Plug-in Suite for Mastering the Production and Playback in Surround Sound and Ambisonics

Matthias Kronlachner
University of Music and Performing Arts, Graz
Email: mail@matthiaskronlachner.com
Web: http://matthiaskronlachner.com

Gold award at the
AES Students Design Competition
136th AES Convention Berlin, April 2014
Category 2 - Graduate Level

***Abstract*—The plug-ins presented here largely simplify the creation and playback of surround sound productions by providing user-friendly access to Ambisonic techniques taken from the most recent research. Hereby, sound designers and composers, the essential end users, can easily employ the newest ways to create, manipulate, and play back Ambisonic recordings in variable spatial resolutions. Additionally, the multichannel plug-in suite provides frequently-used equalization features such as levels, delays, and convolution filters to an arbitrary number of channels within one Digital Audio Workstation. The plug-in suite has already been successfully employed in several performances with various (hemi-)spherical playback facilities (4...43 loudspeakers), in the technical support of computer musicians who deal with spatial music, and in the production of Ambisonic surround recordings.**

I. Introduction

Spatial audio has not only been a hot research topic but an integral part of cinema sound for many years. The new generation of cinema surround sound technologies include loudspeakers on different elevation levels thus allow for the playback of full periphonic surround content. While in private homes such extensive loudspeaker installations are still rare a major amount of music and movies is nowadays consumed through headphones. This work addresses this topic by not only providing playback of surround sound over loudspeakers but also through headphones.

For several years research institutions have been experimenting with loudspeaker arrays in two- and three-dimensional setups. Many individual software solutions were developed that support the creation of content for these loudspeaker arrangements. While most Digital Audio Workstations (DAWs) provide tools to produce for standard surround setups such as *5.1* only, more extensive loudspeaker setups (eg. fig. 1) clearly require tools for a more flexible channel routing. To overcome the limitation of DAWs, most toolkits for spatial audio were developed in graphical programming environments, such as Pure Data and MaxMSP, which basically have no restriction concerning the number of channels and offer quick development and maintenance cycles. However, not only is the integration of such external spatialization toolboxes into a DAW challenging for the developer, these toolboxes are often also challenging to use for non-experts. Audio channels and control data have to be sent between the applications and complex routing scenarios need to be re-established everytime. This frequently causes difficulties when sharing projects and due to the loss of the ability to render audio tracks offline.

The plug-in suite is meant to get rid of the above-mentioned difficulties. It has been developed as an open source project in C++ using the *JUCE*[1] framework which supports building audio plug-ins in all major formats and standalone audio applications for Windows, MacOS and Linux. Currently the author provides VST and standalone binaries for Windows and MacOS. Using flexible DAWs such as Reaper or Ardour allows for a single software controlling the entire production cycle starting from panning the source signals, mixing microphone array recordings until decoding the playback signals and adjusting the loudspeaker signals.

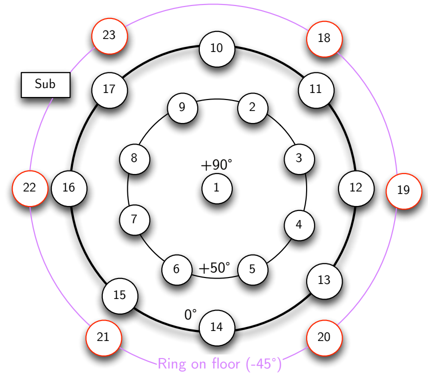

Fig. 1. 23-loudspeaker layout of an examplary Ambisonic concert venue in which the plug-ins have been employed.

II. Ambisonics

Ambisonics can be used to create spatial audio productions for circular or spherical loudspeaker arrangements. In contrast to channel-based standards in surround sound such, as *5.1*, it offers flexibility regarding the loudspeaker setup around the listening area. This makes the production stage largely independent of the loudspeaker arrangement. Furthermore, Ambisonic content can be recorded using its classical and new main microphone array technology.

Ambisonics is based on the expansion of the surround signal into spherical harmonics up to the order N that defines the angular resolution. A full periphonic (3D) Ambisonic signal set of order N results in $(N+1)^2$ channels [1].

[1] http://www.juce.com

III. *ambix* AMBISONIC PLUG-INS

The maximum Ambisonic order N for all *ambix* plug-ins can be defined at compile time and is only limited by the maximum number of channels per track of the host. Currently Reaper supports 64 channels per track resulting in the maximum usable Ambisonic order $N = 7$.

A. Encoding

Ambisonic recordings can be synthesised by encoding (panning) sound sources to specific *azimuth* and *elevation* angles. Each mono or multichannel track in the DAW can be encoded in Ambisonics using *ambix_encoder* and controlled by using the graphical user interface (Fig. 2) or built in automation features of the host.

Keeping track of the position of a number of sources can be challenging and the user might want to have an alternative to opening the graphical user interface of each encoder individually. To overcome this problem, the plug-in has built in Open Sound Control (OSC) functionality, which allows to control the source position from outside the host application (Fig. 3) [2]. An automatically launched application monitors and controls the position of several encoders at once. During a mixing session, the user can keep the overview of the spatial sound scene. It is just a matter of taste whether this application may run on the same computer or another device connected via network.

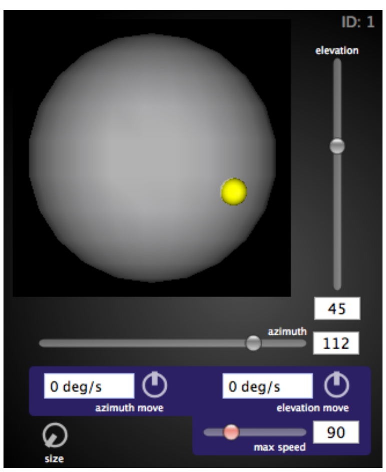

Fig. 2. *ambix_encoder* is used to pan a sound source on the surface of a sphere

B. Manipulation of the spatial image

It might be necessary to adapt the spatial image of a given sound scene after the production has already been rendered to Ambisonic signals or whenever it is based on Ambisonic recordings. One reason might be to compensate for rotated or mirrored loudspeaker arrays during playback, which is an often experienced problem due to mixed up coordinate systems. Moreover during production, a directional loudness modification for microphone array recordings might be needed [1].

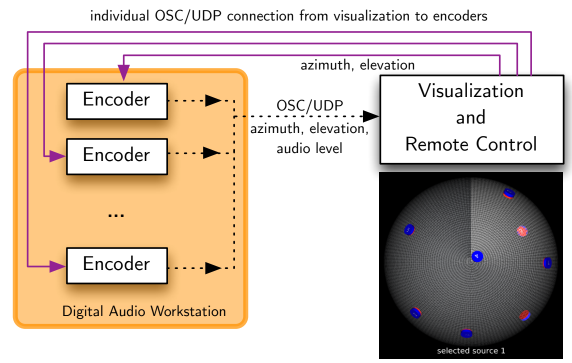

Fig. 3. bi-directional OSC communication for monitoring and controlling multiple encoders, this application may run on the same computer or on a different device connected via network.

1) Warping: *ambix_warp* allows to distort the spatial image towards, or from certain directions on the sphere. Warping the spatial image towards the equator might compensate for an elevated spatial image during playback [1][3][4]. (Fig. 4)

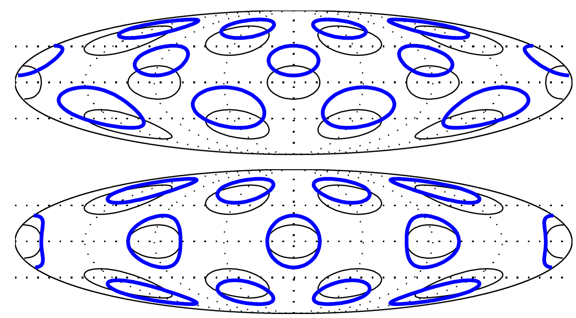

Fig. 4. Warping scheme, thin lines indicates unmodified surround image, warping towards the northpole and warping away from equator.

2) Directional loudness modifications: Most microphone arrays are designed to record the environment without emphasising certain directions. *ambix_directional_loudness* can be used to attenuate or amplify certain regions in the surround recording. This plug-in allows to define several circular or rectangular regions on the surface of the sphere and apply a gain factor to that specific region [1]. This can be used to emphasise the sound from certain directions (eg. instruments) or attenuate the reverberant sound which might reach the microphone array from the back.

3) Source widening: By another new tool, virtual sound sources can be given a width/diffuseness control parameter. The plug-in *ambix_widening* allows to gradually increase the diffuseness of the Ambisonic surround image. The algorithm applies a frequency-dependent rotation, yielding a frequency-dispersed direction of arrival (Fig. 5) [5]. The widening can also be used to synthesise diffuse early reflections.

C. Metering

It is decisive to maintain overview of the audio scene levels when mixing. The necessity became obvious while developing and trying to verify the above mentioned spatial manipulations.

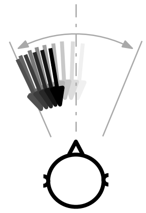

Fig. 5. Frequency-dispersed direction of arrival as Ambisonic widening effect [5].

This led to the development of a surround metering plug-in with rms and peak meters distributed on the surround sphere [1]. A two dimensional representation of the spherical soundfield is obtained by Mollweide projection [6]. (Fig. 6)

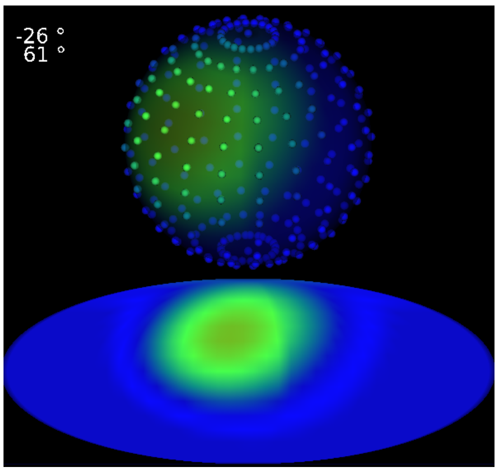

Fig. 6. Ambisonic metering, spherical texture represents the rms level, balls show the peak level.

D. Decoding

An Ambisonic decoder generates loudspeaker signals through linear combination of the individual signal components. Correct decoding for specific loudspeaker arrays is essential for good results with Ambisonics and a challenging task. Computing an appropriate decoder matrix is currently not part of this software and is right now done using a numerical computing software. For a future release the inclusion of an automatically calculated AllRAD decoder [7] from user-definable loudspeaker coordinates is planned.

1) Decoding to loudspeakers: The loudspeaker decoder *ambix_decoder* (Fig. 7) relies on a collection of configuration files containing a decoder matrix for a specific loudspeaker arrangement. The package includes presets for several standard surround configurations. Existing toolboxes by other authors can be used to calculate a decoder matrix from given loudspeaker coordinates. The Ambisonic Decoder Toolbox [8] from Aaron Heller supports writing *ambix_decoder* preset files and can be downloaded for free[2].

2) Decoding to headphones: The binaural decoder *ambix_binaural* generates virtual loudspeaker signals which are

[2]https://bitbucket.org/ambidecodertoolbox/adt.git

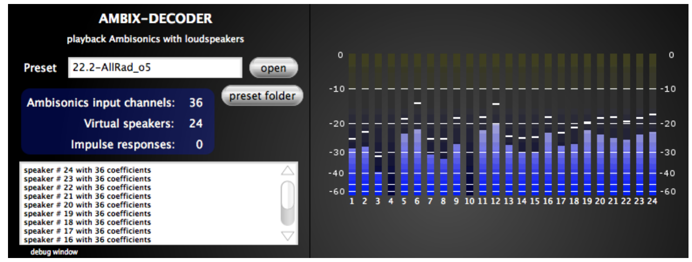

Fig. 7. *ambix_decoder* is used to drive loudspeakers from Ambisonic signals.

then convolved with their associated binaural loudspeaker room impulse responses (Fig. 8). The software includes ready to use binaural decoder presets using the measured impulse responses from venues with 24 up to 46 loudspeakers [9]. (Fig. 9) This allows sound designers and composers to prepare their surround pieces before they enter the actual venue and safe rehearsal time. In combination with a head tracking system and the rotation plug-in *ambix_rotator* an even more realistic simulation can be achieved by taking into account the head movements of the listener.

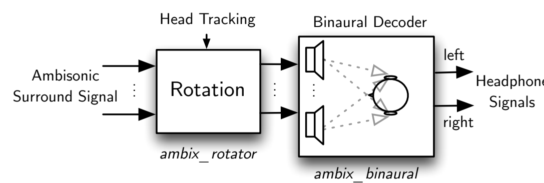

Fig. 8. Listening to Ambisonic surround recordings using headphones and improving localization by employing head tracking.

E. Converting to other Ambisonic conventions

Ambsonics suffers from different existing conventions regarding the sequence and weighting of the signal components [10]. To retain compatibility to software and recordings that are using different conventions *ambix_converter* (Fig. 10) has been developed which is capable of converting between all coexisting Ambisonic conventions.

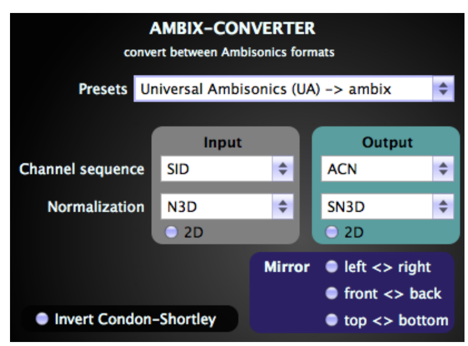

Fig. 10. *ambix_converter* is used to convert between different Ambisonic conventions and to retain compatibility with other software and recordings.

IV. *mcfx* MULTICHANNEL PLUG-INS

Most audio effect plug-ins are restricted to process two channels only, some few can be used for standard surround formats such as *5.1*. While working with multichannel loudspeaker or microphone arrays there is a need for a toolbox that

Fig. 9. Two examples for surround concert venues where the software was already used, *ambix_binaural* includes the binaural impulse responses of each individual loudspeaker and allows to prepare surround recordings before having access to the venue, thus saving time in the mostly busy performance space.

provides basic audio manipulations. Out of this need some important audio processors have been developed without a limitation about the number of input and output channels. The maximum number of channels for the *mcfx* plug-ins can be chosen at compile time and is restricted only by the capabilities of the host application. The following section will outline some classical as well as new areas of application for these multichannel plug-ins.

A. Equalization

mcfx_filter offers a high/low pass, high/low shelf and two peak filters. The filter can be used to modify the sonic characteristic of microphone arrays, Ambisonic recordings or as phase-aligned Linkwitz-Riley multichannel crossover. (Fig. 11, 13)

B. Gain and Delay Compensation

The *mcfx_gain_delay* plug-in can be used to compensate for time alignment or unequal loudness within a loudspeaker array (Fig. 11). Each channel passing the plug-in can be independently scaled and delayed.

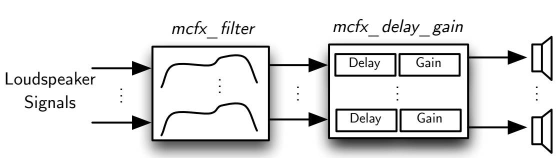

Fig. 11. Using the *mcfx* plug-ins for individual loudspeaker equalization.

C. Convolution

The *mcfx_convolver* plug-in is a multichannel convolution matrix that uses an existing efficient open-source parted FFT convolution algorithm. An arbitrary number of impulse responses can be loaded and assigned to specific inputs and outputs. The convolver plug-in is currently used to drive an icosahedral loudspeaker array from 16 Ambisonic channels [11] (Fig. 12), to encode the signals of a 32 channel microphone array into 25 Ambisonic channels [12] (Fig. 13), and to add Ambisonic reverb to recordings.

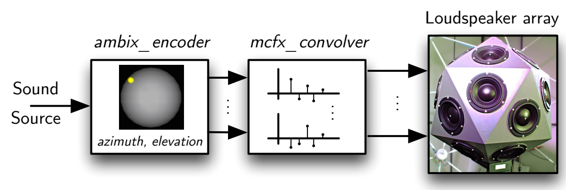

Fig. 12. Using the encoder and convolver to control the sound radiation of an icosahedral loudspeaker.

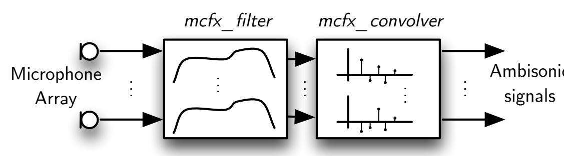

Fig. 13. Adjusting the sonic characteristic of a microphone array and encoding to Ambisonics.

V. Conclusion

A suite of cross-platform Ambisonic and multichannel effect plug-ins has been presented. The software allows to use extensive multichannel microphone and loudspeaker arrays within Digital Audio Workstations. By keeping the whole production within one host, the editing and offline rendering functionality of DAWs can be used to efficiently create and share immersive surround productions. The software has already been successfully used by professionals and students for the production of 3D surround sound and formed an integral part of the concert setup in various venues.

References

[1] M. Kronlachner and F. Zotter, "Spatial transformations for the enhancement of Ambisonic recordings," in *2nd International Conference on Spatial Audio*, Erlangen, 2014.

[2] M. Kronlachner, "Ambisonics plug-in suite for production and performance usage," in *Linux Audio Conference*, Graz, 2013.

[3] F. Zotter and H. Pomberger, "Warping of the Recording Angle in Ambisonics," in *1st International Conference on Spatial Audio*, Detmold, 2011.

[4] F. Zotter and H. Pomberger, “Warping of 3D Ambisonic Recordings,” in *Ambisonics Symposium*, Lexington, 2011.

[5] M. Kronlachner F. Zotter, M. Frank and J.-W. Choi, “Efficient Phantom Source Widening and Diffuseness in Ambisonics,” in *EAA Symposium on Auralization and Ambisonics*, Berlin, 2014.

[6] B. Bernschütz, “Map Projections for the Graphical Representation of Spherical Measurement Data,” in *38th German Annual Conference on Acoustics*, Darmstadt, 2012.

[7] F. Zotter and M. Frank, “All-Round Ambisonic Panning and Decoding,” in *J. Audio Eng. Soc., Vol. 60, No. 10*, 2012.

[8] A. Heller and E. Benjamin, “The Ambisonic Decoder Toolbox,” in *Linux Audio Conference*, Karlsruhe, 2014.

[9] F. Hollerweger and M. Rumori, “Production and Application of Room Impulse Responses for Multichannel Setups using FLOSS Tools,” in *Linux Audio Conference*, Graz, 2013.

[10] E. Deleflie C. Nachbar, F. Zotter and A. Sontacchi, “ambiX - A Suggested Ambisonics Format,” in *Ambisonics Symposium*, Lexington, 2011.

[11] F. Zotter, *Analysis and Synthesis of Sound-Radiation with Spherical Arrays*, Ph.D. thesis, University of Music and Performing Arts, Austria, 2009.

[12] S. Lösler”, “Schallfeldspezische Entzerrung bei Radialfiltern begrenzter Dynamik fur das Eigenmike,” in *Project thesis*, Graz, 2013.