AAMA - DIY AMBISONICS MICROPHONE KITS

Winfried Ritsch *

Institute for Electronic Music and Acoustics Graz
Inffeldgasse 10/3 8010 Graz,
Austria,
ritsch@iem.at

ABSTRACT

The overall target is the making of affordable DIY Ambisonics recording devices with individual flair, targeting streaming, recording, indoor and outdoors field recorders for sound art or environmental monitoring. One central part of this aim is to make a "Affordable Ambisonics Microphones Array" (AAMA). The usage of an embedded Linux computer simplified and accelerated the development of individual solutions of higher order Ambisonics microphone arrays. Hardware add-ons for these embedded Linux systems has been developed, especially for interfacing microphone arrays. Sensors enhancing the usage of these microphones, combined with corresponding software tools enables the construction of individual AAMAs.

1. INTRODUCTION

Within the last decades a lot Ambisonics audio projects as research and artwork have been carried out on our Institute and within this century the theory behind it, even might be mathematically extensive, is well documented [1] and free tools to handle Ambisonics applications are available[1]. One missing link has been for some of these projects, such as permanent sound-installations, performances and field recordings, affordable higher order microphone arrays. To fill this gap, we gave ourselves a mission statement:

> We will assemble a tool-set for making affordable Ambisonics microphone arrays as DIY projects providing hardware kits and hacks and configurable software tool-sets with example applications targeting autonomous Higher Order Ambisonics (HOA) field recorder and autonomous streaming stations.

As a vision of these portable 3D-Audio field recorder, following design targets has been drafted as following:

- microphone interfaces as multichannel ADC boards with optional phantom power
- arrangements of digital microphone arrays interfaced by I2S/TDM[2]
- microphone types: MEMS[3], electret or condenser.
- reasonable quality headphone amplifier for “field monitoring”
- calibration tools of the microphone array, initial and individual settings.

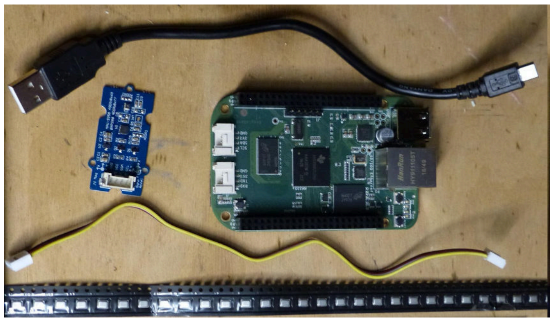

Figure 1: Kit for MEMS AAMA, 16 channel

- mountable calibrator device for on set calibration
- Ambisonics encoder build tools for arbitrary arrangement of capsules
- recording on embedded devices (SD-Card, USB-Sticks, ...)
- live-signals via low latency realtime stream, AOO[4]
- rechargeable battery with battery management system and charger.
- IMU[5]to track microphone position
- IMU to track head/headphone for monitoring.
- embedded ARM device running Debian Linux and Puredata
- ALSA driver for attached audio cards, capes and hats

As current approaches projects are the following projects:

- MEMSAAMA: TDM MEMS microphone array
- AAMA-MIC16: TDM(I2S) interface for targenting electrect microphones
- DIYASB: DIY Audio Streaming-Box based embedded linux systems[6].

2. DESIGN CONSIDERATIONS FOR DIY-MICROPHONE ARRAYS

2.1. The question of quality

Not only in the contemporary computermusic scene the idea of qualities, and especially how to distinguish certain kinds of qualities from

* This work was supported by IEM Graz and Atelier Algorythmics

[1]For example open source developments on the IEM: Plugins http://plugins.iem.at/ and http://www.matthiaskronlachner.com/?p=2015 and as Pd-Library http://git.iem.at/pd/acre-amb/

[2]I2S (Inter-IC Sound) by Philips Semiconductors. June 5, 1996., TDM Time Division Multiplex expands I2S from stereo to multichannel, eg.used by McASP Multichannel Audio Serial Port by Texas Instruments

[3]Microelectromechanical systems

[4]Audio Over Opensoundcontrol, using message based audio stream, see http://git.iem.at/cm/AOO

[5]Inertial Measurement Unit with compass and Gyrator mainly used for tracking the orientation of microphone or headphone

[6]see http://git.iem.at/cm/DIYasb

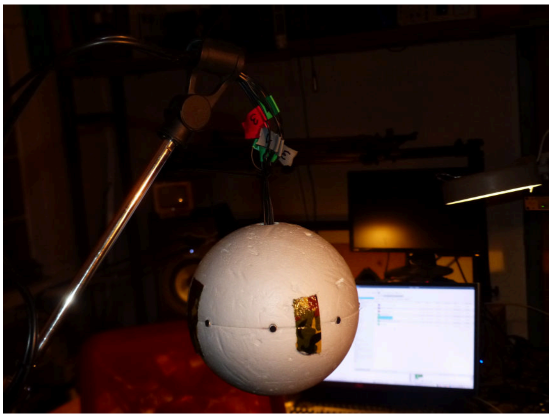

Figure 2: first test DIY 2D-Ambisonics 8 channel 4th order

one another, remains not only controversial but depends most on the usage and the environment and is not discussed here. But we assume some rough minimal criteria and numbers here, take it as examples.

For a usable recording, let us say, the noise equivalent level should be at least some dBs below the environment level. Equivalent noise is expressed as SNR to 94 db_{SPL} reference level. Within appropriate electret microphones the self noise ranges from 62 dB_{SNR} to at best 68 dB_{SNR} resulting in self-noise of 32 db_{SPL} to 26 db_{SPL}. MEMS Microphones claim to have 29 db_{SPL}, condenser microphones typical $18db_{SPL}$.[7]

Within Microphone arrays, if they are close and added, the noise to signal ratio SNR tends to increase with the amount of capsules, since the self-noise does not correlate and ambient signal does, somehow more. But the higher the order of Ambisonics, the more low frequencies are boosted and correction filters are needed for equalization which boost noise on higher frequencies[2]. We can enhance the SNR, using the higher order signals on the array only for the frequency bands needed for spatial information, which start at 200 Hz...$400Hz$ and is relevant up to 2.4 kHz[3] for horizontal spatial information and up to 8 kHz for vertical information.

Additional quality can be gained recording the frequencies bands below and above these limits with lower Ambisonics order, even 0th order, reducing noise in high frequencies, especially with an low noise omni-directional microphone added to the array, which has a much better equivalent noise level.

Recording loud environments, like in concerts, in between crowds and in cities and other loud environment, the noise level is not the problem. But doing field recording in very quiet environments, eg. in nature, MEMS and electret microphones are probably facing their limits of acceptance.

Since we anyway need to perform digital signal processing for decoding the signal, we can also enhance the qualities using equalization of the frequency response of the capsulas, meassured and then manually shaped for the needed frequency range.

2.2. Why a good headphone amplifier ?

The power for headphones coming from most (cheap) embedded devices is mostly inadequate to feed adequate field recording headphones, especially used in noisy environments. Mostly over-driven headphone amplifiers provide not only flat-sounding audio, but also the spatial information gets lost. This can easily be fixed by using a powerful headphone amplifier which provides the power levels needed to deliver enough power to headphones with low impedance ($< 50\ Ohm$) and higher peak voltages for high impedance headphones. Using a good quality headphone amplifier will not only allow you to listen at increased volume levels, but provides a much higher dynamic, full sounding audio which is essential for onsite monitoring of the recording.

2.3. A question of geometry

For different purposes and applications, different geometries of Ambisonics microphones can be chosen. The main types to differentiate are:

- open sphere with cardioid microphones
- rigid sphere with omni-directional microphones

The open construction has fewer problems with surface waves on the solid surfaces, but more with spatial shadowing.

A second decision is the form and distribution of the microphones on the microphone array. It does not need a regular layout, even too regular layouts can make more problems on spatial aliasing, but result in better distributed energy and therefore lower noise. For some situations the usage of more microphones in the "main" direction can help. Also more microphones could be arranged near the horizontal plane for better resolution in the horizontal than the vertical spatial axes or for special purposes. Different geometries can be chosen to the needs.

Figure 3: WILMA box with charger, battery, pre-amplifier ...

2.4. measurement, calculations and calibrations

The material of the mounting construction can also differ a lot and has influence on the separation of the microphone signals. For every chosen solution, there should be a procedure to measure the spatial and frequency response filters for each microphone capsule. This

[7]eg. Panasonic WM-61A 62 dB_{SNR}, PIU Audio POM-3535L-2-R 62 dB_{SNR}, MEMS Invensense ICS-52000 $68dB_{SNR}$, MEMS Infineon IM69D130 $69dB_{SNR}$, Oktava $18db_{SPL}$

and the calculated encoder matrix from the geometry and the spatial filter data can be combined to have the microphone array calibrated. All this data should be processed in the embedded device in realtime, and applied on the A-format stream to be converted in Ambisonics B-format to be recorded or streamed. All this needs at least a multi-dimensional convolution of signals, which the embedded devices are capable of.

For these purposes a special tool-set was developed within the project, once was called Ambilibrium[4] and presented in 2018. New understandings within projects led to new tools as VST-Plugins and/or Pd-Patches. They explain the procedures and provides applications, to calculate the encoder matrices and measure the microphone capsules and output the correction filter matrix for convolution in the microphones signal path.

All this is done by the "firmware" in Pd utilizing Ambisonics plugins and libraries.

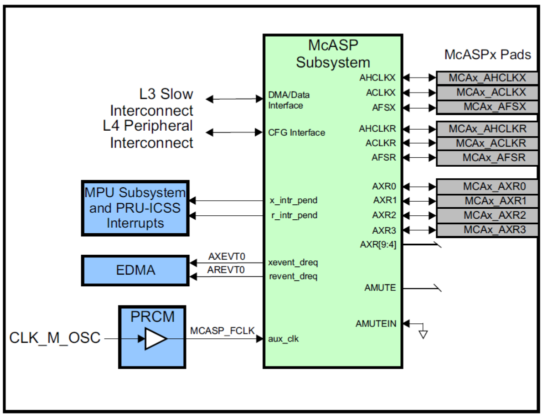

Figure 4: MacASP SOC block diagram

2.5. On set calibration

On different sites, the gain of the pre-amplifier for the microphones has to be adjusted on the sound level situation at recording environment, an on set calibration for the microphones is necessary. One solution is to mount a constant sound source near the Ambisonics microphone array in a fixed position, which can be activated for each recording, before, in pauses or afterwards the gain has been changed. Since the attenuation of the signal path from sound source to capsules has been measured before and remains somehow constant, each microphone should get the same sound level and the gain difference of the microphone can be calculated afterwards and used for the encoder. So the recording is always correctly calibrated. On post production also the microphones can be checked and possible failures identified. Also if the optional additional omni-directional microphone can be used as reference.

2.6. Microphone Stabilizer and head tracker

Doing field recording and using the AAMA as moveable microphone or placing it anywhere, sometimes the horizont is not respected exactly, even worse can move. Also moving the microphone mostly cause changes of its direction. We can rotate any Ambisonics field easily with post processing, but we need the information how. Therefore a IMU[8], with gyrator, preferably with a compass, should be mounted on the microphone and the data stream recorded or processed within the B-Format directly. Also, the position of the head for binaural monitoring can change, to get an stable monitoring sound image, we can use another on the headphone, rotating the binaural encoding with the head movements before the binaural decoder. These sensors can be attached to the embedded device to stabilize the soundfield and also record these information for post production purposes (effects).

3. CURRENT PROJECTS

Here a short list of current projects with first results, mostly under permanent development, is drafted.

[8]Inertial measurement unit, a device that measures acceleration and rotation, used to maneuver air- and spacecraft

3.1. DIYASB - DIY Ambisoncis Stream Box

A predecessor was the Wireless Large Area Microphone Array aka WILMA[5] with Beaglebone Black as an embedded Linux computer, developed at IEM, using the AM335x-core interfacing Microphones arrays via the on chip TDM (MacASP) Interface, AD-converter and calibrated pre-amplifier. It was a proof of concept and dedicated to exact measurements, but became very extensive for a simple field recorder. So for our attempt the "Beaglebone Green" has been chosen, interfacing digital microphone arrays or the more compact AAMA-MIC16 module for use of electret microphone arrays.

For Streaming Boxes located outdoor in harsh environments, the compatible industrial board with the AM3359-SOM-EVB-IND from Olimex was chosen, since it operates within industrial specification from $-40^{\circ}C$ to $+80^{\circ}C$ environmental temperature. Separate hardware was used providing (solar-)charger and battery infrastructure in the box. After using realtime enhanced Linux distributions, they seems to be more unstable and not much faster than the mainline Debian Linux kernel.

A first version with USB Audio Interfaces has been realized for Bill Fontana's project "Reenactment of Sonics Projections" for Kunsthaus Graz and run May-Oktober 2020.[9]

3.2. hamp

Previously named "ampHP" is an audiophile quality headphone amplifier driven by a on board DAC over I2S/TDM (or by an analog line signal feed), dedicated to DIY field recorders to be combined with a head-tracker. "hamp" can be powered by a single power source over 3V or 5V with DC/DC for battery powered devices and is based for now on the on the reference design using the Chip "SSM6322" from Analog Devices.

3.3. MIC8

Escher 8 channel high quality microphone pre-amplifier with 48V phantom-power, based on the "That-Microphone preamplifier" was the first developed originally for a dsPIC microcontroller, but also usable for condenser AAMA solutions. The prototype of the sym-

[9]see doku http://git.iem.at/cm/DIYasb

Figure 5: solar powered DIYasb box

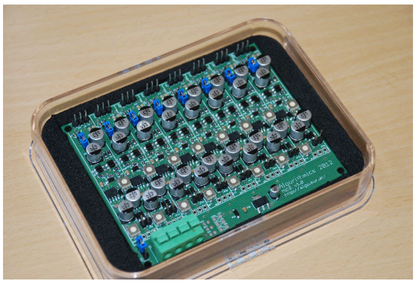

Figure 6: ESCHER-MIC8 8 channel condenser pre-amplifier

metric mic-preamp is shown in the figure 6 is dedicated, but not only for use with on board ADCs.

3.4. MEMSAAMA

MEMSAAMA is an microphone array with digital MEMS microphones like the ICS-52000 from Invensense. They can be daisy chained only using 4 signal lines and operate in a fixed mode. Also claims to be already calibrated, since they are dedicated to microphone arrays, they simplify the electronics a lot. One issue is they are tricky to solder manually and another or maybe a feature is that the opening of the mic is on the solder side, so they can be mounted on any surface and soldered as a daisy chain, directly connected to McASP Interface of the board within 30 cm wire-length. This is a very cost-less approach to an AAMA, since it needs no extra ADC board. For better handling and soldering a simple board was developed, which can be cut and arranged via ribbon cables up to 16 microphones each array. MEMS microphones like the mentioned, can also be used via a wave-channel, like suggested by Invensense, with 2 openings as differential audio, cardioid microphone and using a thin membrane over the microphone openings as hydrophones.

In the meantime MEMSAAMA was also interfaced with Espressif ESP32 devices like the ESP32-DEVKit from Olimex. Two I2S interfaces allow to interface four microphones and make a low cost first order MEMS microphone array. The interfacing via WLAN

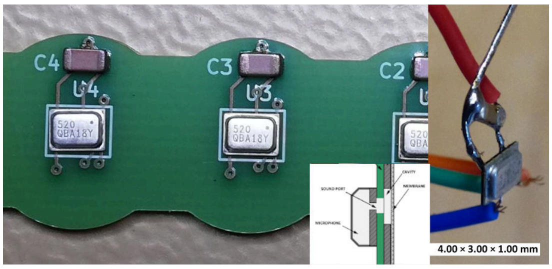

Figure 7: MEMS digital Microphone: board for daisy-chaining, sound channel, manual soldered

(or additional Ethernet with ESP32-PoE from Olimex) can be done via multichannel streams and the control via Open Sound Control (OSC). Using the Audio Library from Espressif ESP-ADF, also filters and decoder matrices can be calculated in realtime in a limited amount.

3.5. MIC16

For electret microphone arrays, a special AD-Converter board was developed, which is only 4 cm diameter and fits inside a small Ambisonics microphone array sphere, providing an multichannel TDM interfaces as well an I2C serial control bus to be connected with the embedded host using 4 dedicated 4-channels AD converter, 2 on each side.

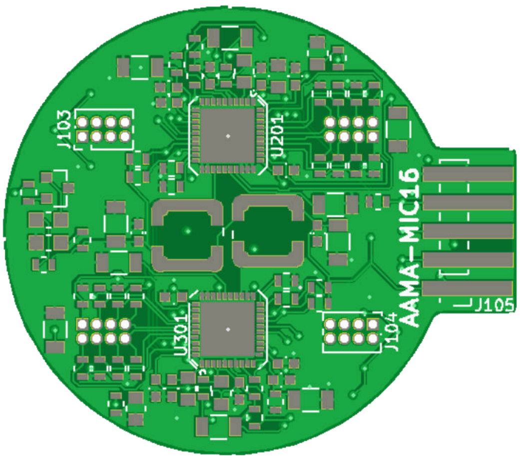

Figure 8: AAMA-MIC16 board 16 AD for electret microphones 4cm

4. LINUX SOFTWARE

The development of special Linux-drivers for the used DIY-hardware is not always needed. Since the structure of the Linux audio driver ALSA is split in ASoC layer, DAI and codec[10], mostly the configuration of "Device tree" is (theorectically) all what it needs to get audio input or output for a special implementation. This means a devicetree entry has to be written as device tree source and compiled to a binary or a proper devicetree overlay has to be provided.

But sometimes, when there is no proper codec definition available, some code for the ALSA driver has to be added. Since mostly there is no need to change many parameters on a calibrated microphone arrays, a dedicated simplified version can be implemented.

4.1. DSP processing

For the prototyping and firmware the graphical computer music language Puredata was used[6], which can be run headless on embedded devices with "-nogui" as a audio-processing application and so special variation of the kits can be managed easily. Additionally individual code can be added for art installation or other purposes. The setup and development is done via ssh over Ethernet, also using X11-forwarding on board. The control of the microphone and interfacing can be done via external devices via OSC[7] over Ethernet or an optional USB WIFI-dongle. Also buttons and LEDs are provided as user interfaces, they are free programmable for the special application. The use of Pd in Debian systems was explored and tested within the DIYasb project.

4.2. Message based audio streams

Streams could be provided using "jack" network streaming possibilities or "zita-njbrige". For a better handling and plug and play on site, the message based audio streaming AOO can be used[8]. The advantage, is there is no need to start/stop a stream on recorder or host side, it just sends the stream via Ethernet. For a field test a special art installation was realized, distributing streamboxes over the city Graz in Februar to May 2020. Therefore an advanced version of AOO has been realized. In the meantime it is also used for the Virtual Rehearsal Room (VRR)[11].

4.3. housings

Figure 9: MEMSAAMA 3D-Printed Ambisonics microphone case, first version

[10] Audio System on Chip and Digital Audio Link, mostly a I2S or TDM bus and codec device

[11] see http://vrr.iem.at/

For housings of microphone capsulas, 3D-printer cases has been used. They are nowadays easy to get but have some drawbacks, like acoustics resonances and surface waves. So nowadays softer material, like foam is used, which can be shaped very nicely manually.

5. CONCLUSIONS

Ambisonics microphone arrays as AAMA, can be realized as very cost-less devices, compared to commercial ones, excluding the working effort building these, but more important can be adapted to the needs of special usages and artworks. Workshops should enlarge the knowledge base and examples of which kind of Ambisonics microphones arrays can be made. The main purpose for a workshop is to give access to open hard- and software via workshops and presentation, this is why this information was written and hopefully a community will form enhancing those.

6. ACKNOWLEDGEMENTS

A study for the DIYASB has been done as bachelor thesis by Patrick Heidegger in autumn 2019, the streaming via AOO has been enhanced by Christoph Ressi, the MEMSAAMA board was done by Manuel Planton, the MIC16 board by Paul Tirk and Kunsthaus Graz encouraged us building Streaming Boxes for an art installation for Bill Fontana. Our thanks goes to the Institute for Electronic Music and Acoustics Graz, which let us experiment with all their tools and develop DIY kits within the education program as open hard- and software and especially Franz Zotter and Matthias Frank for their background Info on Ambisonics, always always have an open door to listen to the wishes and problems of these developments and Daniel Rudrich for his excellent plugins and tools and last but not least IOhannes Zmoelnig for all the Debian packaging and development of streaming software and of course all others who helped in any way.

7. REFERENCES

[1] F. Zotter and M. Frank, *Ambisonics - A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, Springer, Heidelberg (Deutschland), first edition edition, 05 2019.

[2] Jerome Daniel, "Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format," in *Audio Engineering Society Conference: 23rd International Conference: Signal Processing in Audio Recording and Reproduction*, May 2003.

[3] "to be referenced," .

[4] Michael Romanov, "Ambilibrium - an user-friendly ambisonics encoder/decoder-matrix designer tool," in *Audio Engineering Society Convention 144*, May 2018.

[5] M. Zaunschirm C. Schoerkhuber and J. Zmölnig, ""wilma" - a wirelesslarge-scale microphone array," in *In:Linux Audio Conference 2014*, May 2014.

[6] M. Puckette, "Pure Data," in *Proceedings, International Computer Music Conference.*, San Francisco, 1996, pp. 224–227.

[7] Matt Wright, "Open Sound Control Specifications 1.0, online," 2002.

[8] Winfried Ritsch, "towards message based audio systems," in *Linux Audio Conference 2014*, ZKM, Karlsruhe, Germany, May 2014.

[9] Winfried Ritsch, “The House of Sound - SoundLives,” URL.

[10] Winfried Ritsch, “Vision and Context of the Power Flower Bells Network,” 13.1.2012.

[11] Winfried Ritsch, “Vision and Context of the Power Flower Bells Network,” 13.1.2012.

[12] Winfried Ritsch, “Power Flower Bell Network - Idea of an electric sound installation,” internal use only, 11 2011.

[13] Josef Schauer, “Networked Power Flower Bell - Energy Harvesting System for a Cybernetic Sound Installation,” M.S. thesis, Technische Universität Graz, 2013.

[14] Josef Schauer, Winfried Ritsch, and Lothar Fickert, “Networked Power Flower Bell - Energy Harvesting System for a Cybernetic Sound Installation,” in *Proceedings, Sound and Music Computing Conference*, 2013.

[15] Stephen Wolfram, “Statistical Mechanics of Cellular Automata,” 1983.

[16] “Esp8266 module from olimex,” 2012.

[17] David Harel, Addison-Wesley, Cornwall, Great Britain, 1987.

[18] L. L. Turchet, A. McPherson, and C. Fischione, “Smart instruments: Towards an ecosystem of inter- operable devices connecting performers and audiences,” in *Proceedings of Sound and Music Computing Conference*, 2016.

[19] Justin McKeown, “Art and the internet of things: a turning point in creative education,” *The Gardian Online*, vol. may05, 2014.

[20] Patrick Beraud, “The internet of things: The future of the artist,” 2015.

[21] arduino community, “Iot manifesto,” internet.

[22] “Iot design manifesto 1.0,” internet, 2015.

[23] Eric Bonabeau, Marco Dorigo, and Guy Theraulaz, *From Natural to Artificial Swarm Intelligence*, Oxford University Press, 1999.

[24] L. Lessig, *Code and the Commons*, 1999.

[25] L. Lessig, *Code und andere Gesetze des Cyberspace*, Berlin-Verlag, 2001.

[26] T. Clausen and P. Jacquet, “RFC 3626 - Optimized Link State Routing Protocol (OLSR),” 2003.

[27] H. Lee, M. Frank, and F. Zotter, “Spatial and timbral fidelities of binaural ambisonics decoders for main microphone array recordings,” in *Proceedings of 2019 AES International Conference on Immersive and Interactive Audio (2019 AES International Conference on Immersive and Interactive Audio)*, T. Tew and J. Williams, Eds., New York (Vereinigte Staaten (USA)), 04 2019, procedure: peer-reviewed.

Figure 10: Winfried Ritsch mit DIYasb Box mit Mikro, Hotel Daniel Dach Graz