Ambisonics plug-in suite for production and performance usage

Matthias KRONLACHNER
Institute of Electronic Music and Acoustics, Graz (student)
M.K. Čiurlionio g. 5-15
LT - 03104 Vilnius
m.kronlachner@gmail.com

Abstract

Ambisonics is a technique for the spatialization of sound sources within a circular or spherical loudspeaker arrangement. This paper presents a suite of Ambisonics processors, compatible with most standard DAW[1] plug-in formats on Linux, Mac OS X and Windows. Some considerations about usability did result in features of the user interface and automation possibilities, not available in other surround panning plug-ins. The encoder plug-in may be connected to a central program for visualisation and remote control purposes, displaying the current position and audio level of every single track in the DAW. To enable monitoring of Ambisonics content without an extensive loudspeaker setup, binaural decoders for headphone playback have been implemented.

Keywords

Ambisonics, Plug-ins, Digital Audio Workstations, Binaural, Ardour

1 Introduction

It turns out to be difficult finding platform independent audio plug-ins for encoding and decoding Ambisonics. Following section should give a brief overview of still maintained plug-ins.

Fons Adriaensen's AMB-plugins[2] offer encoders and rotators (yaw axis only) until 3^{rd} order. LADSPA can be used with the DAW Ardour under Linux and Mac OS X. There is no Windows host supporting LADSPA. For decoding Ambisonics signals into loudspeaker feeds, Adriaensen's Jack client application AmbDec [Adriaensen, 2005] is often used.

Bruce Wiggins offers his WigWare[3] plug-ins in VST format for Windows and MacOS X. These processors include 2D and 3D encoders until 3^{rd} order as well as 1^{st} order decoders for several fixed loudspeaker arrangements.

Daniel Courville's[4] Audio Unit (Mac OS X only) plug-in suite offers 3D encoders for 1^{st} and 2^{nd} order as well as 2D encoders for 5^{th} order.

For the plug-ins created in this work, the C++ cross-platform programming library JUCE[5] [Storer, 2012] has been used to develop audio plug-ins compatible to most DAWs on Linux, Mac OS X and Windows. JUCE is being developed and maintained by Julian Storer, who used it as the base of the DAW Tracktion. JUCE is released under the GNU Public Licence. A commercial license may be acquired for closed source projects. It is possible to build JUCE audio processors as LADSPA, VST, AU, RTAS and AAX plug-ins or as Jack standalone applications. LV2 (LADSPA version 2) support currently has to be added manually from the separate project DISTRHO[6] [Coelho and Rodrigues, 2012].

Ambisonics suffers from different existing standards concerning channel order and normalization. To overcome this problem, a conversion tool is included in the plug-in suite. Encoders, rotators and decoders from the authors suite are designed for the ACN channel order and SN3D normalization, proposed by [Nachbar et al., 2011]. Conversion between standards of the input and/or output format can be done by the conversion plug-in.

Apart from platform compatibility, some considerations about usability did result in features of the user interface, not available in other surround panning plug-ins. Continuously rotating a sound source results in a discontinuity of the angular representation between $-180°$ and $+180°$. This jump is also reflected when drawing automation curves resulting in a mismatch between the visual representation and auditory perceived movement of a sound source. A solution allowing to define absolute starting points and angular velocities for relative movements is

[1] Digital Audio Workstation
[2] http://kokkinizita.linuxaudio.org
[3] http://www.brucewiggins.co.uk
[4] http://www.radio.uqam.ca/ambisonic/
[5] http://www.rawmaterialsoftware.com
[6] http://distrho.sourceforge.net

proposed.

For headphone monitoring several binaural decoders have been implemented simulating the Ambisonics half sphere of the medium sized IEM Cube with 24 speakers and the concert hall Mumuth[7] with 29 speakers in a elliptical stretched half sphere.

For visualization and external control purposes, a bidirectional Open Sound Control (OSC) communication layer has been implemented in the encoder plug-in. An external program is able to display the current position and audio level of every track in the DAW. The visualization program may also take control over the sources. This can be very useful in performance situations while having a multitrack playback coming from the DAW and a central display to control the position of the individual tracks.

Currently no audio plug-in format can handle dynamic input/output channel counts. Therefore all plug-ins may be compiled for fixed Ambisonics orders.

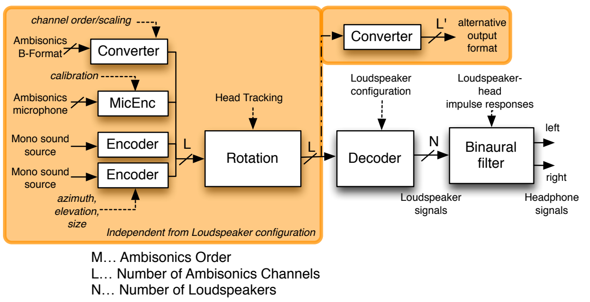

Figure 1: Ambisonics production and playback chain

2 Encoder

The encoder plug-in distributes a mono source signal into Ambisonics channels (derived from spherical harmonic functions), according to *azimuth* and *elevation* settings, representing the coordinates on a spherical surface. An additional parameter called *size* $[0,1]$ may be used to adjust the spatial directivity. Adjusting the *size* parameter from *zero* towards *one* results in a scaling of the higher order components. For a *size* setting of 1, all Ambisonics channels will be zero except the 0^{th} order (also known as W channel), resulting in an equally distributed signal over all loudspeakers.

[7]http://www.kug.ac.at/en/studies-further-education/studies/infrastructure/the-mumuth.html

2.1 Automation parameters

Most current DAWs are limited to represent automation parameters between 0.0 and 1.0 along a time line. Panning plug-ins usually map this range for azimuth and elevation between $-180°$ and $+180°$. A full circle rotation results in a mismatch between the visual representation (Fig. 2) of the automation curve and the perceived continuous rotation of the source. Additionally the plug-in host may interpolate between a jump from 1.0 to 0.0, resulting in a very fast audible jump.

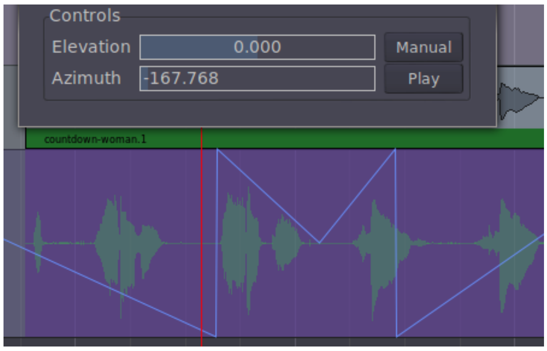

Figure 2: Automation curve for full circle rotation using Ardour and AMB plug-ins

To overcome this problem, automation parameters (Fig. 3) have been added for setting start points *(SetAzimuth, SetAzimuthRel)* and angular velocities *(tgl-rot-azimuth)*. The maximum speed of the angular velocity may be adjusted between 0 and $360 deg/sec$ by an additional parameter (*max-speed*). This guarantees a wide range of adjustment and at the same time accuracy for the rotation speed.

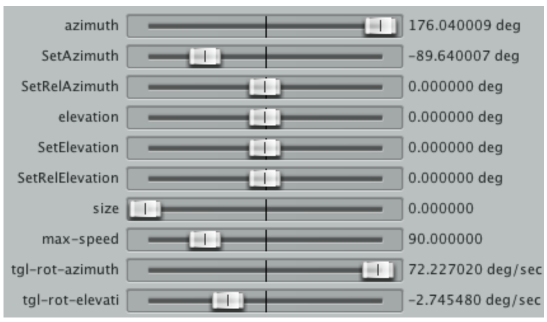

Figure 3: Encoder automation parameters

2.2 Remote control and visualization

Keeping track of all sound source positions within a DAW may be a difficult task for the mixing engineer. To allow a better overview and control of the spatial scenery, a cooperative visualization and control unit has been implemented. All encoder plug-ins are equipped with a bidirectional OSC layer (Fig. 4), sending and receiving control and status parameters. Currently a functional prototype has been implemented in Pd/GEM (Fig. 5) displaying all tracks (encoders) on a sphere, including visualization of their audio levels. Currently the audio level is represented by the variable length of the cylinder representing a source signal. This concept may be extended to a more sophisticated implementation.

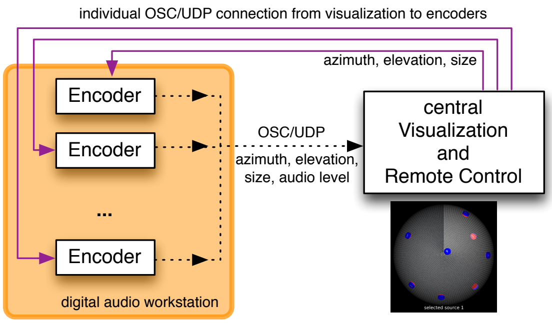

Figure 4: Encoder OSC communication with remote visualization and control

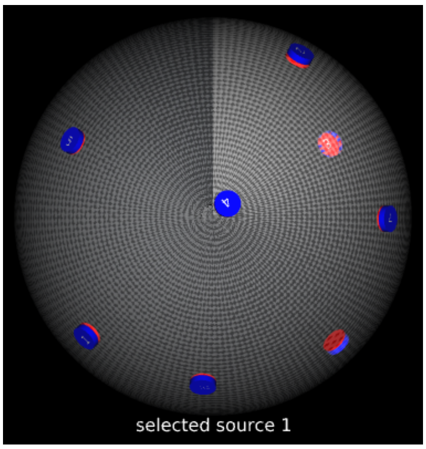

Figure 5: Visualization and remote control of several encoders with Pd/GEM

3 Rotator

The rotation plug-in (Fig. 6) may be used to manipulate the orientation in the Ambisonics domain, as described in [Musil et al., 2003]. An optimized way to calculate rotation matrices can be found in [Rumori, 2009]. Therefore yaw, pitch and roll can be adjusted by ±180°. This is very useful for the incorporation of head movements during binaural playback. The rotation plug-in listens to an adjustable UDP port for incoming OSC messages. This allows to bypass host automation and controlling the rotation directly from the head tracking software (Sec. 4.1).

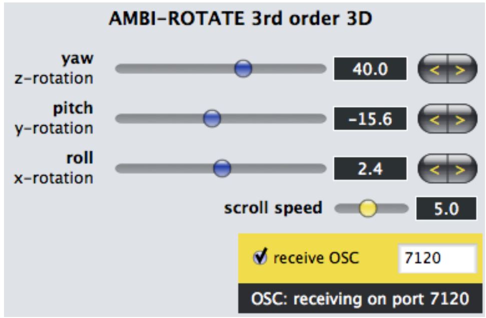

Figure 6: Ambisonics Rotator

4 Binaural decoder

The binaural decoder computes virtual loudspeaker feeds as a linear combination of the Ambisonics signals according to a given decoding matrix. These virtual loudspeaker signals are convolved with their individual stereo impulse responses, modeling the transfer path from the loudspeaker position to the left and right ear of the listener (Fig. 1).

4.1 Head tracking

Head tracking is a significant feature for virtual reality scenes and headphone playback. Small head movements change the relative position of a sound source in aspect to the listeners ears, making localization more easy and removing ambiguity.

The *Kinect*^{TM} sensor as add-on for the gaming console *XBox 360*^{TM} by Microsoft offers a low budget depth sensor. [Fanelli et al., 2011] developed a software to gather head orientation angles and the head position relative to the Kinect sensor (Fig. 7). The author extended this head pose estimation software about sending OSC data to the Ambisonics rotator plug-in[8]. The Ambisonics sound field has to be rotated in opposite direction to keep the sound source positions fixed and suppress the rotation with the listeners head.

[8] `http://github.com/kronihias/head-pose-estimation`

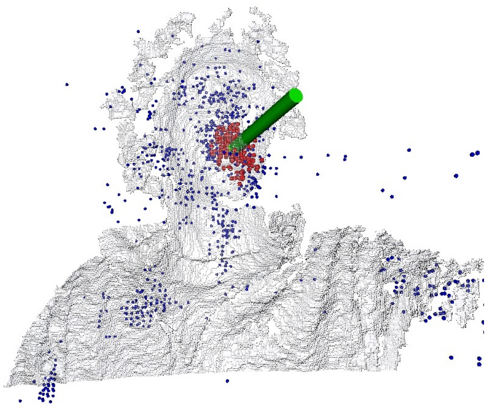

Figure 7: Head pose estimation

4.2 Matrices and virtual venues

To allow monitoring in various virtual scenes, different decoding matrices and sets of loudspeaker-to-head impulse responses were used.

4.2.1 Reduced decoder (3rd order 3D)

This binaural decoder uses specific symmetry relations between the HRTFs and decoder matrices to achieve a reduced set of head releated impulse responses, directly applicable on the Ambisonics Signals without computing virtual loudspeaker signals[Musil et al., 2007]. This approach reduces the numbers of convolutions and therefore the CPU load. The algorithm is implemented in the `iem_bin_ambi` library for Pure data.

4.2.2 IEM Cube and Mumuth decoder (4th and 5th order 3D)

The $120m^2$ IEM Cube (Fig. 8) serves as the main lab of the Institute of Electronic Music and Acoustics Graz. 24 Tannoy coaxial speakers are mounted in a hemispherical arrangement consisting of three rings (12 - 8 - 4). During the years of Ambisonics research at the IEM, several different approaches for finding an optimal decoder matrix have been taken [Zotter et al., 2012; Sontacchi, 2003]. The IEM Cube binaural decoder implemented in the authors Ambisonics plug-in allows to switch between the different available decoder matrices.

The Mumuth (Fig. 9) was opened in 2009 as multi purpose venue for the University of Music and Performing Arts Graz. It houses the $600m^2$ *György Ligeti* concert hall. The 33 Kling&Freitag CA 1001 SP loudspeakers (29 are used for the half sphere, 4 more speakers are located in the corners) may be arranged by a special motor controlled mounting, resulting in a versatile loudspeaker setup that can be changed within a minute. The Mumuth binaural decoder uses the Ambisonics decoder matrix and loudspeaker setting for the IEM Demosuite [Plessas and Zmölnig, 2012] by Thomas Musil.

The impulse responses of the IEM Cube and the Mumuth were recorded by Martin Rumori and David Pirrò as part of the artistic research project *The Choreography of Sound (CoS)* [Eckel et al., 2012].

Figure 8: IEM Cube

Figure 9: Mumuth

5 Ambisonics converter

Since the beginning of Ambisonics in the 1970s and the extension to Higher Order Ambisonics, several different approaches for arranging and normalizing the spherical harmonic components have been taken. A good summary may be found in [Nachbar et al., 2011] and [Chapman et al., 2009]. All plug-ins in this suite are operating with the ACN channel order and SN3D normalization, proposed in [Nachbar et al., 2011]. The Ambisonics converter plug-in allows to interchange the normalization schemes SN3D, N3D, Furse-Malham and the channel order schemes ACN, SID and Furse-Malham. Thus, it is possible to incorporate various standards into the production chain.

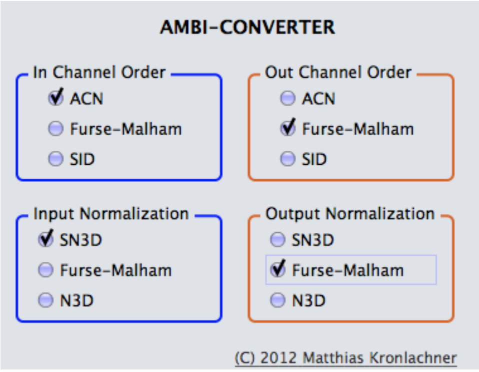

Figure 10: Converter plug-in

6 Summary

This paper presented an Ambisonics processor suite, usable in production or live performance. The JUCE framework proved to be a stable solution for developing cross platform audio plug-ins and standalone applications. By the time of writing this paper, no Mixed-Order systems have been implemented. Mixed-Order Ambisonics uses different orders for the horizontal and the vertical part of the sound field. It is planned to incorporate this into the converter plug-in.

7 Acknowledgements

The author would like to thank Winfried Ritsch, who advised the roots of this work as bachelor thesis. Many people from IEM supported this work by contributing with their valuable experience in talks, software and decoder matrices. Namely Peter Plessas, IOhannes Zmölnig, Franz Zotter and Thomas Musil. Special thanks also to Martin Rumori for providing the author with the impulse responses of the Cube and Mumuth. Last but not least thanks to Fons Adriaensen for his great work on Ambisonics software.

References

Fons Adriaensen. 2005. AmbDec - Jack Ambisonics Decoder for Linux und Mac OS X.

Michael Chapman, Winfried Ritsch, Thomas Musil, IOhannes Zmölnig, Hannes Pomberger, Franz Zotter, and Alois Sontacchi. 2009. A Standard for Interchange of Ambisonic Signal Sets. *Ambisonics Symposium 2009, Graz.*

Filipe Coelho and Antonio Rodrigues. 2012. DISTRHO, Cross-Platform Audio Plugins.

Gerhard Eckel, Martin Rumori, David Pirrò, and Ramón González-Arroyo. 2012. A Framework for the Choreography of Sound. *Proceedings of the International Computermusic Conference 2012, Ljubljana.*

G. Fanelli, T. Weise, J. Gall, and L. Van Gool. 2011. Real Time Head Pose Estimation from Consumer Depth Cameras. In *33rd Annual Symposium of the German Association for Pattern Recognition (DAGM'11).*

Thomas Musil, Johannes Zmölnig, Markus Noisternig, Alois Sontacchi, and Robert Höldrich. 2003. AMBISONIC 3D-Beschallungssystem 5.Ordnung für PD. *IEM Report 15/03.*

Thomas Musil, Alois Sontacchi, Markus Noisternig, and Robert Höldrich. 2007. Binaural-Ambisonic 4.Ordnung 3D-Raumsimulationsmodell mit ortsvarianten Quellen und Hörerin bzw. Hörer für PD. *IEM Report 38/07.*

Christian Nachbar, Franz Zotter, Etienne Deleflie, and Alois Sontacchi. 2011. AMBIX - A suggested Ambisonics Format. *Ambisonics Symposium 2011, Lexington.*

Peter Plessas and IOhannes Zmölnig. 2012. The IEM Demosuite, a large-scale jukebox for the MUMUTH concert venue. *Proceedings of the Linux Audio Conference 2011.*

Martin Rumori. 2009. Girafe a versatile ambisonics and binaural system. *Ambisonics Symposium 2009, Graz.*

Alois Sontacchi. 2003. *Dreidimensionale Schallfeldreproduktion für Lautsprecher- und Kopfhöreranwendungen.* Ph.D. thesis, Graz, Austria.

Julian Storer. 2012. JUCE (Jules' Utility Class Extensions).

F. Zotter, H. Pomberger, and M. Noisternig. 2012. Energy-Preserving Ambisonic Decoding. *Acta Acustica united with Acustica*, 98(1):37–47.