3D speaker management systems – Mixer integration concepts

Etienne Corteel1, Peter Glaettli2, Raphael Foulon1, Reda Frauly1, Ingo Hahn2, Roger Heiniger3, Renato S. Pellegrini1

1 Sonic Emotion Labs, France, Email: etienne.corteel@sonicemotion.com
2 Studer Professional Audio GmbH, Switzerland, Email: Peter.Glaettli@harman.com

Abstract

3D speaker management systems are often only used when creating special spatial effects is required in a show. However, a 3D speaker management system can replace a traditional speaker management system and a delay matrix. Such a 3D system must support main sound reinforcement setups with high power loudspeakers allowing for homogeneous sound level distribution. Spatial effects must be rendered for the entire audience providing consistent auditory and visual impressions and an enhanced intelligibility of the sound stage.

What is the difference in the setup process between these two system types? Which usability concepts are needed on a mixing desk in order to control such 3D speaker management systems? For which situations is a joystick the best usability choice and which other physical interfaces can be used for controlling the positioning of the 3D sources? How can the source selection process made easy and how are the sources displayed to the user?

We are presenting concepts here and a practical implementation for the efficient use of 3D speaker management for everyday production situations combining spatial sound rendering based on Wave Field Synthesis and a mixing console. The system is capable of spatial sound reinforcement in the entire audience using a comparable number of loudspeakers to standard installations. We present tools and installation procedures that illustrate the simplicity of the installation procedure. Furthermore, the integration of the controls in the mixing desk enables optimum control of the spatial rendering by the sound engineer during the performance in a consistent environment.

1. Introduction

Nowadays, 3D sound systems used in theatres and concert halls are usually fully detached from the main sound reinforcement system. We are presenting here a concept for a fully integrated solution that brings 3D to sound reinforcement for improved segregation of sound sources on stage and coherent audio and visual impressions for the entire audience that can be directly controlled from a mixing console.

The integration alleviates the use of multiple components such as a delay matrix, 32 bands equalizer and loudspeaker management systems. All these functions are available in the Sonic Wave 1 processor with additional dynamic 3D control of multiple sound sources based on Wave Field Synthesis. Positioning information of sound sources can be directly controlled within the mixing desk as a standard function for best efficiency during setup and performance.

In this paper, we first introduce Wave Field Synthesis for practical spatial sound reinforcement using a similar number of loudspeakers compared to today’s installations. We then present concrete applications and references. Tools for configuration of the system and visualization of 3D sources are presented. Finally, we describe in details the integration of controls for 3D speaker management within Studer Vista mixing desk.

2. Wave Field Synthesis for Spatial Sound Reinforcement

2.1. Spatial sound reinforcement

Spatial sound reinforcement aims at providing the audience a spatial impression during live performances. Spatial sound reinforcement puts however heavy requirements on the spatial sound rendering technique being used, imposing the spatial rendering to be effective throughout the audience, hence in a very large listening area.

Different sources need to be reinforced (either live or pre-recorded) that may modify the spatial rendering requirements but also the type of interaction for the sound engineer:

- Live sources with a corresponding real source that should be properly localized on stage:
  - individual sources like voice, guitar (single channel, can move);
  - extended objects like choir, piano, drums, orchestra section (most often static but can move during performance);
- sounds with no visual reference onstage:
  - effects or recorded sounds that may be located on stage or all around the room;
  - reverb (multichannel static object).
2.2. Benefits of Wave Field Synthesis

Wave Field Synthesis (WFS) is a sound field rendering technique that allows for spatial sound reproduction over a large listening area [1][2][3]. WFS aims at recreating the sound field of acoustic sources using an ensemble of loudspeakers positioned at the boundaries of the listening area. The target sound sources can be freely positioned in angle but also distance. WFS therefore allows for the proper reproduction of the parallax effect (angular localization changes based on listener position in space).

The parallax effect is illustrated on figure 1 where two sources located on stage are observed/heard from three different listening positions. It can be seen that none of the listeners observe/hear the sources from the same direction.

![Figure 1: parallax effect for sound reinforcement](image)

The “intensity” of the parallax effect depends on the distance of the source from the audience: the closer the more angular localization differences are observed throughout the audience. The distance of the source can therefore be considered as a homogeneity parameter for the audience, more distance meaning more homogeneous impressions for the entire audience (more similar angular localization) [4].

For sound reinforcement, the reproduction of the parallax effect (distance control of sources on stage) provides coherent visual and auditory impressions for any seat in the audience. It therefore guarantees a consistent localization of sound sources on stage throughout the audience. This is a unique property of WFS that no other reproduction technique can offer within an extended listening area.

The distance of the sources may however be exaggerated to guarantee a more homogeneous level coverage along the width of the hall while limiting the angular perception difference. This may result in localization errors for listening positions opposite to the source position. However, the audience may not notice this mismatch for small angular differences between the auditory and the visual localization. The ventriloquism effect allows a disparity of 5 to 15 degrees depending on the type of sound without breaking the audiovisual integration [5].

Beyond localization accuracy, WFS offers natural spatial auditory cues that are known to favour the streaming process of the auditory system [6] improving the ability of listeners to segregate the different sounds being reinforcement. Spatial reproduction also reduces the masking of concurrent sound sources being played concurrently. This so-called “spatial unmasking” is known to improve the intelligibility of sound stage [7] and is closely related to the well-known cocktail party effect [8].

The naturally enhanced intelligibility provided by spatial sound reinforcement limits the use of EQs and compression during the performance in order to unmask concurrent sources (e.g. voice vs. keyboard), which contributes to the naturalness of the sound reinforcement.

2.3. Practical installations with optimum sound level coverage

For practical sound reinforcement applications, WFS however suffers from three major limitations:

1. The large number of loudspeakers required: traditional WFS imposes a loudspeaker spacing that should not exceed 15 to 20 cm requiring tens to hundreds of loudspeakers even for frontal only installations.

2. The use of compact low efficiency/dynamic range loudspeakers: the close spacing imposes the use of compact loudspeakers that offer limited performance and do not match today’s level requirements for sound reinforcement.

3. A necessary trade off between vertical localization accuracy and level coverage: only 1 horizontal array is used for sound reproduction that can be practically arranged above the stage, providing erroneous localization for front seats, or on/below stage thus ending up in poor level coverage for rear seats due to masking and natural attenuation of the loudspeaker array.

The WFS algorithm integrated in the Sonic Wave I processor alleviates these limitations using a small number of high power/efficiency loudspeakers that are often distributed over two levels (above and on/below the stage) for homogeneous level range distribution and spatial sound reproduction keeping all benefits of WFS. Corteel et al. introduced this concept of “multiple level WFS” for sound reinforcement in [8].

The loudspeakers located on or below stage can be considered as a “spatial” front fill array. The loudspeaker ensemble typically spans the entire stage width using 8 to 16 wide dispersion speakers (5 to 8 inches coaxial or asymmetric directivity speakers with wide horizontal dispersion).

3 to 7 loudspeakers are commonly used above the stage targeting homogeneous level coverage further back in the room. Loudspeakers should either be point source loudspeakers with asymmetric directivity (wide horizontal dispersion, directive behaviour along the vertical dimension) or line arrays with preferably wide horizontal dispersion.
Typical LCR configurations are applicable with two large line arrays on either side of the stage and a central cluster. The stage system can be complemented with surround speakers for positioning sources outside of the stage opening.

Sonic Emotion’s algorithm is also offering full 3D WFS rendering by adding loudspeakers in height. The algorithm only requires a reduced number of loudspeakers, similar to the number of surround speakers. The algorithm shows very accurate height localization accuracy within an extended listening area [11].

3. Practical examples

In this section we present practical usage of WFS for spatial sound reinforcement in various use cases.

3.1. Front of house

The Sonic Wave1 processor has been used for two major summer festivals at the Parc Floral in Paris: the Paris Jazz Festival and Classique au Vert. More than 400000 spectators were present over 2 years (2013 and 2014).

The open-air concert stage is located below a tent with a 1400 seats fan shaped area. A standing area behind the seats can host an additional audience up to 2000 extra people.

The system is composed of a standard LCR line array configuration. 10 8 inch coaxial speakers are used on stage as a WFS front fill to complement the system. The sound engineer Jacques Laville (Urbs Audio1) installed and tuned the installation.

More than 60 shows used the system ranging from solo piano to classical orchestra and jazz ensembles. No adjustments in terms of speaker positions were necessary. Sounds were routed from the mixing desk generating up to 24 streams to the Sonic Wave1 (direct out for solo instruments, groups for extended sound sources or reverb).

Jacques Laville acted as a regisseur and sound engineer for about half of the shows. Jacques Laville had prepared the console layout and introduced the system to visiting sound engineers. They adjusted jointly the source positions on stage to fit with the stage layout. The visiting sound engineer could then work as usual, monitoring all other parameters (levels, Eq, compression) directly from the console, leaving out panning information.

The sound engineers appreciated the transparency of the system and the additional intelligibility that resulted from the spatial rendering. The sound engineers but also the production team, the visiting artists and the audience perceived the additional comfort. A testimonial video was created during the 2013 edition of the Classique au Vert festival (https://www.youtube.com/watch?v=b5ZVJZtckNk).

3.2. Theatre

The sound reinforcement system that has been installed at the Piccolo Teatro in Milan for the 12-scene play "The Panic" written by Argentinian playwright Rafael Spregelburd and directed by Luca Ronconi in 2013 [9]. In 2014, a second production directed by Luca Ronconi (La Celestina written by Fernando de Rojas) was created at the Piccolo Teatro using the Sonic Wave 1 counting up to 50 shows in two years.

Ronconi’s request to sound designer Hubert Westkemper was to enable all members of the audience seated anywhere in the venue to hear each amplified voice of the 16-member cast coming from the position of the actor in question even when moving around the stage.

The theatre has more than 500 seats in the stalls and about 400 in the balcony, while the dimension of the stage is about 22 m x 19 m. The speaker setup consisted in 2 horizontal loudspeaker arrays at different height levels to guarantee an optimum level coverage over the entire audience. The lower loudspeaker array was composed of 24 5 inches speakers (wide high frequency dispersion), with an about 70 cm spacing, positioned along the stage front and covering the front half of the stalls. The front speakers were masked with black acoustical fabric, to ensure no visual localization bias. The array was completed by 2 sub-woofers positioned on the left and right side of the proscenium.

Figure 2: Classique au vert/Paris Jazz Festival layout.

Figure 3: Piccolo Teatro, Milano, The Panic 2013

---

1 http://urbsaudio.fr/
The upper array was installed above the proscenium at a height of about 6.5 m. It was composed of 8 pairs of loudspeakers with a spacing of about 2 m and a minimum distance from the audience of 12 m. These loudspeakers, already part of the sound system of the theatre itself, are more directive (70°×40°). The upper speakers of the 8 pairs were aiming to the balcony, while the lower ones covered the second half of the stalls, partly under the balcony for optimum SPL coverage.

All 18 actors were equipped with wireless microphones and had their movements tracked on stage. All actors had one, some even two, Ubisense tags, detected by 5 sensors positioned around the stage, hence to locate the actors' position. The actor's position was automatically transmitted to the Sonic Wave 1 processor that updated the source position accordingly. 6 additional tracks were used for pre recorded sounds that were located at fixed positions on stage.

3.3. Concert hall

Radio France (Paris, France) has equipped the Studio 105 that hosts a large number of live events with an audience. The existing LCR line array installation plus sub has been complemented with 28 additional loudspeakers (8 inches coaxial): 10 for WFS front fill and 18 surrounding. A temporary installation was using different speakers from September 2013 until September 2014 where the installation was completed.

Due to the diversity of the stations of Radio France, the studio is hosting a large variety of events: radio shows, concerts (pop, rock, jazz, world music, classical music, contemporary music, electro, …) or occasionally conferences (Forum International du Son Multicanal 2012-2014 for which the WFS system has become an essential piece).

The installation combines the possibility of frontal sound reinforcement and surrounding effects/reverb. The surrounding speakers can be used to enhance the room perception by using multichannel reverberation units or by simply duplicating stereophonic reverb over multiple virtual sources located all around the audience. Effects can be in the form of a multichannel content (5.1 or 7.1) that can be reproduced using “virtual loudspeakers” that mimic a multichannel setup or individual objects that can be freely positioned in real time.

The studio is used by up to 15 sound engineers from Radio France who are being trained. Over the past year, the full WFS system has been used in average twice a month for live shows with an audience. The studio is now regularly requested specifically for the use of the WFS system.

4. Control options and interfaces

Two software interfaces have been designed to control the Sonic Wave 1 processor. The Wave Designer, is used for configuration of the system (loudspeaker positioning, output routing, system tuning). The Wave Performer is used during the performance for 3D sound positioning and visualization.

4.1. OSC communication

The Sonic Wave 1 processor can be controlled using the OSC protocol [10], a widely used communication standard designed for musical applications. OSC messages are based on the UDP protocol, and consist of a string header followed by additional arguments. A set of messages is specified to offer control of the audio sources 3D position, level, equalization, naming and routing. In particular, “pair sources” messages enable the user to efficiently control stereo pairs as a single entity, by specifying it’s angular width and position.

The use of such a simple protocol and specification enable users to control the Wave processor with OSC-based programs that let the user create custom interfaces, or patches. Such interfaces can be run either on mobile devices using applications such as Line’s Lemur1 and Hexler’s TouchOSC3, or on desktop devices using standalone software such as Max/MSP4 and PureData5.

4.2. Wave Designer, configuration and tuning interface

The Wave Designer presents itself as a computer-aided design interface used to create loudspeaker setups. The Wave Designer has two modes:

1. An offline mode for loudspeaker positioning and output routing

---

1 available on request at pro@sonicemotion.com
2 http://line.net/en/products/lemur
3 http://hexler.net/software/touchosc
4 http://cycling74.com/products/max
5 http://puredata.info
2. An online mode for system tuning.

4.2.1. Offline configuration

In the offline mode, the user creates a loudspeaker configuration by positioning speakers in 3D space. The speakers can be either created individually, or by generating linear or circular arrays of loudspeakers.

The creation of front of house systems is therefore be made very easy by declaring one line for the WFS front fill speakers indicating the stage width and the number of speakers. The upper loudspeaker array is similarly created as well as eventual additional surrounding speakers.

This basic setup being created, positions of individual speakers can be adjusted to match with realistic installation constraints (stairway, doors, other equipment). Sonic Emotion’s WFS algorithm does not require regular loudspeaker spacing to operate in best conditions.

The speakers can be organized in groups, that we call subsystems. For instance, in Radio France concert hall, four subsystems are declared. Each subsystem represents a specific section of the setup: the WFS front fill, the surrounding speakers, the LCR line array and the sub.

The subsystems are assigned a spatial rendering algorithm, which can be either “WFS main” or “WFS support”. All subsystems in “WFS main” mode are then being grouped to form the principal WFS rendering system that can either be fully peripheric or restricted to the front. Subsystems in “WFS support” mode are considered as complementary loudspeaker arrays that are used to bring energy at further locations in the room. Such a “support array” can be the LCR line array configuration and additional under balcony loudspeaker arrays. An adapted Wave Field Synthesis algorithm drives these “support arrays” for creating a consistent spatial sound rendering in support to the main WFS array.

Loudspeakers are assigned to physical outputs in a simple routing matrix. For multi-way loudspeakers, one physical input is assigned for each frequency range (low, mid, high).

Once the configuration is created, it is uploaded to the processor that creates the right number of speakers and outputs and calculates WFS filtering for reducing online pre-processor workload. Prior to the upload, the Designer performs a sanity check of the setup to validate the configuration and indicate to the user the potential errors.

4.2.2. Online tuning

The Wave Designer turns into online mode for further setup & tuning.

A simple test interface offers an automatic or manual tool for testing individual loudspeakers and routing by selecting a given input of the system and routing it directly to a given speaker/output.

The online mode of the Wave Designer contains a complete parametric equalizer for each loudspeaker with high and low shelving filters, high and low pass filters with up to 48 dB slope, and 8 peak filters. Furthermore, each parametric equalizer can be adjusted according to 3 different modes:

- individually per loudspeaker;
- semi-globally by subsystem (LCR, WFS front-fill, surround, …);
- globally for all loudspeakers.

This way, an ensemble of speakers can be tuned very efficiently, offering a comprehensive grouping of speakers of similar type (line arrays or point source) and located in similar acoustical conditions (simply put on stage, hang in height, attached to walls, …).

The equalization can be set manually or by loading equalization presets (from loudspeaker manufacturers or stored by the user).
system. These settings can also be stored with the loudspeaker equalization preset.

An additional panel offers simple level and delay adjustments for each subsystem allowing aligning/balancing the main WFS array and the support subsystems.

The complete configuration (loudspeaker positioning, output routing, equalization, loudspeaker management, level/delay alignment) is stored as a preset for archiving the configuration and loading it at a later time or on a different machine.

4.3. Wave Performer, performance visualizer interface

The Wave Performer serves several purposes: it offers 2D and 3D displays of the scene, and enables sound engineers to move audio sources during a performance.

In order to clarify the sound scene, sources can be grouped, hidden, assigned names and are presented as a list on the right side of the interface.

![Figure 7: Wave Performer source positioning interface](image)

The Wave Performer also includes a master equalization, which can be used to adapt the overall sound colour of the installation. The equalization is fully independent from the one offered in the online mode of the Wave Designer and keeps the configuration unaltered.

This is useful to sound engineers who tour with the same equipment. Moreover, when various sound engineers use the system, for instance, in festivals, the master equalization is used to provide each sound engineer with a specific sound coloration. It can fully replace a typical 32 band graphical equalizer of a classical setup.

Additionally, each source can be directly routed to one or several loudspeakers and/or subsystems, in a routing module. This allows applying a specific audio rendering to each source, in order to adjust the system to various sound reinforcement situations. Finally, the Performer provides a source level mixer with a display of the input levels.

Each performance configuration is storable as an xml preset file, so that user-specific configurations can easily be recalled.

5. Mixing desk integration

As outlined above a 3D speaker management system is an independent (sub-) system within a live performance setup. Especially the configuration tasks are normally carried out completely with the user interfaces of the speaker management system. In case of a 3D speaker management system though a tight integration with the mixing desk significantly simplifies the operation of the complete system during the performance. In the following the essential control functions and integration concepts for an effective operation during the live performance are presented.

5.1. Retrieve source signals; naming of sources

In order to do a proper channel setup on the mixing desk the list of all (signal) sources needs to be retrieved from the 3D speaker management system. Since the mix is normally organized on the mixing desk the naming (labelling) of these ‘shared’ sources needs to be pushed back from the mixer to the 3D speaker management system. The 3D speaker management system imports and displays these labels in the graphical user interface (GUI) of the performance visualizer tool. By these means the sources can be recognized with the same name/label in both systems.

Selecting a source in the performance visualizer tool is not quite straightforward. The sources are normally presented as small circles (or similar marks) with limited naming information. Moreover, several sources are often positioned at almost the same place in the sound space, which makes it very difficult to distinguish them in the GUI. Zooming functions or browsing through a list can help but it is also rather cumbersome in order to quickly select a source in the performance visualizer.

The natural place for the source selection is on the mixing desk where a configurable and a cue recallable strip setup organizes the channels, which are important for a specific scene.

In addition a lot of times the sound engineer does not only want to change the spatial position of a source but also controls other channel processing parameters. The mixing desk therefore distributes the channel selection information to the 3D speaker management system and this source selection information is then displayed within the performance visualizer tool.

5.2. Positioning of sources

As outlined above there are different source types that have different interaction requirements in order to do the spatial positioning: the individual sources, the extended objects and the effects.
The simplest use case is the positioning of the individual sources (single channel). For this source type we propose to patch the direct out of the input channel to the corresponding source in the 3D speaker management system. The standard panning control of this input channel can be reused to calculate the x/y (or azimuth/distance) coordinates within the 3D sound space. A new ‘3D positioning mode’ is added to the panning control and when activated the mixing desk sends the x/y coordinates to the 3D speaker management system. By these means the user has to control only one panning model if the same source signal is also used in a mix within the mixer.

Figure 7: 5.1 panning control in 3D mode.

Using rotaries allows positioning the sources very accurately in the sound space – the sound engineer can change the parameters with a precise tactile control, gets the acoustic feedback from the sound system and can see the source positioning in the performance visualizer tool. As soon as a panning rotary control is touched the mixing desk automatically sends a corresponding ‘source selected’ message to the 3D speaker management system and the user can see, which source position is changed. The rotaries are also ideal for controlling rather precise but slow movements of these sources in a scene.

The joystick, which controls the same panning parameters, is better suited for the control of rather fast and dynamic spatial effects or for the outline positioning of a source. Due to the limited resolution and control area of the built in joysticks of mixing desks an exact positioning of a source is cumbersome though.

Besides these individual sources the 3D speaker management system should also be capable of controlling extended objects or pre-panned/stereo sources. An example of such a pre-panned source would be a choir. The different voices are routed to a stereo group and each voice has its stereo-panned send signals. The picture below shows the setup of these voices within a choir.

Figure 8: Example of a pre-panning for a choir.

The extended object (or stereo group) positioning within the 3D speaker system not only needs the x/y positioning information but also the width and the rotation parameters as depicted below.

Figure 9: Width and rotation parameters; example for a movement of the choir in a scene

If the choir would walk from the back to the front of a stage in a certain scene all of these parameters would have to be controlled accordingly. Depending on the speed of such a movement the control through rotaries is still possible. In a programmed show the crossfade function of the mixing desk automation system can be used defining only the start and end positions and the crossfade time.

If the z-axis is added to the sound space the control of the source positioning gets even more complicated. For the
rather static source positioning the rotaries can still be used. But (faster, dynamic) movements in the x/y/z space are not easy to control with the rotaries only. One approach is to combine the joystick (for the x/y positioning) with a rotary control (z-axis positioning). Another option could be the use of systems, which recognize hand gestures. But these systems lack of a tactile feedback for the user. This is the root cause for most operation errors in such systems. Ultra haptic devices are one solution to this problem. These devices create a haptic feedback for the user with the aid of ultra sonic waves.

Figure 10: Example of an ultra haptic device in laboratory application.

First research projects investigate this technology – but it is too early to evaluate whether these devices are the right choice for controlling sound source movements in the x/y/z sound space. The resolution of these devices is promised to be similar to today’s GUI controls (button, slider, rotary, fader).

5.3. Snapshot and automation

A speaker management system should feature an independent snapshot module (create, update, recall, delete snapshot), which can be accessed through a public control interface. Such an interface is perfect for changing configuration or setup parameters of the speaker management system during scene changes. But the automation of all relevant parameters for the source positioning is simpler within the mixing desk. The source positioning parameters become part of the native snapshot and automation module of the mixing desk and inherit functions like isolation (masking) and parameter crossfade. By these means the interface between the speaker management system and the mixer is simpler (lightweight) and the overall snapshot recall performance is more precise. Editing the snapshot data offline is also simplified by this approach since the user has to work only in one system.

5.4. Interfaces

The integration of a 3D speaker management system to a mixing desk requires an audio and a control interface. Since the signal sources for the speaker management system all stem form the mixer a simple point-to-point connection (MADI) is appropriate. An AoIP interface might allow a more dynamic selection of input signals in a later stage.

A simple control protocol should be designed so that both systems are not coupled to tightly with each other and so that the line up (debugging) is easy. OSC is a standard control protocol, which is widely used for third party control in the professional audio domain and offers all the means in order to implement the functions of the integration interface outlined above.

Figure 11: system diagram with 3D speaker management and mixing system.

6. Conclusion

An effective integration of a 3D speaker management system into the workflow of the mixing desk was presented. It is essential for a simple operation of the overall system during a live performance. A comprehensive interface displays the overall sound stage showing either all or a limited subset of sources in a 2D and a 3D view.

The most important aspect is the management of the source positioning parameters using standard controls of the console surface.

The configuration and tuning of the system is realized very efficiently by easily providing loudspeaker positioning information. The system does not require aligning gains and delays of a large number of individual loudspeakers but only loudspeaker ensembles (so-called sub-systems). Such simplification results in faster setup times and more flexibility during the use of the system. Configurations can be prepared offline and recalled or modified during
installation. This includes loudspeaker positioning/routing but also system tuning information (eq, loudspeaker management, system alignment).

The 3D speaker management system also offers the possibility of spatial sound reinforcement for enhanced intelligibility of the sound stage and reduced masking of competing sources, ensuring an improved overall sound quality.

7. References


