Real-time Algorithmic Timbral Spatialisation: Compositional Approaches and Techniques

Stefano Catena (stefano.catena23@gmail.com)[1]

[1]**Independent Researcher**, Milano, Italy

ABSTRACT

This paper introduces a series of techniques and approaches to algorithmic timbral spatialisation, the real-time processing of audio data and its musical organization in a binaural or multichannel listening space. The intent of the work is to explore different ways for the automated spatialisation of the audio spectrum, especially in contexts of electroacoustic and acousmatic music composition. Typical Digital Signal Processing operations and algorithms will be used in order to create and/or retrieve data for the dynamic positioning of the audio sources. Firstly, in order to describe the processes, the concept of timbral spatialisation will be introduced, describing the compositional interest of such approach. Then the different techniques for data generation will be formalized, describing their "audioparous" process, that is when information regarding musical organization is extrapolated from an audio source. The various approaches to interpretation and usage of the data will be discussed, as well as their implementation in SuperCollider.

1. INTRODUCTION

The concept of "spatiality" has been taken into consideration in the compositional practice since between the 10th and 14th century in the vocal music, particularly with the *antiphonal psalms* [1], later elaborated with the tradition of *polychoral music* [2]. Even during the classical period some works actively use the spatial component of sound ("Serenata 8 in D major for 4 orchestras, K. 286" from 1777 by W. Mozart), and later in the 19th and 20th centuries with composers such as Mahler ("Symphony No. 2") or Charles Ives ("Unanswered Question"). But it is only after the Second World War, with the introduction of electronic instruments and loudspeakers, that spatiality becomes a fundamental aspect of musical production, with some prominent composers, such as Karlheinz Stockhausen, who take great advantage from it [3]. In particular, since the 1970s, the spatial aspect has been one of the most in-depth and researched fields, [4], even with the design of special diffusion systems [5]. In recent years, with the greater accessibility and evolution of spatialisation techniques, several new concepts regarding spatiality have been introduced in the compositional practice. One can argue that spatial movement, even with little formal theory about it, has become a musical parameter just like rhythm, timbre and pitch [2]. One of the approaches to the organization of spatial parameter of sound is "timbral spatialisation" [6]. This process deconstructs a sound source into its individual spectral bands, addressing them as single point-sources, and placing them in the listening space. Timbral spatialisation then recombines the entire spectrum virtually, whether in the concert hall or in headphones, therefore not simply adding the spatial aspect at the end of the composition process, but actually re-composing the space [7]. Timbral spatialisation poses the problem of controlling each individual spectral band in space: such process can require potentially dozens of parameters, all at the same time. Some of the most used techniques for this task have been Wave Terrain Synthesis [8] or granular synthesis [9], but given the compositional nature of the process, the exploration of different algorithmic techniques for musical data generation is of relevant interest.

Sound and space are intertwined irreversibly: whenever sound is recorded, the space where the sound "happens" is recorded as well, inevitably being perceived with the spectral properties of the source [10]. A similar process of linkage can be used as framework for spatialisation, extrapolating data from audio sources in order to influence the spatial aspect of the music. Such technique can be called "audioparity", a composition mode in which musical data originate from sound. A compositional technique, consequently, is audioparous if it defines a projection between a source sound material and an outgoing musical organization [11]. For example, this compositional process has been used by composers such as Messiaen, as he integrated in his works transcriptions of singing birds [12]. With the usage of the same audio data for the spatialisation control, one can extend the concept of "audioparity" to "self-audioparity", where sound and space influence each other. We can therefore say that in "self-audioparous" techniques the musical organization is directly influenced and modified by the audio source itself.

In discussing algorithmic composition, even more when the concept of "audioparism" stands out, it is relevant to point out how human perception processes time scales differently. For example experiencing a simple sinusoid transposed to different time scale would change the perceptual results drastically, but the waveform itself would still remain the same [13]. While this is true for sonic experiences, it does not prevent the composer to cross the boundaries of time scales in order to "apply" data extracted from one temporal region to another.

In the next chapters some of these techniques and approaches will be applied to real-time incoming audio signals in order create complex spatial textures within timbral spatialisation. In particular an exploration in the use of noise algorithms, the use of FFT descriptors and finally an audio snapshot and manipulation technique to automate the creation of auditory scenes.

2. APPROACH

While spatiality can be assimilated to other musical parameter such as pitch, rhythm and timbre, it is sometimes used as a simple post-processing effect, mostly because of the lack of a properly unified and defined spatial language [14]. The formalization of techniques for the generation of spatial data, provides a framework for spatialisation that can be easily integrated inside a composer's work. This is particulary true in electroacoustic and acousmatic composition where the space and spatial experience is aesthetically central [10]. Consequently, the spatial perception of the listener is of fundamental importance: how the movements are generated, how they affect spectromorphology of sound and how the spatialisation is connected, if it is at all, to its source. Even though there isn't a solid framework which might provide a reasonably secure basis for investigating space [10], it is possible to define spatial attributes and characteristics [14] that, in themselves, define and organize the spatial scene. These characteristics are inevitably crucial for the impact on the spatial experience of the listener, and must be taken into account when formalizing procedures for automatic spatialisation.

Timbral spatialisation enables the musical exploration of sound very differently from point-source techniques. It involves the "deconstruction" of sound into spectral bands, by means of several bandpass filters with different central frequencies, allowing for compositional processes to determine how the sound will be spatially distributed for each part of the spectrum. This process is similar to typical FFT synthesis and resynthesis, and even more to analog vocoders [15]. They allow for the deconstruction and reconstruction of the sound based on its frequency content; this process has been linked to the term *spectromorphology* [10]. The concept of "deconstruction" or "decomposition" of sound has been discussed previously [16], and other researches have examined concept and applications of timbral spatialisation as well [6, 17–19]. One can imagine that this particular approach is similar to the concept of orchestration, a term that acousmatic composers are very fond of [5]. They effectively "orchestrate" music on systems like the Acousmonium to achieve the desired spatial and sonic experience, creating and performing gestures based on their own personal taste but also on the spectromorphology of sound [20]. Similarly, timbral spatialisation is capable of creating diffused and immersive sound scenes [18], with the possibility of controlling each part of the sound spectrum algorithmically, creating new possibilities in spatial composition.

However, although the concept of splitting the sound into various frequency bands with bandpass filters is by itself straightforward, it poses several questions of compositional interest that may completely change the resulting spatial scene:

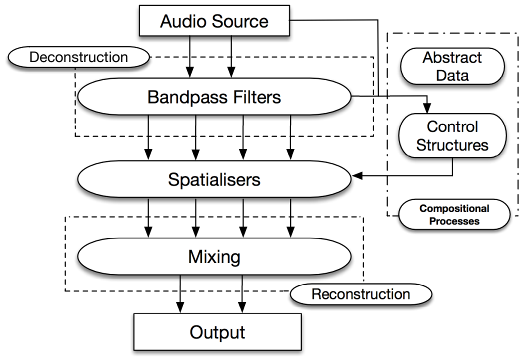

Figure 1. Flowchart of the timbral spatialisation process.

- How many frequency bands should be used?
- Which center frequency for each bandpass filter?
- Why? What's the effect on using different filter settings on the spatial experience?

For sake of consistency it is relevant that the process of deconstruction-reconstruction gives the most faithful result in comparison to the original: this means that is is desirable not to introduce any kind of distortion. In this specific case, different combinations of frequency bands (called Fs or frequency sets) were applied to the incoming signal, in order to experiment on various scenarios, as described in Table 1. In particular:

- Fs_1 frequencies are the preferred octave frequency bands according to the ISO standard [21];
- Fs_2 frequencies from the Random*Source - Serge Resonant Equalizer [1] ;
- Fs_3 frequencies represent a logarithmic scale;
- Fs_4 frequencies are from an API 560 graphic equalizer;

In each set, the lower and upper cutoff frequencies for a single bandpass filter are defined by a pair of values: e.g. [31, 63] are respectively the lower and upper cutoff frequencies for the first bandpass filter in Fs_4 (see Table 1).

The number of used frequency bands has a strong impact on the overall experience: Fs_3, that has only four spectral bands, is more focused and its particular spectral regions are highly localized, while the other frequency sets are more immersive and seem to be a more coherent group. However, different frequency sets can be used depending on the compositional goal and, possibly, on the incoming audio's spectromorphology. In terms of timbral spatialisation, anyway, the process may also be applied to narrower

[1] A unique ten-band filter designed specifically for electronic sound synthesis and processing, where each band, except for bottom and top two frequencies, are spaced at an interval of a major seventh.

Fs_1 (Hz)	Fs_2 (Hz)	Fs_3 (Hz)	Fs_4 (Hz)
22	29	10	31
44	61	100	63
88	115	1000	125
177	218	10000	250
335	411	20000	500
710	777		1000
1420	1500		2000
2840	2800		4000
5680	5200		8000
11360	11000		20000
20000	20000		

Table 1. The various frequency sets (Fs) used.

bands of frequency so that the full audio spectrum of the original source is not reproduced: this would yield non-contiguous but very localized spatial sound-shapes. In this case, the simplest scenario has been taken into consideration, which involves static center frequencies for the bands: further exploration of the spatial implications of this "dynamic" approach can be of relevant compositional interest.

3. TECHNIQUES - CONTROLLING THE SPATIOMORPHOLOGY

Nowadays, many spatialisation systems and techniques are still mixer-oriented [20]: this is because many masterpieces of acousmatic and electroacoustic music were composed for audio systems controlled by a mixer. This means that the interpretation of musical pieces has to be done manually: in the Acousmonium, for example, each fader in a mixer controls the volume of a speaker (or group of speakers) placed strategically in space, with its own frequency response. The history of electronic music interpretation regarding space, is strongly related with the sound-space aesthetic developed in these systems [19]. The development of several new technologies and techniques for controlling spatialisation has changed the way composers approach to this parameter, with very heterogeneous results [22]. The central point of spatialisation is, regardless of the technology used, the control of the spatial environment's attributes. Unfortunately these attributes are not definitive [14], not to mention the lack of a unified musical notation for spatiality and spatialisation [23]. This means that the control of the spatiomorphology of a composition is defined everytime by the composer, perhaps not even coherently with previous compositions. When music is performed by acoustic instruments in an acoustic environment, the physical level of description by itself often provides a workable roadmap to both the listener's experience and the composer's intent [14]: this isn't always true for acousmatic and electroacoustic music, where space is an aesthetically created "environment" [10].

For simplicity, the considered attributes are going to be the geometrical coordinates in a 2D plane (x and y, representing back, front, left, right positions) and a "distance-from-the-listener" (d) attribute: each of these parameters will be applied for every frequency band of the timbral spatialisation process. For example, in the Fs_3, at least twelve parameters (four frequency bands with three attributes each) will be necessary to manipulate the spatial scene. Moreover, the various frequency bands can be seen not just as single, individual point-sources, but also as a group or series of groups, reinforcing the spectral aspects of the spatialisation, with the possibility of "granular" [9] control over the bands. Picking x, y and d attritubes has been a subjective choice of the author. However, this choice provides a more general compositional framework: these attributes can be defined in any spatial environment (in various degrees), making it easier to switch between spatialisation technologies, or even between implementation languages.

As previously noted, audioparous techniques indicate a composition mode in which musical data originate from sound. We can formalize several audioparous procedures that can shape and define different aspects and timescales of the compositions, from micro to macro musical organization: e.g. single generative spatial gestures, the flocking movement of the whole timbral spatialisation etc.

However, differently from other examples of audioparous compositional procedures [11, 12], the application of such techniques in spatial contexts do not produce notation or sound per se, but rather modify and reshape an incoming audio signal. Furthermore, the actual incoming sound is an interesting parameter to explore in order to control sound itself: we can then extend the idea of audioparism to a "self-audioparism", where the incoming audio, perhaps through some other control process, spatialises itself.

3.1 Exploring noise

In the time domain, noise can be defined as sound in which the amplitude over time changes with a degree of randomness. The amplitude is maximally random in the so-called *white noise*. In the spectral or frequency domain, noise can be defined as sound that has a continuous power spectral density over a certain frequency bandwidth. The power spectral density of all frequencies is equal in white noise. [2]. Consequently, there are different "flavours" of noise, and they can be suitable for different compositional strategies.

Some synthesis techniques in analog electronic instruments from the "control-voltage" era are still alive and kicking: for example the use of sample and hold circuits to create random sequences by "feeding" them pink noise [24].

This particular techniques enables to sample a value upon a received trigger and hold it still until another trigger is received: if the sampled signal is noise, then the output would be a sequence of random value. Even more interesting is the interpolation between these values: instead of jumping quickly from one value to another, the transition is smooth, producing all the values in between as well, like in Figure 2. In a spatial context these values can be easily used to automate all the parameters needed for each frequency bands: it's simple to create a rich and immersive spatial scene by generating several of these functions. Furthermore, by controlling the sampling rate, the speed of

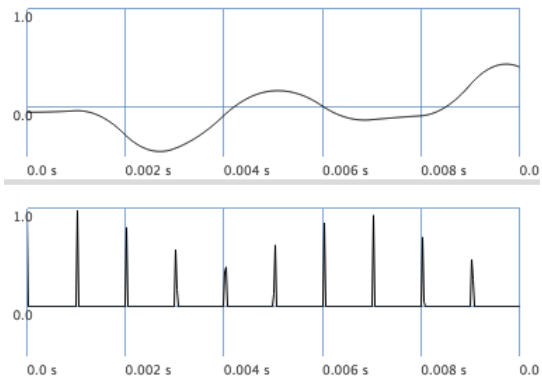

Figure 2. Quadratically interpolated random values sampled from a noise function with relative trigger inputs.

the change of spatial attributes would dramatically change: nesting noise functions into other noise functions to dynamically change the *x* and *y* positions of the frequency bands adds fluctuation and unpredictability (Figure 3).

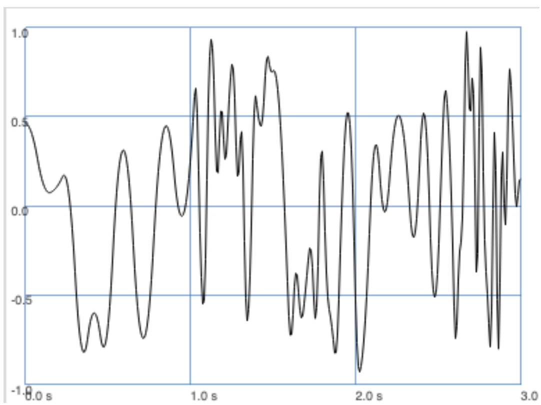

Figure 3. A sampled noise function with its sampling rate modulated by another nested noise function.

3.2 FFT controlled walk

The Fast Fourier Transform (FFT) inspects the frequency content of a signal, and can be extremely useful for audio analysis or frequency-domain sound processing. By "windowing" a real-time signal, a succession of overlapped spectral frames are obtained: the FFT focuses attention on the magnitude and phases values for all of the different frequency bin resulting from this operation [15]. By using this data, it is possible to spectrally analyze sound, extrapolating some of its intrinsic characteristics: these attributes can be used compositionally in a spatial context by mapping them to the relevant values, in our case the *x*, *y* and *d* attributes described in Section 2.

In this context, the so-called *Instantaneous Spectral Descriptors* [25] are used. They are a set of instantaneous attributes obtained from the FFT analysis that describe the spectral shape of sound in a certain moment in time: we could say that they are a photograph of the spectromorphology of audio. In particular, three descriptors were used:

- *Spectral Centroid*: the weighted mean frequency, or the "centre of mass" of the spectrum. It can be a useful indicator of the perceptual "brightness" of an audio signal;
- *Spectrum Roll-off point*: the frequency below which lies the 90/95 percent of the signal energy. This somewhat indicates the harmonic/noise cutting frequency;
- *Spectral Flatness*: it measures of "noisiness" or "sinusoidality" of the spectrum (or parts of it);

These descriptors are used as dynamic controllers for random walks, also called Brownian motion [26]. Specifically, two random walk functions generate the *x* and *y* position for each of the frequency bands: respectively, the Spectral Flatness and Spectral Centroid descriptors control the frequency and step's amplitude for the previously defined functions. This means that depending on measures of "noisiness" and the perceptual "brightness" of the incoming audio, there will be a changing number of steps per second and each of them will move closer or further away from the previous one. The Spectrum Roll-off point descriptor, instead, will control the distance attribute for the whole group of frequency bands acting as a global modifier and effectively treating them as a collective group that behaves coherently. The use of FFT analysis in this particular compositional technique, defines what was previously discussed as "self-audioparism": using the spectral descriptors values obtained by the analysis of the incoming audio onto itself, we define a self-sustaining, ever-changing algorithmic spatial texture.

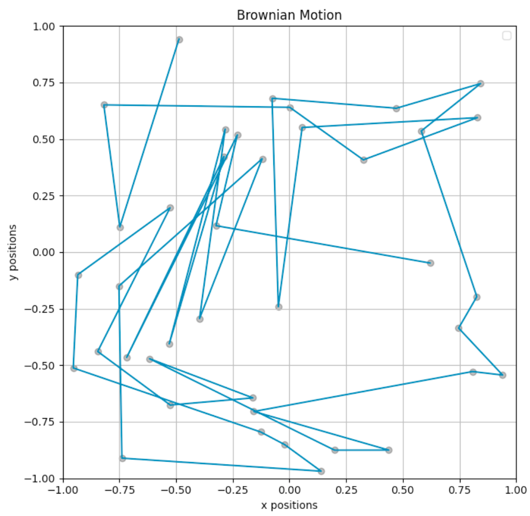

Figure 4. A 40 steps walk for a single frequency band in the 2D space, where (0,0) is the listener's virtual position.

3.3 Audio snapshot technique

The audio snapshot is yet another audioparous technique that involves sampling into a buffer a rather small snippet of sound, using this data to drive the spatialisation. This techniques is actually self-audioparous, because the sound being recorded is indeed from the source itself. The basic idea is to use an interpolating buffer player set to very low playback rates in order to read data: in this way we "transpose" information from a *micro* timescale (a few milliseconds of audio) to a *sound object* or *meso* time scale (from several tenths of a second to a few seconds of audio) [13] .

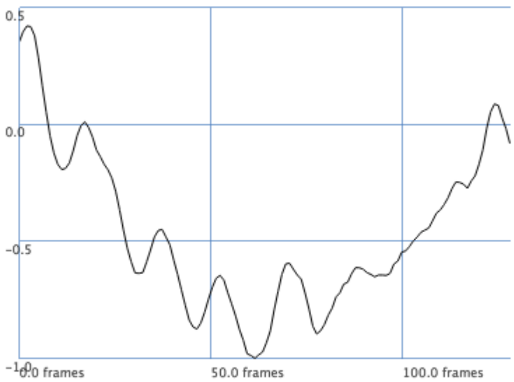

Figure 5. One snapshot of 128 frames, taken from *Solar Ellipse* by Barry Truax.

This particular technique of data generation, in practice, produces pseudo-random spatial gestures, depending on the length of the buffer: from 16 to 64 frames the gesture is definitely short, while from 128 onwards the spatial phrase starts being long enough to reach into the *meso* timescale. This, of course, depends on the playback rate of the buffer players: extending its frequency into audio territory (over 20Hz), one could even have audio-rate spatial modulation.

More precisely, for each of the frequency band, three buffer players will read the data: one for x position, one for y position and one for the distance attribute. Each of these players will read the buffer from random initial positions in order to get coherent but varied results: the combination of these three functions in time will be the actual gestural output applied to the spatial texture. Randomizing both the playback rate and the initial reading position, will also ensure that the resulting spatial gesture is always different for each of the spectral bands, giving the impression of a richer spatial scene. At any time the initial buffer can be resampled, loading new information and starting back from scratch with new audio data, producing completely new spatial motions.

Furthermore, once the buffer has been filled, typical Digital Signal Processing operation can be applied to the recorded data: for example, using sub-audio or audio generators with appropriate parameters will yield interesting and extreme results (see Figure 6).

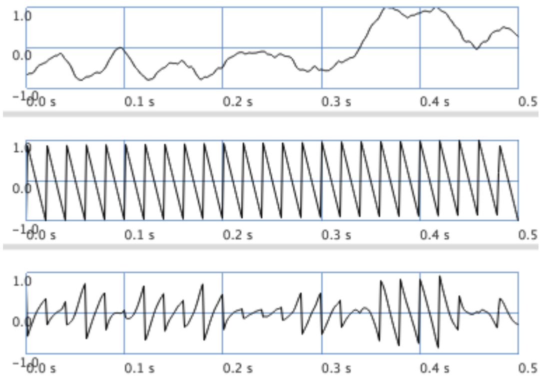

Figure 6. In the upper graph is the original sampled signal; in the middle graph 50Hz sawtooth wave acting as a modulator; in the lower graph the resulting signal from the multiplication of the two previous signals.

4. IMPLEMENTATION

The implementation of both the timbral spatialisation and the control techniques poses some challenge in the organization of the data flow, and can be relatively CPU heavy due to the spatial rendering of many frequency bands: in experimenting with these techniques, an Ambisonics [27] binaural approach was used. However, other technologies (such as DBAP [28], for example) may be more suitable for multichannel setups, both from perceptual aspect (no sweetspot) and computational aspects (much cheaper).

All the software was implemented in the SuperCollider environment [29] which features an Object Orientend programming language that controls a powerful audio synthesis server. The use of an audio dedicated programming language is particularly fitting because it comes pre-loaded with algorithms and Unit Generators that can be easily integrated in the workflow. Moreover, the SuperCollider community provides a series of free classes and add-ons [2] that dramatically extend the capabilities of both the language and the synthesis server. One of these free plugins has been used to implement the binaural spatial rendering, the "Ambisonics Toolkit [3] " (Atk), and more precisely the FOA ("First Order Ambisonics"). With a process of encoding, transforming and decoding (Figure 7), the Atk effectively allows the user to easily spatialize sound by controlling the transformation procedures, and specifying the nature of encoder and decoder(s). In this case, the spatial

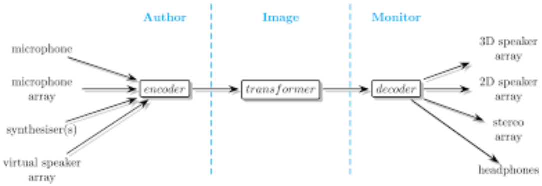

Figure 7. The Ambisonics workflow.

[2] https://github.com/supercollider/sc3-plugins
[3] https://www.ambisonictoolkit.net/

rendering is obtained through a binaural decoder based on the IRCAM's Listen HRTF database [4] which provides various equalization for different head's widths, allowing the user to choose their preferred settings. The spatial render is interchangeable with any appropriate binaural render, providing several of this decoders with different properties: for example a Synthetic spherical head model HRTF [30] or a CIPIC HRTF database from the University of California Davis [5].

It is important to mention that the SuperCollider architecture is based on the dualism between language (*sclang*) and server (*scsynth*) [31]: many processes can be implemented within both the interpreted language or the synthesis server, depending on the user's need. Specifically, the techniques and approaches presented here, have been implemented server side: in other words most of the functions such as noise generators, oscillators, FFT analysis etc.. are specific Unit Generators that are allocated and managed dynamically on *scsynth*. The advantage of this organization is that, other than the readiness of use, the low-level (C++) implementation of the Unit Generators allows for a much more optimized use of the computational power. Furthermore, *scsynth* offers a flexible and multichannel bus system which is perfect for sending/receiving the large number of computed data in the spatialisation system.

In SuperCollider there is a great choice of Ugens that can be used inside this compositional framework, or even substitute the ones presented here. For example:

- a number of "coloured" noise Ugens or stochastic generators such as WhiteNoise, PinkNoise, BrownNoise, GreyNoise or Crackle (a noise generator based on a chaotic function);
- various degrees of interpolated or non-interpolated sample and hold Ugens such as LFNoise0 (non interpolated), LFNoise1 (linearly interpolated) or LFNoise2 (quadratically interpolated);
- different types of Ugens suitable for real-time audio analysis such as Pitch (autocorrelation pitch follower), Amplitude (envelope follower), Loudness (extraction of instantaneous loudness in sones) etc..
- several add-ons are present, ranging from the sc3-plugins to a large suite of audio analysis tools called "Fluid Decomposition Toolbox" [32], all freely available;

While for each spatialised frequency band the x and y positions are straightforward (representing front, back, left and right), the d attribute described in Section 3 has been implemented according to [33] on distance cues. In order to emulate these cues, the d parameter scales the amplitude of each frequency band so that the direct signal decreases in amplitude more with distance than does the reverberant signal.

[4] http://recherche.ircam.fr/equipes/salles/listen/

[5] http://interface.cipic.ucdavis.edu/sound/hrtf.html

The implemented techinques are available publicly on the author's GitHub [6].

5. CONCLUSIONS

Timbral spatialisation is a signal processing technique that has a great potential for creating rich and fascinating spatial textures, but it can also be viewed as a tool for composing space and effectively considering it a part of the compositional workflow. Together with algorithmic techniques for the control of the spatial environment, it is possible to automate the set of attributes that define such virtual space. Furthermore, the introduced concept of "self-audioparity" adds another layer of complexity and coherence to the whole spatialisation process.

The possible applications for the discussed techniques are multiple and in different contexts:

- live spatialisation of electroacoustic and acousmatic performances;
- the reinterpretation in multichannel setups or binaural rendering of fixed media compositions;
- as a standalone tool for spatial composition and for the integration of space inside of a composer's workflow;

It is important to notice that the implemented techniques are described from a compositional point of view, which means that the final user can adjust the internal parameters and mappings according to its own taste: the specifications collected here are a representation of what is possibile, but are by no means definitive. Moreover, many other techniques can be implemented, or perhaps expanding the ones that have been presented: for example, further explorations of different spectral descriptors and its mapping to spatial attributes; the implementation of "dynamic" instead of "static" frequency bands; the creation of completely new audioparous algorithms from the ground up.

The next steps will involve the development of a framework for easily switching between techniques in real-time; inclusion of the "height" attribute in the spatialisation process; the implementation of a GUI for visual feedback; the integration of machine learning techniques for intelligent spatialisation; the creation of a set of hardware tools for the performative control of the spatial textures.

6. REFERENCES

[1] W. J. Solomon, "Spatialisation in music: the analysis and interpretation of spatial gestures," Ph.D. dissertation, University of Georgia, 2002.

[2] C. Roads, *Composing electronic music: a new aesthetic*. Oxford Univerisity Press, 2014.

[3] K. Stockhausen, "Music in space," *Die Reihe*, vol. 5, pp. 67–82, 1961.

[6] https://github.com/StefanoCatena/Timbral-Spatialization/tree/main

[4] F. Schumacher and C. Fuentes, "Space–emotion in acousmatic music," *Organised Sound*, vol. 22, pp. 394–405, 12 2017.

[5] S. Desantos, C. Roads, and F. Bayle, *Acousmatic morphology: an interview with Francois Bayle*. Computer Music Journal, Fall 1997, vol. 21, no. 3, pp. 11–19.

[6] R. Normandeau, "Timbre spatialisation: the medium is the space," *Organised Sound*, vol. 14, pp. 277 – 285, 12 2009.

[7] S. James and C. Hope, "2d and 3d timbral spatialisation: Spatial motion, immersiveness, and notions of space," *International Computer Music Conference*, pp. 77–84, 01 2013.

[8] S. James, "Spectromorphology and spatiomorphology of sound shapes: Audio-rate aep and dbap panning of spectra," *International Computer Music Conference*, pp. 278–285, 09-10 2015.

[9] D. Kim-Boyle, "Spectral and granular spatialization with boids," *Computer Music Journal*, pp. 139–142, 01 2006.

[10] D. Smalley, "Space-form and the acousmatic image," *Organised Sound, 12(1)*, 2007.

[11] A. Valle, "Sampcomp: sample-based techniques for algorithmic composition," in *Proceedings of the 22nd CIM*, November 20-23 2018, pp. 128–135.

[12] O. Messiaen, *The Techniques of my Musical Language*. Paris: Leduc, 1956.

[13] C. Roads, *Microsound*. MIT Press, 2001.

[14] G. Kendall and M. Ardila, "The artistic play of spatial organization: Spatial attributes, scene analysis and auditory spatial schemata," *Computer Music Modeling and Retrieval. Sens of Sounds, 4th International Symposium, CMMR*, pp. 125–138, 08 2007.

[15] M. Dolson, "The phase vocoder: A tutorial," *Computer Music Journal*, vol. 10, pp. 14–27, 12 1986.

[16] K. Stockhausen, *Vier Kriterien der Elektronischen Musik*. DuMont Buchverlag, 1970- 1977, vol. 4, pp. 360–401.

[17] R. Torchia and C. Lippe, "Techniques for multi-channel real-time spatial distribution using frequency-domain processing." *Conference on New Interfaces for Musical Expression (NIME04)*, pp. 116–119, 06 2004.

[18] S. James, "A classification of multi-point spectral sound shapes," *Proceedings of the ACMC/AFAE Conference*, pp. 57–64, 07 2016.

[19] A. Barberis, S. Mungianu, and S. Sapir, "Ormé, a tool for automated spatialization of fixed-media music based on spectrum contents annotation," *Colloqui d'Informatica Musicale (CIM)*, 09 2016.

[20] A. Vande Gorne, *L'interprétation spatiale. Essai de formalisation méthodologique*. Université de Lille, December 2002.

[21] "Acoustics, Preferred frequencies for measurement," ISO 266:1997(E), 1997.

[22] N. Peters, G. Marentakis, and S. Mcadams, "Current technologies and compositional practices for spatialization: A qualitative and quantitative analysis," *Computer Music Journal*, vol. 35, pp. 10–27, 03 2011.

[23] A. Vidolin, "Suonare lo spazio elettroacustico," *Quaderni di Musica*, vol. 51, 2002.

[24] A. Strange, *Electronic Music: Systems, Techniques and Controls.* Dubuque, Iowa: William C. Brown Company, 2nd Ed., 1983.

[25] G. Peeters, "A large set of audio features for sound description (similarity and classification) in the cuidado project," Institute for Researche and Coordination in Acoustics/Music, Tech. Rep., 01 2004.

[26] A. Valle, "Algorithmic spatialisation. interfacing supercollider to the wavefield system," CIRMA/DAMS, Department of Humanities (StudiUm), University of Torino., Tech. Rep., 2016.

[27] R. K. Furness, "Ambisonics-an overview," in *Audio Engineering Society Conference: 8th International Conference: The Sound of Audio*, May 1990.

[28] T. Lossius and P. Baltazar, "Dbap – distance-based amplitude panning," *International Computer Music Conference*, 01 2009.

[29] S. Wilson, D. Cottle, and N. Collis, *The SuperCollider Book*. Cambridge, Mass.: The MIT Press, 2011.

[30] R. O. Duda, "Modeling head related transfer functions," *Proceedings of the Twenty-Seventh Annual Asilomar Conference on Signals, Systems and Computers*, 1993.

[31] A. Valle, *Introduction to SuperCollider*. Logos Verlag, 2015.

[32] P. Tremblay, O. Green, G. Roma, and A. Harker, "From collections to corpora: Exploring sounds through fluid decomposition," *International Computer Music Conference and New York City Electroacoustic Music Festival - ICMC-NYCEMF*, 2019.

[33] J. Chowning, "The simulation of moving sound sources," *Computer Music Journal*, vol. 1, 06 1977.