New Directions in Soundscape-based Sound-Art: Hybridising Autoethnography with Computational Analysis

Natasha Barrett
Norwegian Academy of Music
natasha.barrett@nmh.no

Abstract

Real-world sound environments feature throughout sound-art. Fixed media soundscape compositions, electroacoustic music addressing eco-structuralism, site-specific sound installations, performers that use the sound landscape as their stage, and architects and city-planners that design sound-sensitive structures are increasingly common.

Despite this intense interest in soundscape-based sound-art, the creative approach developed in the early years of recording generally persists in current practice: that is, a process guided by experiential self-reflection. This approach has served for many decades, yet tends towards a closed loop limited by the ear's and body's variable awareness.

I propose a practice that hybridises autoethnography with computational analysis to overcome the limitations of experiential personal reflection, without denying the value of the artist as the creative force. In this paper I have chosen to focus on the affordances of spatial sound in outdoor environments framed in a musical perspective. The work draws on the perceptual theories of Gibson and Ingold, and on the possibilities that computational tools such as 3D sound decomposition and music information retrieval have to offer. To illustrate the hybridisation approach, I present non-technical examples from my current experiments. The work is part of the Reconfiguring the Landscape project funded by the Norwegian Artistic Research Council.

1. Introduction

For many composers the allure of the soundscape is omnipresent in artistic practice. It may take the form of a small detail that leads the listener inside an otherwise abstract sound-word, a keynote sound that may evoke an outdoor narrative, the inference of environmental sound archetypes through the behaviour of texture or motion, the sonification of data captured from a real-world system, or an explicit phonographic portrayal of the real world. Sound-walks as both composition and participatory action also lend self-narrative to the subject, and ubiquitous technologies transform the traditional sound-walk into an interactive experience using GPS mobile phone apps such as Echoes. Moreover, 3D audio technologies such as Ambisonics and Wavefield synthesis are stimulating a new dimensionality in the capture and communication of our compositional ideas.

Despite this diversity, the creative approach developed in the early years of recording generally persists in current practice: that is, a process guided by experiential self-reflection. This approach has served soundscape-based sound-art for many decades, yet by listening to the proliferation of soundscape compositions available online – and how similar many of them appear to be - could be trending towards a closed-loop. Are our creations more profiting from the call to document immediate soundscapes in a time of rapid environmental change, while artistic works are themselves tending to be short-lived?

In this paper I propose a practice that hybridises autoethnography with computational analysis to overcome the limitations of experiential personal reflection, without denying the value of the artist. Although the proposal can embrace many facets of the soundscape, here I will focus on 3D sound and discuss how the theories from Gibson and Ingold can be transferred into creative practice.

In 2012 Barry Truax (Truax 2012) wrote, "A purely aesthetic approach may be problematic when artists wish to deal with the external world as part of their work". Truax was referring to composers grappling with non-musical sounds and their relationship to the environment and acoustics. Nowadays, technical and scientific knowledge proliferates via education programs, self-learning internet resources, open access publications and even social media groups, such that compositional aesthetics, technical knowledge and an understanding of outdoor sound are highly integrated. Yet there are other reasons why a purely aesthetic approach may be problematic. Westerkamp writes (Westerkamp 2002), "Each new recording will create a totally new piece… a specific moment and place can contain all the ingredients out of which a meaningful language can emerge... and it is that material, and not some predetermined musical structure or context, that will contribute significantly to a work's unique character." Although this approach has resulted in some important repertoire works such as Westerkamp's own, its sustainability seems contingent on one or more factors: (a) composer and listener are similarly engaged culturally, socially, politically or spiritually, (b) that the chosen material is sufficiently unique to hold interest beyond the first few listenings, (c) that the specific moment and place is special in the ears of the listener, or that the composer is able to communicate what is special.

2. Background concepts

2.1. Autoethnography

The term autoethnography originally referred to a study of cultural phenomena where researchers were active members and provided personal and subjective insights. More recently, autoethnography has been used to describe a creative approach to environmentally driven electroacoustic music where qualitative arguments are guided by self-reflection, and where personal experience and beliefs are in focus. This approach has of course existed for as long as soundscape composition, and autoethnography is simply a useful label.

2.2. Computational analysis

In this article I have chosen the general term ‘computational analysis’ rather than the specific term ‘artificial intelligence’. The computational analysis of audio is about extracting information with computer tools. Amongst the many applications, it can be used to systematise audio experiences that may be intuitive yet difficult to quantify via aural approaches, or to highlight potentially interesting information which may go unnoticed. Although machines struggle with listening skills that humans take for granted, methods for music information retrieval (MIR) and machine learning (ML) are increasingly tailored to the challenges of complex audio signals. These techniques have been applied extensively in bioacoustics and computational musicology for some time for example (Bellisario, et al. 2019).

2.3. Time and space

The temporality of the sound landscape is ‘long play’: although some sounds are persistent, interesting events are often intermittent and distributed many minutes, hours, days or even months apart. Experientially, over a long period of time information accumulates and collapses into a familiar memory and a set of expectations which we take for granted. The real soundscape is also incredibly spatial, where any one listening position may reveal a new understanding of the scene, or a gust of wind may carry sounds towards or away from our ears.

However, the ear’s and body’s short-term experiential awareness limits the best of our capacity to harness these features in composition. As the composer, I am the ‘filter’: my decisions determine the temporal collapse, decide where to place the microphone, or make the decision to wear clothing that is inappropriate for the weather conditions. Fluctuating background noise, climatic conditions and our state of mind or body result in capricious day to day judgements. How can we reconcile ourselves as ‘capricious filters’ in search of a more stable ‘truth’?

3. The significance of ecological theory

Gibson’s ecological theory proposes how perception, the organism and the environment are connected. One of the core ideas is that objects and events 'afford' or furnish possibilities for action, and that this action is coloured by the relationship between individual needs and the already structured environment. Windsor’s elaboration of these ideas in acousmatic theory (Windsor 1995) place what was originally routed in visual experience in a sounding context, elaborating that our perception of the music is directly sensitive to the structured world that we inhabit. This is a useful springboard into a more extensive discussion.

In "Against Soundscape" (Ingold, Against soundscape 2007) Ingold highlights some fair problems for soundscape composition. Besides a general critique of terminology, he questions how we artificially slice up the environment 'along the lines of the sensory pathways by which we enter into it'. Instead he argues that 'the world we perceive is the same world, whatever path we take'. Our normal behaviour as composer sound-gathers, particularly when we are tourists to a sonic location, confirms Ingold critique.

'Against Space: Place, Movement, Knowledge' (Ingold, Against space: Place, movement, knowledge 2009), influenced by Gibson's work, argues that 'place' is not a space occupied by things, but rather a woven habitat that we understand though motion: existence unfolds, affected by encounters with the surroundings where knowledge is gained through movement that unravels 'the meshwork'.

Ingold is emphasising Gibson's theory of 'reversible occlusion' (Gibson, 1979: 198, in Ingold, 2002: 238): 'one knows the way in terms of the specific order in which the surfaces of the environment come into or pass out of sight as one proceeds along a path', and this perception changes as we move. Wonderfully put, he says, 'as you turn a corner a new set of vistas come into view while the previous ones disappear.… It is through the ordering of these vista that the structure of the environment is progressively disclosed to the moving observer, such that he or she can eventually perceive it from everywhere at once.' These ideas were developed with visual perception in mind, yet can easily be applied to auditory perception: (a) the emphasis on spatial movement leads to us hear the acoustic reflections change, resonate and diffract, and inform us of a spatial totality beyond that delivered by our visual senses, (b) although sound travels and wraps around us differently to reflected light, vistas appearing and receding are nevertheless apparent, for example through acoustic masking, baffles and enclosures, (c) wayfaring by sound may result in a different path to wayfaring by sight, by the principles of information gathering remain.

3.1. Where do these ideas lead in practice?

How may we overcome problems of transient participation, or of information being acoustically, perceptually or cognitively masked in the moment we listen? Can we reveal information, that for whatever reason we fail to be receptive to? How can we 'hear' from everywhere at once over a time-span greater than our patience and memory can process? These questions imply a practice that incorporates long 3D sound recordings capturing a high spatial and spectral precision, for these recordings to be made from listening locations that together capture the totality of the area of interest, and 3D impulse response recording to reveal acoustic fingerprints. The extent of the materials calls for new methods for exploring content and meaning.

4. Hybridising autoethnography and computational analysis

Artistic research on the sound landscape is common. Examples that include analytical methods that allow a transfer of procedural knowledge are few. In Cerwén's application of what he calls "analytical autoethnography" (Cerwén, Kreutzfeldt og Wingren 2017) he explains how sonic experiences are discussed in relation to previous research and supported by data collected on site. The findings are then used to gain insights for a theoretical framework that he calls Soundscape Actions. Applied to design, they can be thought of as a template with which to understand a new sound space and to guide an autoethnographic approach, allowing opportunities for both subjectivity and consistency. This template of Soundscape Actions is

static. Computational methods are more interesting if they are in some way 'reactive', such as being able to train on the global context.

4.1. 3D sound and music information retrieval

In 2017, while working on the composition The Weathered Piano (Barrett, The Weathered Piano u.d.) I analysed the spatial information in first-order 3D Ambisonics recordings and used this information to form structures in the composition. Part of the method involved applying simple MIR algorithms for segmentation and spectral-temporal analysis, and analysing for patterns and trends. At that time, MIR algorithms operated only on mono sources and I integrated some simple methods for spatial analysis. In 2019 I then developed a more accurate approach to 3D sound analysis by drawing together my technical knowledge of Ambisonics and my artistic practice in the same field (Barrett, Deepening presence: probing the hidden artefacts of everyday soundscapes 2020). In summary, the method decomposes high resolution 3D signals from a microphone such as the MhAcoustics EM32 into spatial objects, by applying beam-forming and content analysis. I have since improved my method by analysing for directional energy in frequency bands, and by including audio descriptors such that the system can more successfully distinguish between spatial objects. Other computational methods for separating direct and diffuse sources exist, but few have been implemented in software. (The Compass tracker (McCormack, et al. 2021) shows some success in decomposing spatial scenes, and other researchers are developing new methods for the spatial analysis of 3D room impulse responses). Sound-field decomposition is the first stage in my proposed method. The output generates spatial objects in *real-time* – the significance of which will be clear shortly. This is the starting point for a process that probes time, space, acoustics, interaction and motion, auditory perception, and to some extent sound identity.

4.2. An iterative process that converges on a method

How can we know what to look for in sound materials that exceed our perceptual capacity? A blind analysis using commonly used audio descriptors for spectral energy over a specified temporal duration will likely reveal some meaningful principle features, but analysing the complete 3D spherical sound picture makes the task more complex. Thresholding the data allows for selective focus, but knowledge of the data range is needed before a threshold can be set. We should further consider if we are interested in the content of one recording, a collection of recordings from one location, or a collection of locations. An unsupervised method involving MIR and ML could reveal the values of scales and thresholds. But first it is useful to apply the power of autoethnography.

Figure 1 suggests one approach to this problem. On the left side is everything to do with sound capture: recording, basic sound-field decomposition and 3D impulse responses. The centre panel explains the analysis in two parts. On the left I present two possibilities: unsupervised analysis using audio descriptors alone (currently still in an experimental phase due to the complexities involved), and a selective analysis guided by a memory of the original

environment and by the audition of short extracts of recordings. Data ranges and thresholds also need to be set, and then complete system verified as suitable for the source. To do this, the system is implemented in a real-time MaxMSP patch. Controllers are mapped to variables and an iterative process of real-time listening and adjustment continues until key features in the source begin to emerge. This solution can then be used in the unsupervised analysis of larger amounts of audio (in my current work this process is however still running in real-time). The right-hand section of the figure shows the output. This includes movement archetypes, information about space, time and spectrum, other information that can be used as control values for various sound processing techniques, and a timestamp indicating the location of key sounds in the original file.

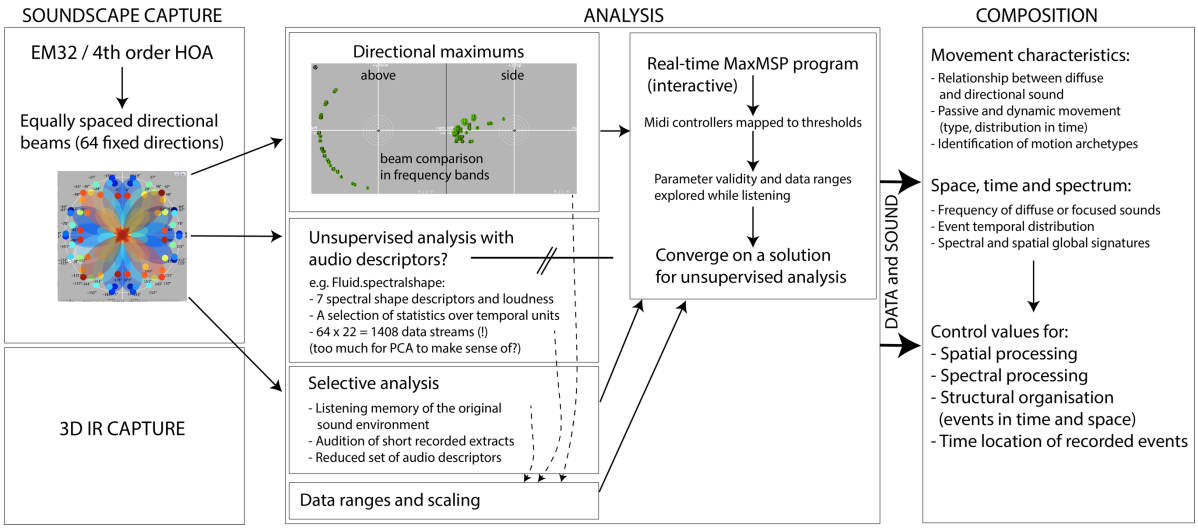

Figure 1. Sound capture and analysis, converging on a method.

4.3. Invariants and attunement

So far, I have suggested how we can explore space and time over scales beyond our experiential capacity and embrace some of the interesting features of ecological theory in compositional practice. Working in the acousmatic tradition, it would be intriguing to explore how Schaeffer's ideas on listening can be harmonised with this process. In some ways Gibson and Schaeffer can be regarded as opposites: one developed theories of a meshwork, the other used the phenomenological lens to separate sounding matter. Ignold draws attention to what Gibson calls invariants: '…sensations do not, as such, constitute the data for perception (Gibson 1979: 55). Rather, what the perceiver looks for are constants underlying the continuous modulation of the sensory array as one moves from place to place'. He says that it is the, "...invariants that underlie this transformation, and not the momentary patterns of stimulation themselves".

In Schaeffer's discussions on the modes of listening (Schaeffer 1966 / 2017) he draws attention to the problem of attunement, arguing that we are not always correctly attuned, even though the information, upon which 'novel perceptions arise', is available to anyone attuned to pick it up. Although Schaeffer looked towards the intrinsic features of sound revealed through reduced

listening, when we perceive the environment we are not only listening causally but also enjoying movement, changes in spectrum and resonance, texture and articulation. The method proposed above allows an investigation of both invariants and the objects of attunement, each characteristically reflected in the data.

5. Discussion

As an artist, I find it compelling that the results of computational analysis could influence a compositional genre. In my own work I have so far applied these methods in three site specific compositions on the sound landscape, each of which are part of the Reconfiguring the Landscape Research project based at the Norwegian Academy for Music in Oslo. (Documentation videos are found at (Barrett 2021). The artistic work will be documented in detail in forthcoming publications). Over the course of these three projects the analytical methods have remained constant, yet on each iteration I am more satisfied with the artistic output. What has changed? My own understanding and handling of the technical aspects has indeed improved, but moreover, the roll of autoethnography and compositional decisions are clearer: how I relate to the results of the MIR, what I decide to trust in the data, the excitement of discovery and how the information is recycled into the artistic process.

6. References

Barrett, N. 2021. 10 December. Last accessed December 10, 2021. https://www.researchcatalogue.net/shared/841fe95df9608bb1c446a72f05792646.

Barrett, N. 2020. “Deepening presence: probing the hidden artefacts of everyday soundscapes.” *Proceedings of the 15th International Conference on Audio Mostly, Association for Computing Machinery, New York, NY, USA* 77-84.

Barrett, N. s.d. “The Weathered Piano.” *Leap Seconds.*

Bellisario, K M, J VanSchaik, Z Zhao, A Gasc, H Omrani, et B C Pijanowski. 2019. “Contributions of MIR to Soundscape Ecology. Part 2: Spectral timbral analysis for discriminating soundscape components.” *Ecological Informatics* 51: 1-14.

Cerwén, G, J Kreutzfeldt, et C Wingren. 2017. “Soundscape actions: A tool for noisetreatment based on three workshopsin landscape architecture.” *Frontiers of Architectural Research* 6 (4): 504-518.

Ingold, T. 2007. “Against soundscape.” Dans *Autumn leaves: Sound and the environment in artistic practice*, 10-13.

Ingold, T. 2009. “Against space: Place, movement, knowledge.” Dans *Boundless worlds: An anthropological approach to movement* , 29-43.

McCormack, L, L Politis, A Särkkä, et V Pulkki. 2021. “Real-time tracking of multiple acoustical sources utilising Rao-Blackwellised particle filtering.” *29th European Signal Processing Conference.*

Schaeffer, Pierre. 1966 / 2017. “Traité des objets musicaux (Treatise on Musical Objects).” 69-111. Oakland, California: University of California Press.

Truax, B. 2012. “Sound, Listening and Place: The aesthetic dilemma.” *Organised Sound* 17 (03): 193-201.

Westerkamp, H. 2002. “Linking soundscape composition and acoustic ecology.” *Organised Sound* 17 (1): 51-56.

Windsor, L. 1995. *Windsor, William Luke. A perceptual approach to the description and analysis of acousmatic music. PhD diss.* City University London.