Editorial

Environmental Sound Synthesis, Processing, and Retrieval

Andrea Valle

CIRMA, Università di Torino, via Sant'Ottavio, 20, 10124 Torino, Italy

Correspondence should be addressed to Andrea Valle, andrea.valle@unito.it

This special issue of the EURASIP Journal on Audio, Speech and Music Processing is dedicated to Environmental Sound Synthesis, Processing, and Retrieval. It aims at targeting the multifaceted area of research devoted to the complex relation between environment and sound, a relation that still needs to be investigated. Indeed, we are literally immersed into sound: as Handel says, "Listening puts me in the world" ([1], xi). In this sense, a phenomenology of listening, delving deeply into the philosophical and psychological aspects of sound perception is mandatory in order to clearly understand the specific features of "auditory events" with respect to other perceptual modalities. With a classic example, while vision is in some sense always external to our body (emphasizing the separation between subject and object in perception), sound on the contrary resonates through our body, literally embodying the information that it carries on as we vibrate through the world we are surrounded/placed into. That is, "Listening is centripetal; it pulls you into the world. Looking is centrifugal, it separates you from the world" ([1], xi). Even if embracing a technologically oriented perspective, these philosophical and psychological aspects are not to be omitted while researching on environmental sounds. Through a historical tendency which origins date back to 20 years ago, sound is becoming more and more relevant in our media environment. As an example, one can think about sound systems in cinema, now a standard and pervasive solution on the market. Another example is new complex, multimodal, integrated displays now pervasively built into portable devices. Not by chance, the first edition of ICAD, the International Conference on Auditory Displays dates back to 1992, and from 2000 it is held on an annual basis. But, in order to effectively exploit sound, we have to avoid the risk to simply borrow principles and models from vision and to adapt them to listening, without a real

understanding of specific perceptual features of the audible domain. Indeed, this applies to sound in general, but is particularly relevant for sound materials that have not been extensively studied by cultural practices with a long tradition, as it typically happens with music and speech.

So, what is environmental sound? In some sense, sound is necessarily environmental as it is strictly coupled with its physical medium (including the listener). I have already anticipated a definition ex negativo that better specifies our field of interest: environmental sound is neither music nor language. Indeed, such a definition is at the same time too strict and too broad, as it supposes that there are three distinct realms of sound, while theoretical researches and productive practices have shown that these three aspects of sound perception/production are deeply intermingled. It can be noted that such a definition has been historically at the basis of sound at cinema, where "music", "voice", and "sound" have always been treated in specific ways [2, 3], and thus it is, in some sense, "classic", as it simply establishes that "sound" (here intended as the fictional acoustic environmental scene) has a residual nature with respect to speech and music. Indeed, in contemporary cinema too this categorization is becoming more and more unsatisfactory as sound is receiving an increasing attention. That is, the internal complexity of this third category—"sound"—is increasingly emerging in cinema studies and practices thanks to technological developments. This is not an accidental aspect: from the 19th century going on through 20th (and 21th century), technology is continuously stimulating the research on audio, radically challenging different contexts of perception and production [4, 5].

In order to deal with this complexity, it is possible to cite at least some fields and authors that have provided a general frame for the understanding of (environmental)

sound. First of all, the notion of "sound object" has been proposed by Schaeffer [6] with the specific goal to describe all possible sounds. Even if problematic in many respects [7], Schaeffer's "morphotypology" is still unsurpassed, as it is the only theoretical framework trying to be at the same time analytical and exhaustive. It could be noted that Schaeffer's perspective is deeply technologically rooted, as the French author started his journey into sounds thanks to the possibility, provided by recording, of listening again and again to the same sound. More, the theoretical framework by Schaeffer was originally aimed at providing a conceptual tool for the organization of sound objects in music composition, thus linking listening practice to sound manipulation. Partly moving from Schaeffer, R. Murray Schafer firstly introduced (or, at least, theoretically discussed) the term "soundscape" in his famous book *The Tuning of the World* [8]. Now an ubiquitous term, soundscape at least covers three different domains and relative applications (eco/anthropology, music/sound design, architecture/urban planning, [9]). Again, the interest in soundscape emerges from the technological possibility of field recording and of accurate, iterated, analysis of the obtained soundscape through editing and playback. Soundscape studies, in the context of acoustic ecology [10], have shown the complexity, variety, and internal articulation of acoustic environments coming from all the world, showing many aspects that were completely neglected before. From Murray Schafer, the diffusion of the term has continuously increased, and the relevance of soundscape in the actual "mediascape" cannot be disputed, as currently the concept of soundscape plays a pivotal role at the crossing of many sound-related fields, ranging from multimedia [11] to psychoacoustics [12], from working environment studies [13] to urban planning [14], from game design [15, 16] to virtual reality [17], from data sonification [18] to ubiquitous computing [19, 20]: soundscape is a fundamental notion for acoustic design [21, 22], electroacoustic composition [23], and auditory display studies [24]. The integration of soundscape in a landscape documentation/simulation is crucial in order to ensure a believable experience in human-computer interaction [25]. Moving on in this fast run through relevant approaches to environmental sound, "everyday listening" has been proposed by Gaver [26] as a specific modality of listening to sound, mainly based on a re-construction of some features of the sound sources. As well said by Gygi et al. in this volume: "Although what Gaver termed "everyday listening" is a frequent activity, the nature of the experience has been remarkably underscrutinized, both in common discourse and in the scientific literature". Listening to every day sound also requires specific perceptual strategies, that cannot be described in the usual theoretical framework of psychoacoustics: in this sense, Bregman's summa [27] has established the notion of "Auditory Scene Analysis" (ASA) as a pivotal psychological basis for the perception of complex sound mixtures like the ones we experience in "natural" environments (even if highly anthropized, e.g., a city). Finally, the Sounding Object project has pioneered the study and the application of an ecological approach to sound and perception to the design and production of

interactive auditory displays based on physical models of audio production/perception [25].

Following the threads I have tried to individuate in the previous paragraphs, in this issue we have selected seven contributions that indeed demonstrate the multifaceted nature of environmental sound studies. Quite approximatively, they can be grouped into three areas. The first subset includes the papers by M. Takada et al. and L. Wang et al. Takada and colleagues propose a research on the relation between onomatopoeia and sound. Indeed, the use of voice to reproduce sounds allows to study the way sounds are perceived, represented and reproduced by the subjects. It might be assumed that these features are particularly relevant for example, in auditory display applications as they can be embodied directly by the user and easily shared among other users (as they can be easily reproduced through the voice). From a strict signal processing perspective rather than from a psychological/semiotic one, the work by Wang and colleagues discusses a method for improving source separation in reverberant environments. Indeed, the contribution deals with a typical and crucial problem of the auditory domain, the fact that, to speak with a visual metaphor, "sound is transparent" ([27], 619). In this sense, it can be seen as a contribution to Computational ASA (CASA, [28]), a field that aims at computationally implementing Bregman's approach for automated perceptual analysis of acoustic environments. The paper by B. Gygi and V. Shafiro discusses the creation of a large database of environmental sounds. Mainly aimed at providing researchers a tool for the investigation of ecologically based sounds, it shares with the following two papers (with which it can be grouped) the interest into large collections of sounds, indeed a major topic in actual research, as social networking increasingly allows users to provide and share audio contents. The database proposed by B. Gygi and V. Shafiro still implements a topdown perspective, as categories related to ecological features of sounds necessarily have to be under the control of the database managers in order to be effective. The papers by G. Roma et al. and by G. Wichern et al. both deal with the problem of exploring large databases of sounds. While the contribution by G. Roma and fellow researchers is mainly focused on automatic classification of sounds based on acoustic ecology's principles, G. Wichern et al. 's contribution is characterized by an explicit ontological focus. An interesting point lies in the fact that both papers study, as one of their test beds, the user-contributed database of the Freesound project, thus providing the readers the possibility of comparing the proposed approaches on the same experimental situation.

Finally, the two papers by R. Nordhal and Menzies both concern the integration of audio into virtual reality applications in order to enhance user experience. In both cases, the main problem is to provide ecologically-based sound models, thus allowing a more immersive and plausible experience to the users. Not by chance, they both share the use of physical models of sound synthesis, a very promising approach pioneered by the aforementioned Sounding Object project. Apart by the specific solutions proposed by the authors, the reader's perspective is enriched also by the different focus

of the two contributions. While Menzies is mainly oriented toward production (that is, sound designers), Nordhal takes into account the evaluations by final users, in order to compare physically based synthesized sounds and recorded ones.

Andrea Valle

References

[1] S. Handel, *Listening: An Introduction to the Perception of Auditory Events*, The MIT Press, Cambridge, Mass, USA, 1989.

[2] V. LoBrutto, *Sound-on-Film: Interviews with Creators of Film Sound*, Praeger, Westport, Conn, USA, 1994.

[3] M. Chion, *L'audiovision. Son et Image au Cinéma*, Nathan, Paris, France, 1990.

[4] J. Sterne, *The Audible Past*, Duke University Press, Durham, UK, 2003.

[5] D. Kahn, *Noise, Water, Meat. A History of Sound in the Arts*, The MIT Press, Cambridge, Mass, USA, 1999.

[6] P. Schaeffer, *Traité des Objets Musicaux*, Seuil, Paris, France, 1966.

[7] A. Valle, *Preliminari ad una semiotica dell'udibile*, Ph.D. thesis, Università di Bologna, Bologna, Italy, 2004.

[8] R. Murray Schafer, *The Tuning of the World*, Knopf, New York, NY, USA, 1977.

[9] A. Valle, V. Lombardo, and M. Schirosa, "Simulating the soundscape through an analysis/resynthesis methodology," in *Proceedings of the 6th International Symposium on CMMR/ICAD*, S. Ystad, M. Aramaki, R. Kronland-Martinet, and K. Jensen, Eds., vol. 5954 of *Lecture Notes in Computer Science*, pp. 330–357, Springer, 2009.

[10] B. Truax, *Acoustic Communication*, Greenwood, Westport, Conn, USA, 1984.

[11] M. Burtner, "Ecoacoustic and shamanic technologies for multimedia composition and performance," *Organised Sound*, vol. 10, no. 1, pp. 3–19, 2005.

[12] F. Fontana, D. Rocchesso, and L. Ottaviani, "A structural approach to distance ren-dering in personal auditory displays," in *Proceedings of the International Conference on Multimodal Interfaces (ICMI '02)*, Pittsburgh, Pa, USA, October 2002.

[13] I. McGregor, A. Crerar, D. Benyon, and C. Macaulay, "Soun-fields and soundscapes: reifying auditory communities," in *Proceedings of the International Conference on Auditory Display*, Kyoto, Japan, July 2002.

[14] B. U. Rubin, "Audible information design in the new york city subway system: a case study," in *Proceedings of the International Conference on Auditory Display*, Glasgow, UK, 1998.

[15] M. Droumeva and R. Wakkary, "The role of participatory workshops in investigating narrative and sound ecologies in the design of an ambient intelligence audio display," in *Proceedings of the 12th International Conference on Auditory Display*, London, UK, 2006.

[16] J. Friberg and D. Gardenfors, "Audio games: new perspectives on game audio," in *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*, pp. 148–154, ACM Press, New York, NY, USA, 2004.

[17] S. Serafin, "Sound design to enhance presence in photorealistic virtual reality," in *Proceedings of the International Conference on Auditory Display*, Sidney, Australia, July 2004.

[18] T. Hermann, P. Meinicke, and H. Ritter, "Principal curve sonification," in *Proceedings of International Conference on Auditory Display*, 2000.

[19] A. Butz and R. Jung, "Seamless user notification in ambient soundscapes," in *Proceedings of the International Conference on Intelligent User Interfaces (IUI '05)*, pp. 320–322, New York, NY, USA, January 2005.

[20] F. Kilander and P. Lonnqvist, "A whisper in the woods—an ambient soundscape for peripheral awareness of remote processes," in *Proceedings of the International Conference on Auditory Display*, Kyoto, Japan, July 2002.

[21] A. A. VV, "The tech issue ... to be continued," *Soundscape*, vol. 3, no. 1, 2002.

[22] A. A. VV, "Acoustic design," *Soundscape*, vol. 5, no. 1, p. 19, 2004.

[23] H. Westerkamp, "Linking soundscape composition and acoustic ecology," *Organised Sound*, vol. 7, no. 1, 2002.

[24] B. S. Mauney and B. N. Walker, "Designing systems for the creation and evaluation of dynamic peripheral soundscapes: a usability study," in *Proceedings of the 48th Annual Meeting on Human Factors and Ergonomics Society*, New Orleans, La, USA, 2004.

[25] D. Rocchesso and F. Fontana, Eds., *The Sounding Object*, Edizioni di Mondo Estremo, Firenze, Italy, 2003.

[26] W. Gaver, "What in the world do we hear? an ecological approach to auditory event perception," *Ecological Psychology*, vol. 5, no. 1, pp. 1–29, 1993.

[27] A. Bregman, *Auditory Scene Analysis. The Perceptual Organization of Sound*, The MIT Press, Cambridge, Mass, USA, 1990.

[28] D. Wang and G. J. Brown, Eds., *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*, Wiley-IEEE Press, New York, NY, USA, 2006.