SpatGRIS/ServerGRIS, Creative tools for 2D and 3D sound spatialization

Robert Normandeau
Université de Montréal
robert.normandeau@umontreal.ca

Olivier Bélanger
Université de Montréal
belangeo@gmail.com

Christophe Lengelé
Université de Montréal
christophe.lengele@umontreal.ca

David Ledoux
Université de Montréal
david.ledoux@umontreal.ca

ABSTRACT

The *Groupe de Recherche en Immersion Spatiale* (GRIS) is a research group based at Université de Montréal. It has been involved in spatialization software development since 2009. *SpatGRIS* is a multichannel sound spatialization plugin available in Audio Unit (Mac) and VST formats (Mac and Windows). This free and open-source plugin allows movement of multiple sound sources on a variable set of speakers. Different modes of grouped movement and a trajectory system are provided for spatialization of mono, stereo and multichannel sources. The spatialization is made directly within the DAW without the need for external software. *SpatGRIS* has also a mode called *OSC Spatialization*. In this mode, the audio is not spatialized within *SpatGRIS* but externally by the *ServerGRIS*. In this case, *SpatGRIS* sends only Open Sound Control (OSC) data to the Server to place the sounds in a set of speakers. The sound itself is sent from the DAW to the Server via *JackRouter*. Finally we propose a standard for the circulation of multichannel works.

1. INTRODUCTION

The two main ideas that have led to the development of GRIS spatialization tools are the following: 1. The vast majority of electroacoustic music composers use a DAW as their main composition tools; 2. The composition of space must be concurrent with the composition of time in the general process of development of the work. The space is not an added flavor, a color sprinkled at the end of the course when the work is already finished on the temporal plane. It is an integral part of the electroacoustic work of the 21st century, it is even the most original and exclusive component. For these two reasons, we decided to develop not a new sound spatialization software - there are already many - but rather a plug-in that integrates with the daily work tool of composers.

The first plug-in that we developed in 2010 was *Octogris*, intended for octophonic sound spatialization. Version 2 was released in 2014 and version 3 in 2016. Meanwhile, in 2012, we released the first version of *ZirkOSC*, a plug-in designed to control *Zirkonium* MKI from ZKM, a 3D spatialization software using VBAP[1]. Version 2 of *ZirkOSC* was released in 2014 and version 3 in 2015. In 2017, we merged these two plug-ins into one, named *SpatGRIS*. Finally, in 2018, we released the first version of *ServerGRIS*, a 2D and 3D spatialization software.

All the combined versions of *Octogris* have been downloaded more than 5000 times, while those of *ZirkOSC* have been more than 1000 copies[2]. *SpatGRIS*, since its launch in May 2017, has exceeded 1,200 downloads[3].

2. SPATGRIS

The *SpatGRIS* is a plugin (Mac AU/VST and VST Windows format) designed to compose multichannel space. It allows the user to spatialize the sound in 2D (up to 16 speakers) or in 3D (up to 128 speakers) under a dome of speakers with the ServerGRIS,

SpatGRIS is a fusion of two former plugins by the GRIS: *OctoGRIS* and *ZirkOSC* with a lot of new features.

SpatGRIS offers two audio modes:

- o Free volume
- o Pan span

It also offers an OSC Spatialization mode that is described in the *ServerGRIS*.

[1] Vector Base Amplitude Panning, for dome of speakers

[2] Information compiled on GoogleCode and SourceForge.net

[3] Information compiled on SourceForge.net

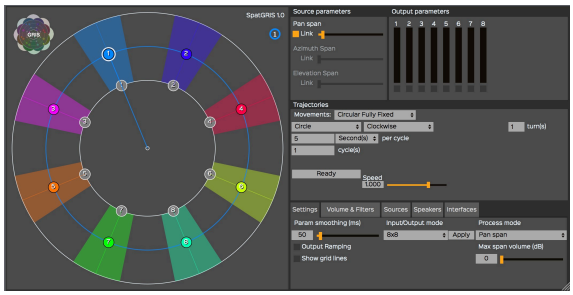


Figure 1: SpatGRIS

2.1 Interface

SpatGRIS has a graphical interface consisting of five panels:

- 2D Spatializer
- Source parameters
- Output parameters (VU meters)
- Trajectories
- Configuration panel, subdivided in tabs.

2.2 The 2D Spatializer

The 2D Spatializer is a two-dimensional representation of the sound sources distributed on a speaker system. The two concentric circles represent:

- The equidistant circle of speakers;
- The far limit of the spatialization system;
- Gray dots each represent a speaker;
- Coloured dots represent the sound sources.

2.3 Source parameters

When *SpatGRIS* is used with the Free Volume spatialization mode, each source has its own emission area. The level of the signal sent to the speakers is then proportional to their distance to the centre of the source.

In the Pan Span mode, the Elevation Span and Azimuth Span parameters replace the Surface parameter.

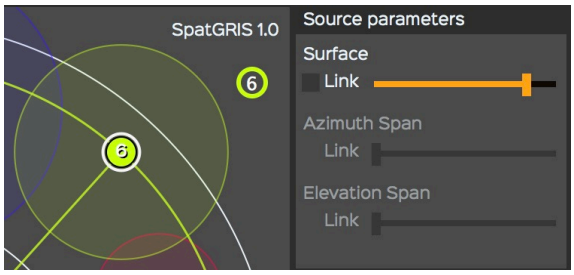


Figure 2: Source parameters

2.4 Trajectories

It is possible to automate the movement of the sources using predefined trajectories. Within the Trajectories control panel, you can set and adjust these movements to the musical context. The concept behind the movements is that there is a source master while the others are slaves. This way, the number of parameters written in the DAW is then limited to Source 1 X and Y, and not to every source.

SpatGRIS offers eight different sorts of trajectories: Circle, Ellipse, Spiral, Pendulum, Random, Random Target, Sym X target and Sym Y Target.

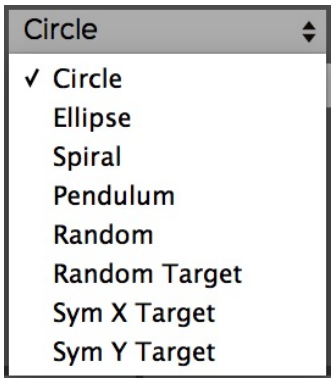


Figure 3: Trajectories

2.5 Group Movements

Within the Trajectories section, under the Movements drop-down menu, it is possible to choose different ways of linking all sources together.

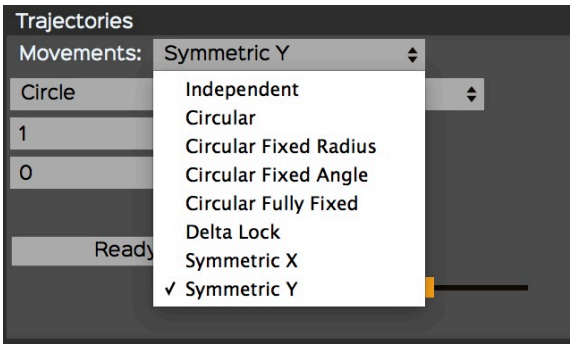


Figure 4: Group Movements

2.6 Configuration Panel

The configuration panel gives access to a set of plugin parameters. These settings are grouped under tabs with explicit names:

- Settings
- Volume & Filters
- Sources
- Speakers
- Interfaces

2.7 Interfaces

SpatGRIS is controllable via an external Open Sound Control (OSC) controller. The first version of a TouchOSC™ interface is already available. Under the Interfaces tab, you can configure the connection between the OSC controller and the plugin. It is also possible to control *SpatGRIS* with a Leap Motion™ controller. Finally, it is possible to control *SpatGRIS* with a joystick or almost any USB gaming device.

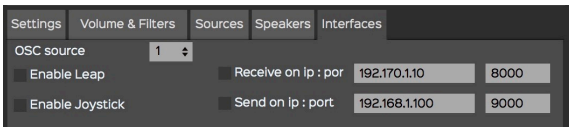


Figure 5: Interfaces

3. SERVERGRIS

3.1 Why a new VBAP software?

Because most of the already existing ones are either autonomous piece of software — meaning that you have to use them with stems, or already composed works —, or very difficult to use on a daily basis — most of the time because their development was stopped at one point or another.

3.2 Presentation

The *ServerGRIS* is an external multichannel sound spatializer for different configurations of speakers in 2D (plane mode: X and Y axis) or 3D (vertical mode: X, Y and Z axis). It is based on the JackRouter HAL plugin which is installed and controlled with the Server. The Server may include up to 128 inputs and outputs. The movements are sent from the SpatGRIS plugin in OSC mode to the Server. The audio spatialization itself is made by the *ServerGRIS* and sent to the audio interface.

The main difference between the *SpatGRIS* plugin in audio mode and the combination *SpatGRIS*/*ServerGRIS* is that the former is limited by the DAW's maximum number of outputs per track, which usually varies from 2 to 8[4]. This means that the user is limited to octophonic spatialization (hence the original name of the plugin: *OctoGRIS*). *SpatGris* also doesn't take into account the vertical dimension of the spatialization (2D only). *ServerGRIS* does have a practical limitation of up to 128 outputs (but theoretically up to 256[5]) and is a 3D software.

ServerGRIS comes with a complete manual and a Quick Start Guide that helps the users to progress very rapidly through a limited number of steps before obtaining satisfying results. We think that this easiness may help to see more and more multichannel pieces dealing with 3D parameters.

3.3 Architecture

The *ServerGRIS* setup is made of three elements:
• The *ServerGRIS* itself that spatializes the sound
• The Edit Speakers page that designs the setup of the speakers
• The *SpatGRIS* that designs and records the trajectories.
The whole architecture includes these elements (Audio and OSC are working in parallel):

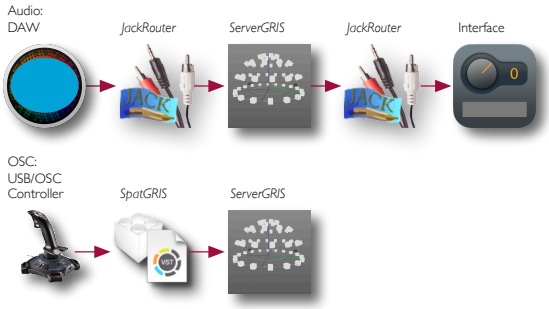


Figure 6: Architecture

3.4 Jack and Multi-client

JackRouter is at the heart of the *ServerGRIS* architecture, but unlike other spatialization software that uses it, it is here entirely managed by *ServerGRIS*, which makes it all the easier to use. *ServerGRIS* is very flexible and various changes either in the Speaker Setup or in the DAW are dynamically reflected in the software. There is no need to turn off the system in cases of changes, like with JackPilot.

ServerGRIS is also multi-client:

Client	Start	End	Available	On/Off
Digital Perfo...	1	20	20	<->
REAPER	21	148	128	<->

Figure 7: Multi-client

3.5 ServerGRIS Zones

The *ServerGRIS* window is divided into 4 zones:
• Speakers and Sources
• Inputs/Outputs
• Parameters
• Client

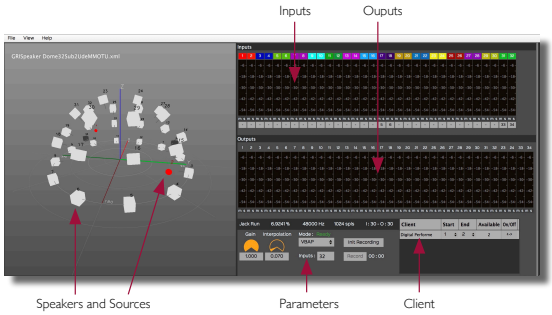


Figure 8: ServerGRIS Zones

3.6 Algorithms

ServerGRIS uses a number of algorithms and there will be more in the future.

[4] Except for Reaper, 64 outputs.

[5] The limitation of the ServerGRIS are those of Jack, which is a mono-core software.

VBAP

Vector Base Amplitude Panning (VBAP) is an algorithm created by Ville Pulkki in 2001. It allows the user to spatialize the sound under a dome of speakers by placing the sound according to the relative amplitude of three speakers (instead of two in 2D spatialization).

LBAP

Layer-Base Amplitude Panning (LBAP) is used to spatialize the sound in 3D environment based on different layers in a CUBE. Almost any set of speakers can be configured with this algorithm that is currently under development.

BINAURAL

Head Related Transfer Function (HRTF) is a function that reproduces the way we perceive the localization of sounds in the space. It is a set of phase and amplitude calculations for listening on headphones. Primarily, it is designed for 5.1 reproduction or immersive listening on headphones, situations found in the gaming and the VR industries.

STEREO

There is a simple stereo mode to listen to a complex project on a pair of speakers. All the sounds are sent to the corresponding speakers depending on their location (left to left, right to right).

3.7 Speaker Setup for a dome

A speaker configuration is created by determining the number of speakers in each ring and their location (Zenith, Radius and Offset Angle). An Output Patch connects the speakers to the output number of the audio interface (a 32-speaker dome plus 2 subs (direct outs) in this example).

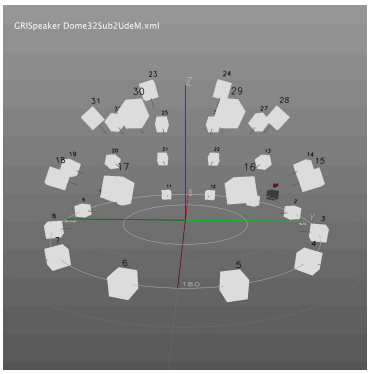


Figure 9: A dome of 32 speakers

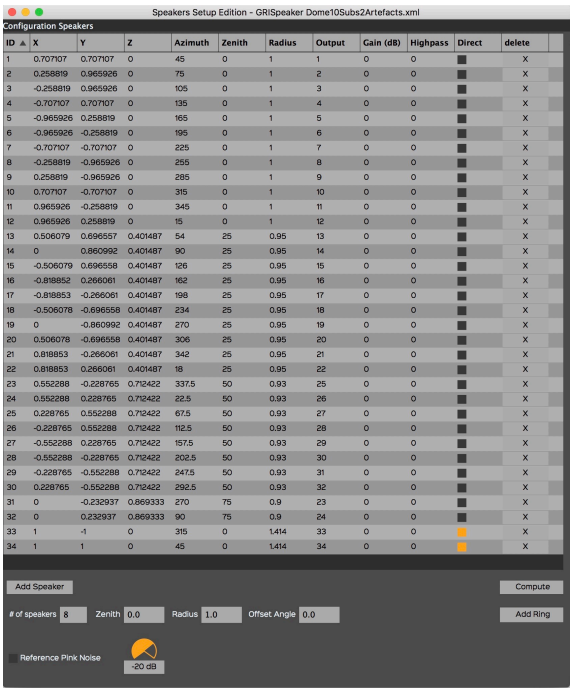
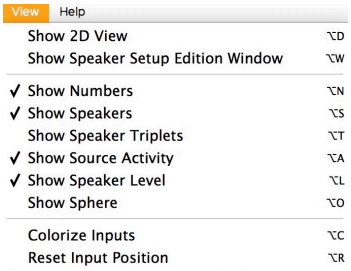


Figure 10: The Speaker Setup Edition Window

• The output Volume and a Highpass filter is available on each speaker to calibrate the setup.
• Azimuth is the plane angle, from 0˚ to 360˚.
• Zenith is the elevation angle from 0˚ to 90˚ (or to -90, for a complete sphere).
• Radius is the distance from the speaker to centre of the sphere.
• It is possible to add speakers that are assigned to direct outputs, which means that they are not part of the spatialization algorithm, like the subwoofers. These speakers are identified with an orange rectangle in the setup.

3.8 View Menu

This is where what is shown in the 3D view is adjusted:



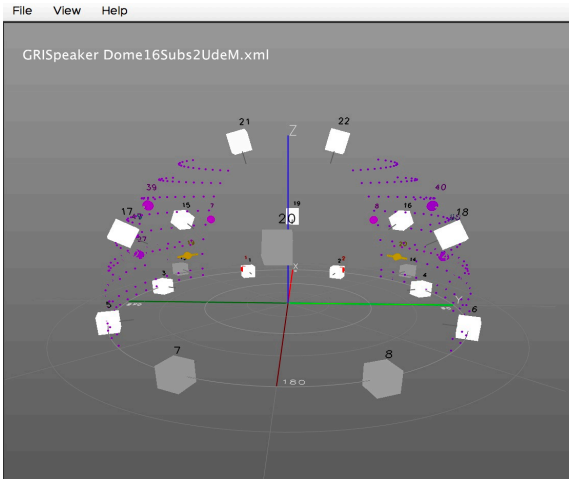


Figure 11: View Menu: Show Numbers, Speakers, Activity and Level

• Show 2D view: A 2D view from the top of the dome is given with only the sources showing.
• Show Speaker Setup Edition Window: A window access to all the parameters taken into account for a valid configuration of speakers (See 4.7. Speaker Setup).
• Show numbers: Show or hide the numbers of the sources and speakers.
• Show Speakers: Show or hide the speakers in the 3D view.
• Show Speakers Triplets: Show or hide the triplets in the 3D view. Triplets reveal how the speakers are connected to each other, a very useful tool for troubleshooting
• Show Source Activity: This option allows you to see the trajectories of the sources (big dots) as well as their energy, displayed according to the Azimuth and Elevation Span (small dots) sent via the SpatGRIS plugin. Note that there is nothing to see when the DAW is stopped. The threshold is set at -70 dB. When not selected, all the sources are shown in a static 3D view, even when the DAW is stopped.
• Show Speaker Level: Shows how much energy each speaker delivers. From grey (nothing) to white (maximum).
• Show Sphere: If you have the chance to play in a full sphere!
• Colorize Inputs: This option allows you to set all the inputs to a different colour within the visible spectra from red to purple. Be careful, it erases all the custom colours already in place.
• Reset Input Position: When changing the project in your DAW, sources may stay in place until you use this option to clear up the 3D view.

3.9 Recording

When the automation of the spatialization is completed, there are two options:
1. To play the piece "live" with the DAW and the ServerGRIS.
2. To record the spatialization in separate files that correspond to the number of speakers in the setup.

The ServerGRIS records mono files (AIFF or WAV according to your preferences).

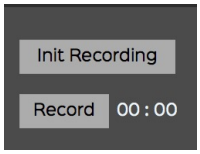


Figure 12: Recording

4. FUTURE

Future development of these tools is divided into different projects:

• SpatGRIS will be rewritten in the next few months from the ground up. Despite the fact that it works properly, some features are still missing and before adding them, we need to clean up the code that was designed by four different developers since 2010. One of the new features is the bidirectional relationship between *SpatGRIS* and *ServerGRIS*.

• Trajectories will be enhanced by adding math functions to their evolution over time. At the moment, only linear trajectories are available.

• The HRTF mode will be greatly enhanced to encourage the users work at home on headphones.

• A LBAP[6] algorithm that would give the *ServerGRIS* a very flexible way to deal with all different speaker setups like standard acousmonium, sound installation, non-conventional setups, etc.

• More external controllers like the Lemur and more templates for existing ones like the iPad and the Leap Motion.

5. A STANDARD FOR MULTICHANNEL DIFFUSION

One of the major problems of multichannel diffusion is the lack of standardized formats. Each composer has his own rules, often dictated by the setup he intends his work for, and each of these setups has unique characteristics. How to circulate the works then, without having to have to completely redo the spatialization? Already with the concept of a dome of speakers, we are approaching a certain form of standardization since the VBAP algorithm allows to move from one dome to another without having to change the space. The design of the spatialization is independent of the configuration of the speakers. But if this applies to concert situations, what about individual listening, webcasting, or the organization of multichannel competitions? We propose a solution to this situation.

5.1 An autonomous multi-channel file reader

We added a multichannel file reader to ServerGRIS. This makes it standalone and independent of a DAW. It is

[6] Layer-Base Amplitude Panning

therefore possible to read directly a file containing several tracks. But how to make sure that spatialization respects the composer's will?

5.2 Proposition of a standard at 16.2, 24 bits, 48kHz

We propose to the international community the establishment of a standard of spatialization with 16 channels plus 2 subwoofers (format 16.2) which will allow a better diffusion of multichannel works. This standard is the result of many years of experience and it seems to us the result of the optimization of a certain number of parameters:

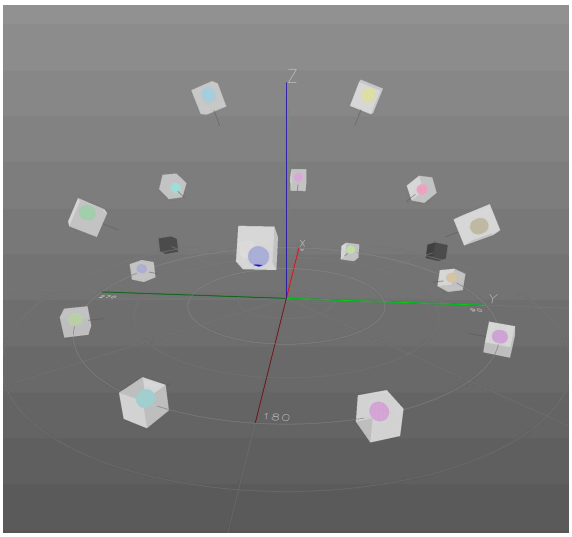
- Sufficient number of channels for good resolution;
- Interleaved file format of reasonable size (156 Mb / min);
- Reasonable number of speakers for a concert setup.

5.3 Binaural listening or speakers

The 16.2 file in the ServerGRIS player will use its default location of 16.2 for binaural listening (this is already the case with the Binaural_Speaker_Setup) or for listening with speakers.

5.4 Listening online

An audio content provider could place the ServerGRIS upstream of its broadcast system and upload only the binaural contents of a multichannel file for headphone listening, reducing the bandwidth required.



6. CONCLUSION

We have been involved in the spatial composition under dome of speakers for over twenty years now. We have gone through all kinds of technologies from the least sophisticated to the most sophisticated and we have spent a lot of time diverting these tools more or less well adapted to the reality of today's composers. There are more and more large permanently installed sets of speakers around the world — just to mention Virginia Tech, BEAST, Audiorama, ZKM, Graz University, University of Montreal, etc. - and therefore more and more opportunity to present very elaborate music in terms of spatial immersion. We hope that with these tools, and especially with the newest, the ServerGRIS, just released in February 2018, we will contribute to the ease of composing for this new medium.

Acknowledgments

Thanks to the actual team of developers and also to the former teams: Antoine Missout, Vincent Berthiaume and Nicolas Masson.

The development of these tools was made possible thanks to research grants from Hexagram, CRSH and FRQSC.

7. REFERENCES

[1] N. Barrett, "Spatio-musical composition strategies," Organised Sound, 7(3), 313–323, 2002.

[2] D. L. Barreiro, "Considerations on the handling of space in multichannel electroacoustic works", Organised Sound, 15(3), 290–296, 2010.

[3] T. Carpentier, "Tosca : un plugin de communication OSCpour le mixage spatialisé orienté objet," JIM, 2015.

[4] D. Kostadinov, J. D. Reiss and V. Mladenov, "Evaluation of Distance Based Amplitude Panning for Spatial Audio," ICMC, 2009.

[5] C. Miyama, G. Dipper and L. Brümmer, "Zirkonium MKIII - A Toolkit for Spatial Composition," Journal of the Japanese Society for Sonic Arts, Vol.7 No.3 pp.54–59.

[6] R. Normandeau, "Timbre spatialisation: the medium is the space," Organised Sound, 14(3), 277–285, 2009.

[7] R. Peñha and J. Oliveira, "Spatium, Tools for Sound Spatialization," SMC, 2013.

[8] C. Ramakrishnan, J. Goßmann, L. Brümmer, and B. Sturm, “The ZKM Klangdom,” Proceedings of the 2006 Conference on New Interfaces for Musical Expression, pp. 140–143, 2006

[9] V. Pulkki, "Virtual Source Positioning Using Vector Base Amplitude Panning," Journal of the Audio Engineering Society, vol. 45, pp. 456-466, 1997.

[10] C. Ramakrishnan, "Zirkonium: Noninvasive Software for Sound Spatialisation," Organised Sound, vol. 14, pp. 269-276, 2009.

[11] D. Wagner, L. Brümmer, G. Dipper and J. A. Otto, " Introducing the Zirkonium MK2 System for Spatial Composition," Proceedings of the 2014 ICMC-SMC Conference.

[12] M. A. Baalman, "Spatial Composition Techniques and Sound Spatialization Technologies," Organised Sound, vol. 15, no. 3, pp. 209-218, 2010.

[13] A. Roginska and P. Geluso, "Immersive Sound, The Art and Science of Binaural and Multi-channel Audio", Routledge, New York, 2017.