of Sanctuary [Projection]. Traveling Installation
Mototake Y, Ikegami T (2015) A simulation study of large scale swarms. SWARM 2015, Kyoto University, Kyoto, pp 446–450
Mung J, Mann S (2004) Using multiple graphics cards as a general purpose parallel computer: applications to computer vision. In: Proceedings of the 17th international conference on pattern recognition (ICPR2004). Cambridge, United Kingdom, vol 1, pp 805–808
Nagel T (1974) What is it like to be a bat? Philos Rev 83(4):435–450. https://doi.org/10.2307/2183914
Norihiro M, Doi I, Masumori A, Oka M, Ikegami T, Vesna V, Taylor C (2014) Evolution of artificial soundscape in a natural environment. In: Exploiting synergies between biology and artificial life technologies: tools, possibilities, and examples, ALIFE, p 14
OpenGL (2018) OpenGL—the industry standard for high performance graphics. https://www.opengl.org/. Accessed 10 Feb 2018
Pijanowski B, Villanueva-Rivera L, Dumyahn S, Farina A, Krause B, Napoletano B, Pieretti N (2011) Soundscape ecology: the science of sound in the landscape. BioScience 61(3):203–216. https://doi.org/10.1525/bio.2011.61.3.6
Pompei FJ (1999) The use of airborne ultrasonics for generating audible sound beams. J Audio Eng Soc 47(9):726–731
Postel J (1980) RFC 768: user datagram protocol. https://tools.ietf.org/html/rfc768. Accessed 10 Feb 2018
Reas C, Fry B (2006) Processing: programming for the media arts. AI & Soc 20:526. https://doi.org/10.1007/s00146-006-0050-9
Renderheads (2018) AVPro Video. Retrieved February 10, 2018, from http://renderheads.com/product/avpro-video/
Reynolds CW (1987) Flocks, herds, and schools: a distributed behavioral model. In: ACM SIGGRAPH computer graphics, pp 21–25
Sasahara K, Cody ML, Cohen D, Taylor CE (2012) Structural design principles of complex bird songs: a network-based approach. PLoS One 7(9):e44436. https://doi.org/10.1371/journal.pone.0044436
Schmeder A, Freed A, Wessel D (2010) Best practices for open sound control. In: Linux audio conference, Utrecht, NL
Shi C, Gan WS (2010) Development of parametric loudspeaker. IEEE Potentials 29(6):20–24. https://doi.org/10.1109/MPOT.2010.938148
Sick-Leitner M (2015) Deep Space 8k—the next generation—Ars Electronica feature. https://www.aec.at/feature/en/deep-space-8k/. Accessed 10 Feb 2018
Simon T (2014) Birds of the West Indies. Gagosian Gallery, Los Angeles
Sosolimited H, Design P (2016) Diffusion Choir. Biomed Realty, Cambridge
Stengers I, Goffey A (2015) In catastrophic times: resisting the coming barbarism. Open Humanities Press, London
Sullivan GJ, Ohm JR, Han WJ, Wiegand T (2012) Overview of the high efficiency video coding (HEVC) standard. IEEE Trans Circuits Syst Video Technol 22(12):1649–1668. https://doi.org/10.1109/TCSVT.2012.2221191
Sumitani S, Suzuki R, Arita T, Naren, Matsubayashi S, Nakadai K, Okuno HG (2017) Field observations and virtual experiences of bird songs in the soundscape using an open-source software for robot audition HARK. In: abstract book of 4th international symposium on acoustic communication by animals, pp 116–117
Suzuki R, Matsubayash S, Hedley R, Nakada K, Okuno HG (2017) HARKBird: exploring acoustic interactions in bird communities using a microphone array. J Robot Mechatron 29:213–223
Takatori H, Enzaki Y, Yano H, Iwata H (2016) Development of the large scale immersive display LargeSpace. Nihon Virtual Reality Gakai Ronbunshi 21(3):493–502
Taylor C, Brumley JT, Hedley R, Cody ML (2017) Sensitivity of California thrashers (*Toxostoma redivivum*) to song syntax. Bioacoustics 26:259–270. https://doi.org/10.1080/09524622.2016.1274917
Thrift N (2007) Non-representational theory: space, politics, affect. Routledge, London
Tuchman M (1971) Art and technology: a report on the Art and Technology Program of the Los Angeles County Museum of Art, 1967–1971. Los Angeles County Museum of Art; distributed by the Viking Press, New York
Unity3D (2018) Unity Game Engine [online] http://unity3d.com/. Accessed 1 Feb 2018
Wainwright J, Mann G (2018) Climate Leviathan. Verso Books, London
Wark MK (2016) Molecular Red: theory for the Anthropocene. Verso Books, London
Wilson S, Cottle D, Collins N (2011) The Supercollider Book. The MIT Press, Cambridge
Woodford C (2018) Directional loudspeakers—how they work. http://www.explainthatstuff.com/directional-loudspeakers.html. Accessed 19 Feb 2018
Yoneyama M, Fujimoto J-i, Kawamo Y, Sasabe S (1983) The audio spotlight: an application of nonlinear interaction of sound waves to a new type of loudspeaker design. J Acoust Soc Am 73(5):1532–1536
Yu K, Yin M, Luo J-A, Wang Y, Bao M, Hu Y-H, Wang Z (2016) Wireless sensor array network DoA estimation from compressed array data via joint sparse representation. Sensors (Basel, Switzerland) 16(5):686. https://doi.org/10.3390/s16050686