HIGH-ORDER SURROGACY FOR THE AUDIOVISUAL DISPLAY OF DANCE

Stefano Kalonaris

RIKEN
Music Information Intelligence Team
Tokyo, Japan
stefano.kalonaris@riken.jp

Iannis Zannos

Ionian University
Department of Audiovisual Arts
Corfu, Greece
zannos@gmail.com

ABSTRACT

The current pandemic (COVID-19) has had considerable impact on many fronts, not least on the physical presence of humans, affecting how we relate to one another and to the natural environment. To investigate these two interactions, the notion of *surrogacy*, originally described by Smalley as remoteness between source and sonic gesture, is considered and extended to include bodily gesture, for the rendering of contemporary dance performances into abstract audiovisual compositions/objects. To this end, for a given dance performance, sonification of the motion capture data is combined with video-frame processing of the video recording. In this study, we focus on higher order surrogacy and associate this with 1) a soundscape ecology-inspired approach to sonification, whereby three species of sounds coexist and adapt in the environment according to the symbiotic paradigm of mutualism, and 2) a wave space method to sonify their coevolution. Aesthetic implications of this procedure in the context of multimodal, telematic/remote and virtual systems are discussed as disembodied presence emerges as a dominant trope in our daily experience.

1. INTRODUCTION

Auditory display (or *sonification*), is the practice of rendering, representing or interpreting data through the aid of sound processes. It has been said that "sound becomes sonication when it can claim to possess explanatory powers" [1, p. 213], and it is generally agreed upon that whenever it is intended that the listener "understands extra-musical information" [2, p. 178] through the sound process, then one can speak of sonification.

Inferring information or eliciting meaning from a different domain thanks to sound appears to be a *sine qua non* for sonification, however, the more abstract or metaphoric the mapping function between the data and the sound is, the more removed the two become. This relationship has been described in [2], perhaps the most comprehensive attempt to integrate a multiplicity of approaches, goals and motivations that converge in this discipline. There, a two-axes aesthetic perspective system to understand and design sonification is put forward, outlining a continuum between *sonification concrète* and *abstraite*.

1.1. Surrogacy

Abstract, removed, metaphorical, conceptually blended, analog, and so forth, are all related terms to indicate the relationship between the signified and the signifier, with varying degrees of distance. To query the space between signified and signifier, we consider the notion of *surrogacy*. This is not to be solely intended as one process acting in place of another, but it specifically refers to the notion formulated by Dennis Smalley in his theory of *spectromorphology* [3], whereby one can distinguish four types of surrogacy: first, second and third-order, and remote surrogacy. In Smalley's words, these are associated to a continuum between whether one can "recognise source [. . .] and type of gestural cause" to when "either gesture or cause becomes dubious" [3, p.112]. In other words, a path from the literal through to the inferred and, finally, to the vestige.

2. AIMS & MOTIVATION

Interested in higher order surrogacy, we conceptually extend Smalley's definitions to encompass both the sonic and the visual domain, using contemporary dance performance motion capture (hereinafter, mocap) and video data as input sources for the blending of audiovisual structures into an abstract narrative. The choice to investigate surrogacy in the context of contemporary dance is due to an ongoing collaboration with practitioners of this style, which might lead to an interactive implementation of our procedure in the near future. Such a dance idiom, compared to other styles, is inherently more abstract (*e.g.*, less dependent on strong beat marking, defined rhythms, and so forth), which might make it more conducive to our framework.

We posit that *audiovisual surrogacy* can be a useful notion in the context of multimodal environments where tele- or virtual presence is featured, if one wishes to depart from explicit correspondences, to "elicit sound-image disjunctures" [4, p. 95]. This can be particularly relevant in a world that has been profoundly changed by the COVID-19 pandemic. The latter, has dramatically re-configured the role of physical presence and interaction, prompting enquiries beyond the social sphere, to include performing arts and aesthetics. With specific reference to our domain of interest (dance), a 'disembodied' dimension of gesture, an audiovisual object which maintains but a residue of the original physical component, might help elicit novel avenues for representation, interpretation and fruition of the aesthetic experience.

Motivations behind auditory display are seemingly polarised between empirically-based or arts-driven, although more comprehensive views [5] can help to include gradations between the two.

In [6], it is posited that it might be helpful to express stronger intentions, either towards one or the other, and that trying to occupy the space in between might risk to dilute both. With this in mind, this study leans firmly towards a sonic arts-derived sensibility, foregrounding the implications of using a sonification approach which increases the disjuncture between its sonic output and the data used to produce it.

3. CONTEXT

The procedure proposed in Section 6 is twofold: the output object is synthesised by combining a visual and an audio element which are currently obtained separately and from different data relating to the same dance performance (video and marker-based mocap data, respectively). Therefore, although at times difficult to separate these domains, we divide the related work accordingly.

3.1. Visualisation of Dance

From the Forsythe's *Improvisation Technologies* [7] to more recent applications of abstract visualisation of dance movement [8, 9], a wealth of approaches and strategies have been employed for technologically mediated movement analysis and representation. For the visualisation of mocap data, the leading approach is to render it using using plotting libraries, web-based scripting[10], animation software and game engines. The latter are increasingly ubiquitous and often employed for the rendition of *virtual dance*, offering creative potential in shaping novel approaches to choreography [11]. While non-photorealistic rendering (NPR) [12] techniques have been used to display images or videos in an artistically stylised manner, the remoteness with respect to the input data is rarely such as to mask the source, or even to render it unrecognisable. Some have experimented with pushing the intelligibility limit further, as in the case of the *[mViz]* project [13], or in the works of digital artist Stas Sumarokov[1]. Of particular relevance regarding fully-fledged artistic works that blend audiovisual display of dance performances is the work of Nicolas Salazar Sutil, for example *REACH*[2], which features the remote interaction of a musician and a dancer, via audiovisual feedback, or *Flatland*[3], featuring interactive graphics and choreographic human movement analysis.

3.2. Sonification of Dance

In the context of dance mocap data sonification, numerous multimodal systems have been developed, most notably Rokeby's *Very Nervous System* [14], the Multisensory Expressive Gesture Applications (MEGA) project [15, 16], *MotionComposer* [17] and the *il-SoP* system [18]. While there are several methods of sonification, the vast majority of these systems are limited to approaches that map data to sound synthesis parameters. Examples of such techniques include *Audification* [19] and *Parameter Mapping Sonification* (PMS) [20]. In the former, data of interest would be normally sonified directly (as a series of sound pressure values), although processes of re-sampling or filtering might be applied (*e.g.*, to accommodate the human hearing range, etc.). In PMS, on the other hand, data preparation and *mapping* or *transfer* functions are employed, so as to reveal structures in the data. Less common is the use of *Model-Based Sonification* (MBS) [21], which focuses on the interaction with a *parameterised sound model* derived from the data. In other words, a virtual sound object akin to an instrument is being "played" through the interaction with the user, according to some interaction specification. Even less common, is the use of *Wave Space Sonification* (WSS) [22], which scans a scalar field along a data-driven trajectory. Other sonification methods, such as *Earcons* or *Auditory Icons* (sound placeholders for an event or action, typically used as auditory aids) are not apt for artistic rendition of data, and are of little interest to the authors.

We note that these different sonification methods are also subtending a path from the concrete to the abstract: the relationship between the data and the sounds becomes increasingly less explicit as one moves from Audification to WSS. To date and to our knowledge, no application of WSS methods in contemporary dance has been tried; we embark in such a task, drawing from soundscape ecology and coevolution, which we look at now.

[1]https://www.youtube.com/user/exsstas/videos
[2]https://vimeo.com/74704777
[3]https://vimeo.com/45150759

4. INSPIRATION

Why soundscape ecology and coevolution? Disembodied presence as a novel social interaction trope is not the only effect that has arisen from the spread of COVID-19. For example, the reduced human physical presence in the natural environment (due to lockdown control measures) has prompted an overdue reconsideration of our relationship with it. This has allowed quantifiable insights into human-wildlife interactions [23], climate change [24], and sustainability [25]. The effects of the recently reduced human mobility have manifested in all facets of our interaction with the environment, not least in the sound domain, with repercussions on, for example, high-frequency seismic ambient noise [26] or on urban noise pollution [27]. Soundscape ecology is the study of the relationship between living organisms (human or not) and the environment, through sound, and it made sense to draw inspiration from this field of enquiry in the context of our study.

4.1. Soundscape Ecology

Although sometimes the term soundscape ecology is used interchangeably with *acoustic ecology*, the two differ. While definitions are flexible and non-exclusive, the latter, which originates from the work of R. Murray Schafer [28] and Barry Truax [29], seems more concerned with the human perception of everyday sounds and with the relationships between humans and the sonic environment, whereas the former considers the interactions of three basic different 'sound species' that constitute the environment. According to soundscape ecology, the sound domain is divided into *geophony*, *biophony*, and *anthrophony*. Examples of the first category might include water sounds (marine waves, waterfalls, streams, rain, etc.), wind, thunderstorm, lightning, and so forth. Biophonies are instead related to the sound production of vocalising species in a given environment. Finally, anthropophonies are normally intended as the sounds originating from activities related to humans, with focus on technology. Examples of this category might be sounds from machinery, industries, locomotion vehicles, explosions, and urban areas. However, anthropophony might also include music and language.

4.2. Mutualism

Mutualism [30, 31] (a flavour of *symbiosis*) describes an ecological interaction where two or more species (might) benefit from the interaction itself. Mutualism can be useful to model coevolution and, in our framework, it is used as a proxy for the *niche hypothesis* [32], which is defined only within biophony. The simplest way to model the mutualistic interaction of three species is through the Lotka-Volterra equations [33]. In the type I functional response model [34], which allows unbound growth, these are expressed as follows:

$$\begin{aligned}\frac{dN_1}{dt} &= r_1N_1 - \alpha_{11}N_1^2 + \beta_{12}N_1N_2 + \beta_{13}N_1N_3\\ \frac{dN_2}{dt} &= r_2N_2 - \alpha_{22}N_2^2 + \beta_{21}N_1N_2 + \beta_{23}N_2N_3\\ \frac{dN_3}{dt} &= r_3N_3 - \alpha_{33}N_3^2 + \beta_{31}N_3N_1 + \beta_{32}N_3N_2\end{aligned} \quad (1)$$

where N_i is species *i*'s density, r_i the intrinsic growth rate of species *i*, α_i the negative effect of within-species *i* crowding, and β_{ij} the benefit of mutualistic interaction between species *i* and *j*.

5. SONIFICATION METHOD

To develop our approach, we employ the definitions and notions of soundscape ecology and mutualism described above. For the auditory display, we use WSS.

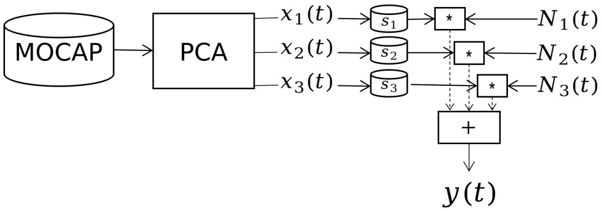

Figure 1: Diagram of the sonification model.

More precisely, a static sample-based flavour which, in [22], defines the wave space (V) as follows:

$$V(\vec{x}) = \frac{1}{d}\sum_{i=1}^{d} s_i(c_i \cdot x_i) \quad (2)$$

for some time-indexed data $\vec{x}(t) \in \mathbb{R}^d$, and where s_i is a sound file and c_i a scaling coefficient. Breaking down the different elements of this equation for this particular sonification, $\vec{x}$ is obtained by means of feature projection, more precisely by computing the first three principal components (PCA). Correspondingly, there are three sound samples (s), one for each principal component and for each sound species described in Section 4.1: geophony, biophony, and anthropophony. The choice of these sound files is arbitrary and specific to the application of the method. For example, they could be recorded ad-hoc by the sound designer or sourced via other means. Each scaling factors c_i is derived from the sound species population N_i in the mutualistic model described in Section 4.2, and thus is also a function of time. These coefficients can then be mapped to any sound property of interest. For example, the coefficients could be used as the gain factors for three band-pass filters (BPF), each with a corresponding band matching the spectral content of each species. These mappings are case-specific and bound to the sound designer's aesthetic, desires and goals. The normalised values of $\vec{x}$, instead, are mapped to both the length of the samples in the sound files (so that the trajectories subtended produce sample scanning/scrubbing) and to panning positions in the stereo field. Finally, the dynamically scaled and panned sample scans for all three sound species are combined into the output sound scene. The general scheme for the sound process is shown in Figure 1. As for the initial condition of the sound species population model, it is necessary to have a heuristic measure of the weights for the species. A common measure used in ecoacoustics is the *Normalised Difference Soundscape Index* (NDSI) [35]:

$$NDSI = \frac{\beta - \alpha}{\beta + \alpha} \quad (3)$$

In the NDSI equation, α represents the power spectral density (PSD) of sound in the range 12 kHz, normally associated with anthropophony, while β is the PSD associated to biophony, with a range of 211 kHz. It is important to note that these ranges are approximate and not exclusive. Biophony, in fact, might include frequencies below the 2 kHz. The NDSI ranges from -1 to +1 (anthropophony to biophony, respectively), and it does not account for geophony. This is due to the latter's wide spectral range. Inspired by this measure, and needing to include geophony, the PSD for all three sound species is calculated; then, the area under these curves is computed and scaled between 1 and 10 (population range). This is a rather arbitrary procedure, however, lacking objective data as to the real-world ratios between the different **phonies*, it was deemed "good enough" for a proof of concept, in the awareness that a more rigorous strategy should be sought after.

6. A PRACTICAL EXAMPLE

For a practical example, we proceeded as follows: spatial information of a dance performance obtained with a camera/marker-based system was used for the sonification, while the video recording of the same performance was used to obtain an abstract representation of the bodily gestures. Subsequently, sound and visual output components were then recombined to render the final audiovisual object, as shown in Figure 2.

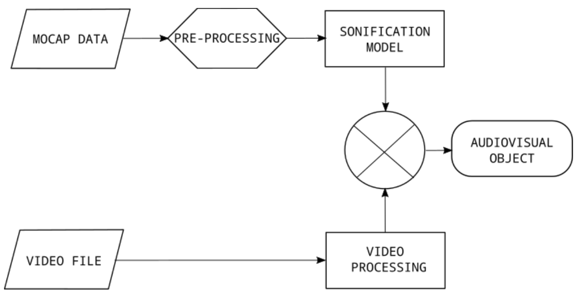

Figure 2: Procedure's flow diagram.

6.1. Data

The mocap data and video file of the dance performance used for a prototypical implementation was selected from the contemporary dance subset in the University of Cyprus' Dance Motion Capture Database[4]. There are 132 dance performances in this subset, recorded using 8 double array cameras and with dancers wearing a mocap suit of 38 markers. All performances are short, between 15 and 150 seconds, thus apt for rapid prototyping. These performances are offered in several formats (*e.g.*, FBX, C3D, etc.), although not consistently, and comprise spatial information in 324 dimensions (joints' position and rotation), sampled at 30 frames per second. Of these performances, the 36 featuring both as Biovisual Hierarchical (BVH) data and in mp4 video format were selected as candidates and, eventually, the first version of Andria Michaelidou's mini-dance with the "afraid" tag was chosen for testing our procedure. The BVH data (which provides detailed skeleton information) was parsed using the PyMo[5]Python module. The performance mocap data was preliminary processed by means of smoothing (with a Butterworth low pass filter of 3rd order, with a cutoff frequency of 0.1 Hz), and normalisation, to make it invariant to differences in the joints' ranges. Subsequently, the first three principal components were calculated from the 885 data points in the mocap file (see Figure 3).

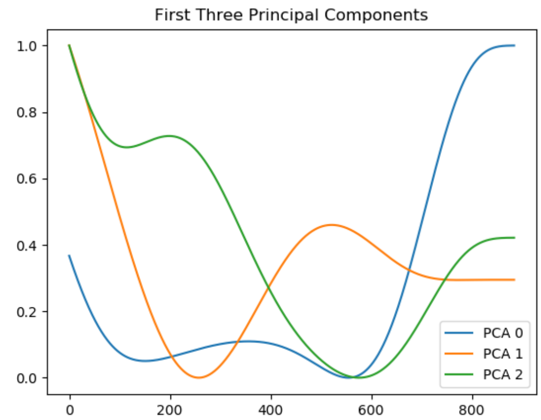

Figure 3: Feature projection with PCA.

As for the sound data, we used three stereo files [36, 37, 38] obtained from Freesound[6], recorded with a 48kHz sampling rate and with a bitdepth varying from 16 to 32 bit.

6.2. Sonification

Using the chosen audio recordings, we obtained the PSD shown in Figure 4, which provided us with the initial condition for our population model.

Running the coevolution simulation for t time steps equal to the data points in our mocap file yields the dynamics shown in Figure 5. In our example the coefficients' trajectories so obtained affect only the amplitude of the corresponding sound files.

[4]`http://dancedb.eu/`
[5]`https://github.com/omimo/PyMO`
[6]`https://freesound.org`

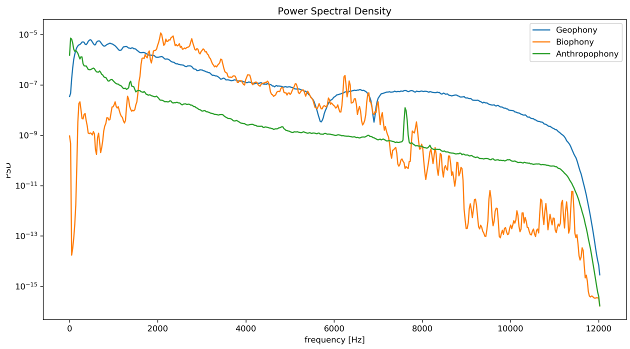

Figure 4: Power Spectral Density (PSD) of the three recordings.

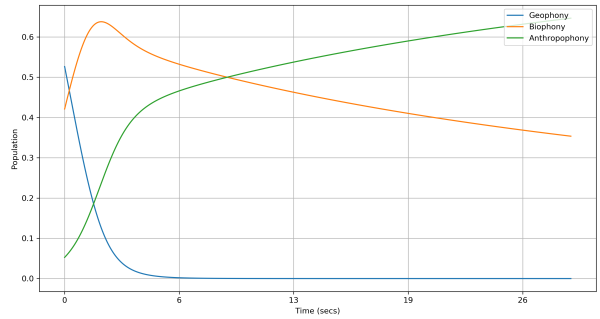

Figure 5: Three-species population model with initial conditions obtained from PSD analysis of the sound data.

6.3. Visualisation

In our coarse prototype we simply applied video processing to the dance performance, by means of edge detection, background subtraction [39] and optical flow estimation using the Lukas-Kanade algorithm [40], to increase the *remoteness* effect. A sample results is shown in Figure 6.

The rendered audiovisual object described so far can be accessed at `https://vimeo.com/558887679.` whereas the code is hosted at `https://gitlab.com/skalo/hoavsur`.e

7. DISCUSSION & FUTURE WORK

This study can be framed as an artistic investigation of two factors foregrounded by the current pandemic: disembodied (human) presence as a new *de facto* viewpoint of daily experience (to include aesthetic), and dialectic enquiry on the impact of human presence on the natural environment. This investigation was carried out in the domain of sound and dance, to render surrogate audiovisual compositions/objects. Particular emphasis was given on the abstract rather the concrete, capitalising on sonification and mapping procedures that seemingly increase the remoteness of the source from the output. This endeavour was treated as a generative task rather than an empirically-based sonification/visualisation,

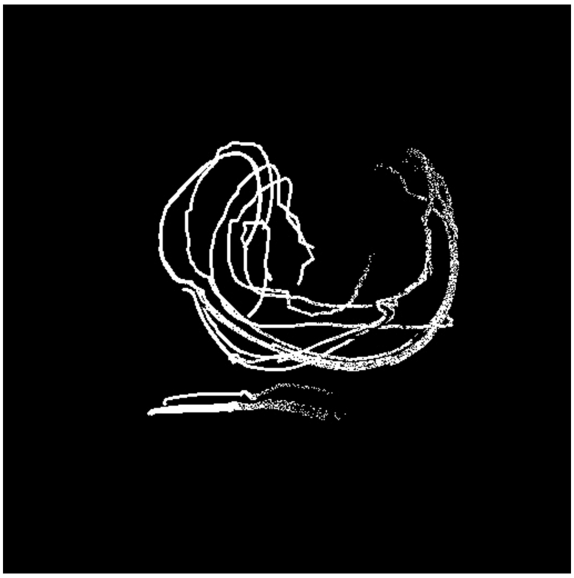

Figure 6: A screenshot of the visual rendition of a dance performance.

embracing a sensibility that, in the broader field of aesthetics, has moved towards a pragmatist stance, which "accepts not only the existence but the value of indeterminacy" [41]. Thus, relating back to the sonification's power of explaining or uncovering information (see Section 1), we are more inclined to adopt Shusterman's concept of interpretation, whereby one does not seek to discover/uncover meaning but to construct it, instead.

The study presented is but a proof of concept, an initial exploration of the countless possibilities of working with this approach, and much remains to be done. Most notably, the visual factor of the composite audiovisual output object was not the primary focus of this study and it is somewhat underdeveloped at this stage; there are many sophisticated techniques one could use to this end (see Section 3.1) and we endeavour to undertake more experimental investigations on the visual front. Regarding the sound species, original material would be preferable to samples collected over the internet. Moreover, it would be interesting to capitalise on the possibilities of 3D audio, by using available field recordings of this type [42] or record some anew ourselves. Further experimentation with more sophisticated sound processes, *e.g.* multi-source audio morphing, is also planned. There are plans to implement a real-time interactive version of audiovisual surrogacy for dance performance, however, this would pose challenges as for online feature projection and would require more allowance for default parameters or alternative strategies.

On a different front, much could be done regarding gesture or emotion recognition. The mocap data used in this experiment are emotion-labeled, lending itself to machine learning techniques to add a semantic layer to our framework. This would make the model strongly dependent on the training data, and recording a bespoke training set featuring specific dance collaborators would be advisable, if necessary.

8. REFERENCES

[1] F. Grond and T. Hermann, "Aesthetic strategies in sonification," *AI and Society*, vol. 27, no. 2, pp. 213–222, 2012.

[2] S. Barrass, "The aesthetic turn in sonification towards a social and cultural medium," *AI & SOCIETY*, vol. 27, pp. 177–181, 2012.

[3] D. Smalley, "Spectromorphology: explaining sound-shapes," *Organised Sound*, vol. 2, no. 2, p. 107126, 1997.

[4] A. McArthur and S. Kalonaris, "Foregrounding expectation: Spatial sound-image composition in virtual environments," *The New Soundtrack*, vol. 7, no. 2, pp. 95–110, 2017.

[5] P. Vickers and B. Hogg, "Sonification Abstraite/Sonification Concrète: An 'Aesthetic Perspective Space' for Classifying Auditory Displays in the Ars Musica Domain," *arXiv e-prints*, p. arXiv:1311.5426, Nov. 2013.

[6] J. G. Neuhoff, "Is sonification doomed to fail?" in *International Conference on Auditory Display*, Northumbria University, 2019, pp. 327–330.

[7] C. Ziegler, "William Forsythe's improvisation technologies and beyond: A short design history of digital dance transmission projects on CD-ROM and DVD-ROM, 1994-2011," in *Transmission in Motion: The Technologizing of Dance*. Taylor and Francis, 2016, pp. 41–51.

[8] K. Lee, "Misplaced Euphoria:Developing Kinesthetic Empathy Through Interactive Performance," in *Proceedings of the 25th International Symposium on Electronic Art (ISEA19)*, 2019, pp. 616–617.

[9] J. M. Moura, N. Barros, and P. Ferreira-Lopes, "From real to virtual embodied performance - a case study between dance and technology," in *Proceedings of the 25th International Symposium on Electronic Art (ISEA19)*, 2019, pp. 370–377.

[10] O. Alemi, P. Pasquier, and C. Shaw, "Mova: Interactive movement analytics platform," in *Proceedings of the 2014 International Workshop on Movement and Computing*, ser. MOCO 14. New York, NY, USA: Association for Computing Machinery, 2014, p. 3742. [Online]. Available: https://doi.org/10.1145/2617995.2618002

[11] R. Cisneros, K. Wood, S. Whatley, M. Buccoli, M. Zanoni, and A. Sarti, "Virtual reality and choreographic practice: The potential for new creative methods," *Body, Space and Technology*, vol. 18, no. 1, pp. 1–32, 3 2019.

[12] J. E. Kyprianidis, J. Collomosse, T. Wang, and T. Isenberg, "State of the "art: A taxonomy of artistic stylization techniques for images and video," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 5, p. 866885, May 2013. [Online]. Available: https://doi.org/10.1109/TVCG.2012.160

[13] "[mViz] Human Movement Visualizations," https://mviz.omid.al/, accessed: 2020-05-16.

[14] D. Rokeby, *The Construction of Experience : Interface as Content*. ACM Press/Addison-Wesley Publishing Company, 1998.

[15] A. Camurri, B. Mazzarino, M. Ricchetti, R. Timmers, and G. Volpe, "Multimodal analysis of expressive gesture in music and dance performances," in *Gesture-Based Communication in Human-Computer Interaction*, A. Camurri and

G. Volpe, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 20–39.

[16] A. Camurri, G. D. Poli, A. Friberg, M. Leman, and G. Volpe, "The MEGA project: Analysis and synthesis of multisensory expressive gesture in performing art applications," *Journal of New Music Research*, vol. 34, no. 1, pp. 5–21, mar 2005.

[17] A. Bergsland and R. Wechsler, "Composing interactive dance pieces for the MotionComposer, a device for persons with disabilities," in *15th International Conference on New Interfaces for Musical Expression, NIME 2015, Baton Rouge, Louisiana, USA, May 31 - June 3, 2015*, E. Berdahl and J. T. Allison, Eds., 2015, pp. 20–23.

[18] M. Jeon, M. T. Smith, J. W. Walker, and S. A. Kuhl, "Constructing the immersive interactive sonification platform (iISoP)," in *Distributed, Ambient, and Pervasive Interactions*, N. Streitz and P. Markopoulos, Eds. Cham: Springer International Publishing, 2014, pp. 337–348.

[19] F. Dombois, "Using audification in planetary seismology," in *Proceedings of the 8th International Conference on Auditory Display (ICAD 2001)*, 2001, pp. 227–230.

[20] D. Worrall, *Parameter-Mapping Sonification of Tick-Data*. Cham: Springer International Publishing, 2019, pp. 237–252.

[21] T. Hermann and H. Ritter, "Sound and meaning in auditory data display," *Proceedings of the IEEE*, vol. 92, no. 4, pp. 730–741, April 2004.

[22] T. Hermann, "Wave space sonification," in *Proceedings of the 24th International Conference on Auditory Display (ICAD 2018)*, Michigan, US, 2018.

[23] C. Rutz, M.-C. Loretto, A. E. Bates, S. C. Davidson, C. M. Duarte, W. Jetz, M. Johnson, A. Kato, R. Kays, T. Mueller, R. B. Primack, Y. Ropert-Coudert, M. A. Tucker, M. Wikelski, and F. Cagnacci, "Covid-19 lockdown allows researchers to quantify the effects of human activity on wildlife," *Natural Ecology and Evolution*, vol. 4, pp. 1156–1159, 2020. [Online]. Available: https://doi.org/10.1038/s41559-020-1237-z

[24] P. M. Forster, H. I. Forster, M. J. Evans, M. J. Gidden, C. D. Jones, C. A. Keller, R. D. Lamboll, C. L. Quéré, J. Rogelj, D. Rosen, C.-F. Schleussner, T. B. Richardson, C. J. Smith, and S. T. Turnock, "Current and future global climate impacts resulting from covid-19," *Nature Climate Change*, vol. 10, pp. 913–919, 2020. [Online]. Available: https://doi.org/10.1038/s41558-020-0883-0

[25] T. Rume and S. D.-U. Islam, "Environmental effects of covid-19 pandemic and potential strategies of sustainability," *Heliyon*, vol. 6, no. 9, p. e04965, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405844020318089

[26] T. Lecocq, S. P. Hicks, K. Van Noten, K. van Wijk, P. Koelemeijer, R. S. M. De Plaen, F. Massin, G. Hillers, R. E. Anthony, M.-T. Apoloner, M. Arroyo-Solórzano, J. D. Assink, P. Büyükakpınar, A. Cannata, F. Cannavo, S. Carrasco, C. Caudron, E. J. Chaves, D. G. Cornwell, D. Craig, O. F. C. den Ouden, J. Diaz, S. Donner, C. P. Evangelidis, L. Evers, B. Fauville, G. A. Fernandez, D. Giannopoulos, S. J. Gibbons, T. Girona, B. Grecu, M. Grunberg, G. Hetényi, A. Horleston, A. Inza, J. C. E. Irving, M. Jamalreyhani, A. Kafka, M. R. Koymans, C. R. Labedz, E. Larose, N. J. Lindsey, M. McKinnon, T. Megies, M. S. Miller, W. Minarik, L. Moresi, V. H. Márquez-Ramírez, M. Möllhoff, I. M. Nesbitt, S. Niyogi, J. Ojeda, A. Oth, S. Proud, J. Pulli, L. Retailleau, A. E. Rintamäki, C. Satriano, M. K. Savage, S. Shani-Kadmiel, R. Sleeman, E. Sokos, K. Stammler, A. E. Stott, S. Subedi, M. B. Sørensen, T. Taira, M. Tapia, F. Turhan, B. van der Pluijm, M. Vanstone, J. Vergne, T. A. T. Vuorinen, T. Warren, J. Wassermann, and H. Xiao, "Global quieting of high-frequency seismic noise due to covid-19 pandemic lockdown measures," *Science*, vol. 369, no. 6509, pp. 1338–1343, 2020. [Online]. Available: https://science.sciencemag.org/content/369/6509/1338

[27] B. Basu, E. Murphy, A. Molter, A. Sarkar Basu, S. Sannigrahi, M. Belmonte, and F. Pilla, "Investigating changes in noise pollution due to the covid-19 lockdown: The case of dublin, ireland," *Sustainable Cities and Society*, vol. 65, p. 102597, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670720308143

[28] R. Schafer, *The Soundscape: Our Sonic Environment and the Tuning of the World*. Inner Traditions/Bear, 1993.

[29] B. Truax, *Handbook for Acoustic Ecology*. Vancouver, Burnaby, B.C.: A.R.C. Publications, 1978.

[30] J. Bronstein, *Mutualism*. Oxford Press University, 2015.

[31] M. B. Dhakne and A. B. Munde, "Stability analysis of mutualistic interactions among three species with limited resources for first species and unlimited resources for second and third species," *Differential Equations and Dynamical Systems*, vol. 20, no. 4, pp. 405–414, Oct 2012.

[32] E. Stone, "Separating the noise from the noise: A finding in support of the niche hypothesis, that birds are influenced by human-induced noise in natural habitats," *Anthrozoös*, vol. 13, no. 4, pp. 225–231, 2000.

[33] R. M. May, "Models for two interacting populations," in *Theoretical Ecology. Principles and Applications*, R. M. May, Ed. Sunderland, Massachusetts: Sinauer Associates, 1981, pp. 78–104.

[34] C. S. Holling, "The components of predation as revealed by a study of small-mammal predation of the european pine sawfly," *The Canadian Entomologist*, vol. 91, no. 5, p. 293320, 1959.

[35] E. P. Kasten, S. H. Gage, J. Fox, and W. Joo, "The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology," *Ecological Informatics*, vol. 12, pp. 50–67, 2012.

[36] habbis92, "Picadilly Circus Ambience," https://freesound.org/people/habbis92/sounds/240233/, June 2014.

[37] NightVoice, "Forest Birds," https://freesound.org/people/NightVoice/sounds/123999/, July 2011.

[38] ciccarelli, "Ocean Waves," https://freesound.org/people/ciccarelli/sounds/132736/, October 2011.

[39] A. B. Godbehere, A. Matsukawa, and K. Goldberg, "Visual tracking of human visitors under variable-lighting conditions for a responsive audio art installation," in *2012 American Control Conference (ACC)*. IEEE, jun 2012.

[40] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of Imaging Understanding Workshop*, 1981, pp. 121–130.

[41] R. Shusterman, "The Pragmatist Aesthetics of Richard Shusterman: A Conversation - Interviewed by Günter Leypoldt," *Zeitschrift für Anglistik und Amerikanistik: A Quarterly of Language, Literature, and Culture*, vol. 48, no. 1, pp. 57–71, 2000, online; accessed 27 February 2020.

[42] M. C. Green and D. Murphy, "Eigenscape: A database of spatial acoustic scene recordings," *Applied Sciences*, vol. 7, no. 11, 2017.