

Electric Audio Unit

VIRTUALMONIUM

The world's first acousmonium emulated in in higher-order ambisonics

Natasha Barrett 2017

User Manual

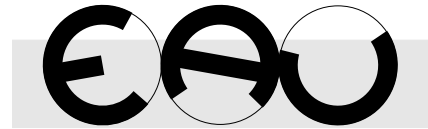


KULTURRÅDET
Arts Council
Norway

The Virtualmonium

User manual

Natasha Barrett 2017



CONTENTS

1. About the Virtualmonium	2
Features	2
2. Versions	3
CPU / OS requirements	3
Knowledge requirements	3
Availability and terms of use	4
3. Setting up your custom Virtualmonium	4
SAVING and RECALLING SETTINGS	4
AUDIO SETUP	4
VM ARRAY SETUP	5
FADER CONNECTION	6
FADERS and CHANNEL MAPPING	6
VIRTUAL SPEAKER EMULATION	7
4. Background ambisonics theory	9
Setting up the ambisonics array	9
Choosing the correct Virtualmonium version	9
5. Ongoing work	10
Bibliography	10

1. About the Virtualmonium

The Virtualmonium emulates the classical loudspeaker orchestra, or Acousmonium, in higher-order ambisonics.

The idea was formulated by Natasha Barrett in 2015 while researching in spatial audio at the University of Oslo. Further research, practical implementation and all subsequent testing was carried out by EAU with support from Arts Council Norway (2016-2017).

Features

- Graphical interface for setting up a virtual acousmonium of up to 30 virtual loudspeakers.
- Additional sends to sub-bass speakers as separate input channels or routed from the main source.
- Room model and virtual-loudspeaker rotation options.
- User selected settings for virtual speaker modelling including dynamic response and impulse responses.
- 2-D and 3-D array possibilities.
- Presets for multiple users (array layouts can be changed mid-concert).
- Flexible fader to virtual-loudspeaker mapping.
- Options for distributing multichannel sources and performing stereo diffusion in the same setup.
- Direct 'real-speaker' sends as well as 'virtual-speaker' sends.
- Off-site rehearsal possibilities over headphones (binaural).

The Virtualmonium is built in MaxMSP and uses IRCAM's higher-order ambisonics Spat MaxMSP [1] objects for spatial encoding and decoding. When using the Virtualmonium, sound diffusion is performed in the traditional way, but instead of mixing desk faders, MIDI faders are used.

2. Versions

There are currently four versions of the Virtualmonium:

- 24-channel horizontal ambisonics array encoded in 11th order 2-D HOA.
- 16-channel horizontal ambisonics array encoded in 7th order 2-D HOA.
- 24(25)-channel 3-D array encoded in 4th order 3-D HOA.
- Binaural version in 3-D for headphones and home listening / private rehearsal. (See section 3 concerning limitations).

For all versions, the HOA decoder applies the dual-band energy preserving method at a cross-over frequency of 400 Hz [2]. By default, NFC-HOA is not implemented (to preserve CPU), but can be added on request.

All versions are set to a sampling rate of 44.1Khz. Changing this value will result in incorrect impulse response speaker emulation.

For future versions and ongoing work, please refer to section 5.

CPU / OS requirements

The Virtualmonium runs on MacOSX. It has been tested on the following systems:

- MacOS Sierra
- MacOS El Capitan.
- CPU - 3.1 GHz Intel Core i7
- Memory - 16 GB 1867 Mhz DDR3

Knowledge requirements

A basic knowledge of MaxMSP is required for connecting your own MIDI faders.

It is not necessary for knowledge of ambisonics, but the user is advised to consider section 5 for selecting the correct Virtualmonium version and for setting up the real loudspeakers of the ambisonics array.

Availability and terms of use

Please contact nlb@natashabarrett.org or admin@electricaudiounit.no

3. Setting up your custom Virtualmonium

Open one of the Virtualmonium versions and you will see a simple front-end consisting of five coloured panels. We'll look at each of these panels separately.

1. SAVING and RECALLING SETTINGS

User settings are saved individually in each sub window. For an example about working with presets, open 'Fader to virtual speaker mapping' in the section 'Faders and Channel Mapping'. On the left side, you will see the 'presets' panel.

On start-up, a default set of presets are loaded. To make your own presets, click on the button 'write'. You will be prompted to name your preset file, and this will then contain the eight available presets.

Important: ensure that you save your own preset file in the same folder as the VM program.

Follow the onscreen instructions and save your setting in one of the 8 possible slots. Likewise, follow on screen instructions for recalling settings. *TIP:* use the same slot number for all sections (sub windows) throughout your complete setup. All sections provide preset saving options unless otherwise stated.

Now close the sub window currently open and we can look at the right side of the main front interface called 'GLOBAL RECALL'. Here you will see a similar panel and 8 possible slots in which to save a global recall. The slot numbers in this section correlate with the slot numbers in other areas of the program.

GLOBAL RECALL will function correctly if settings are consistently stored in the correct pre-set slots.

- ➔ Activate global recall.
- ➔ Follow instructions on screen.

2. AUDIO SETUP

Open DAC

Here you select your audio interface and other parameters concerning audio processing in MaxMSP. For CPU optimisation, it is recommended to use an i/o and signal vector size of 2048. 'Scheduler in Overdrive' should be turned on, while 'Audio in Interrupt' turned off. To turn on audio processing, press the button in the top left corner.

Stereo player

Open the *Stereo player* to load a stereo source.

Multichannel player

Open the *Multichannel player* to play a multichannel source. Multichannel sources can contain from 4 to 26 channels and should be formatted as one multichannel sound-file. The channels in the source can be routed to virtual loudspeakers in the Virtualmonium as well as directly to real loudspeakers in the ambisonics array.

Master gain

Controls the final output volume of the ambisonics decoding (this value can be adjusted by hand, or be mapped to a fader).

Output meters

The output meters show the output of the ambisonics decoding – or in other words the audio levels sent to the real loudspeakers.

3. VM ARRAY SETUP

Open VM-layout

Open the layout and a new window will appear. This window shows the ambisonics loudspeaker array (the black squares with numbers) and the virtual loudspeakers (in green), using IRCAM's `spat.viewer`. Click and drag the green virtual loudspeakers to place them in virtual space. If using the 3-D array version of the Virtualmonium, 'v' on the keyboard will toggle between the view from above and from the side.

- With *Speaker mirroring* turned on, moving odd-number speakers will move mirror-image even-number speakers.
- The closest proximity at which we can clearly hear the distance of a virtual-loudspeaker will be at the perimeter of the ambisonics array. You may wish to test virtual loudspeakers positioned inside the ambisonics array, but listen carefully as their proximity will appear vague. For all distances, angular location will be precise with respect to the chosen version of the Virtualmonium (see information provided in section 5).
- In general, we suggest locating virtual loudspeakers outside of the ambisonics array on the visualisation (zoom out of the visualisation using the mouse wheel or track-pad to see a larger area). More interesting results will be achieved by increasing the distance range between near and far, and increasing this to an extent greater than that of a real acousmonium.¹

Load an array

On opening the VM for the first time, there are three default layouts which are overwritten when you save your own array to one of these three slots. Click on the pop-up to select a layout.

Speaker mirroring

With *Speaker mirroring* turned on, moving odd-number speakers will move mirror-image even-number speakers.

VM-Yaw settings

Yaw settings control the direction of face for each virtual loudspeaker.

- Open the yaw settings and a new window will appear.
- The boxes along the top show yaw settings for each virtual speaker. Drag or type numbers into the boxes and see how the virtual-loudspeaker rotation changes in the *VM-layout*.
- If *Speaker mirroring* is turned on, the rotations will be mirrored in pairs.

Save array

Save your new array in one of three slots. *Sure?* then press the button, which will now show red to indicate that the array has been saved. This array has now overwritten the default layout.

User defined ambisonics speaker locations

Only change these settings if you are familiar with higher-order ambisonics. Follow on-screen instructions.

¹ See Barrett 2016 [7] for an explanation.

Important note for the binaural version

In binaural, direct speaker sends outside of the ambisonics domain make no sense. In the binaural version, the graphics are still present, but only the ambisonics domain processing is enabled. Users should also be aware of the spatial-perceptual limitations of binaural listening.

4. FADER CONNECTION

Users are required to patch in their own MIDI or OSC hardware faders. Long-throw faders are recommended. It is assumed that the user has a sufficient knowledge of MaxMSP to connect their own hardware.

- Open the subpatch 'Faders_sends.maxpat'.
- Set up an appropriate MIDI-IN in MaxMSP. Most fader-hardware will send control messages on different channels.
- Follow the onscreen instructions. Scale the fader range to output between 0. to 1. as floats.
- You may wish to scale the fader range with an exponent so as to emulate the character of your favourite mixing desk.
- Connect the individual data streams from your MIDI faders to each of the boxes (f1 = fader 1, f2 = fader 2... etc)

5. FADERS and CHANNEL MAPPING

In this section, you decide the mapping of midi faders to virtual loudspeakers, decide if direct speaker sends are also to be used, and decide on multi-channel source routing.

Fader to virtual speaker mapping (stereo sources)

Open the fader mapping window and you'll see a matrix on the right side.

- The top line shows the fader numbers as routed in the 'Faders_sends.maxpat'.
- The side shows the numbers of the 30 virtual loudspeakers.
- Click on a point on the matrix to map a fader to one or more virtual speakers. *TIP:* think about an ergonomic setup, where fader 1 may not be best mapped to speaker 1. For further reading, see [3]
- Scroll down and you'll see a section where faders can be mapped to up to eight direct speaker-sends (i.e. not in the ambisonics domain). *TIP:* direct sends will result in the real speaker being heard, and contrast to the virtual speakers. This can be effective, but should be used with care. Note that so far, we have not specified which of the real-speakers these direct sends will be routed to (see below).
- At the very bottom is a choice to map faders to the sub-bass and to the master volume.

Location of direct sends from stereo midi fader mapping

In the 'Fader to virtual speaker mapping' above, we decided if any of our faders controlled sound routed directly to the loudspeakers of the ambisonics array (i.e. not to any of the virtual-speakers of the Virtualmonium). In the current section, we decide which speakers these are. At the top of the right side we see the eight possible feeds which are controlled by the faders already selected. On the left of the right side you can decide which of the 24 real loudspeakers receive any of these eight direct-speaker sends.

Multichannel sources routing

This window is used to map channels from a multichannel source to virtual loudspeakers in the VM-array.

- The top line shows the channels of the multichannel source.
- The side shows a choice of 24 virtual speakers and two sub-bass speakers (the two sub-bass are direct sends and not in the ambisonics domain).

Visual fader levels

The *Visual fader levels* pop-up shows the value of virtual speaker gains from 0-1. These values are for visual reference to show the state of the MIDI faders.

- ➔ Move a fader and see that the correct virtual speaker gain also varies.
- ➔ The values in the *Visual fader levels* can also be entered by hand, but be aware that these values do not feedback to control motorised MIDI faders.²

6. VIRTUAL SPEAKER EMULATION

Although the loudspeakers of the Ambisonics array will colour the sound, we assume that they are neutral. (The user may decide to test their own loudspeaker compensation, but in our own tests using a Genelec system we found that a deconvolution method was of no advantage).

The virtual loudspeakers emulate real loudspeakers using three methods:

- Directional impulse response convolution.
- Single impulse response convolution.
- Room models.
- Dynamic response modelling.

Impulse responses convolution

Virtual speakers 1-8 are modelled using directional IRs. Direction IRs from the online database at the 3-D Audio and Applied Acoustics Lab at Princeton are used for convolution [4]. In the database, measurements were made for every 5-degree speaker rotation. To reduce computation demands, the Virtualmonium combines IRs from each 60-degree rotation. Therefore, each virtual speaker is represented by six directional virtual sources³, where each direction is convolved with the correlated IR. In the Virtualmonium these speakers are called ‘complex’ virtual speakers.

Virtual Speakers 9-20 are modelled with single impulse responses. The convolution for these speakers can also be bypassed via the interface.

Virtual Speakers 21-30 are not convolved with IR.

TIP: It is useful to remember which virtual speakers are modelled by different methods as they will stimulate the room model differently (the eight complex speakers being the most interesting).

Selecting impulse responses

A default set of IRs are loaded at start up. To select your own IRs, first open the *Impulse responses* window. On the right-hand side is a list of 10 loudspeaker names, and three possible user-created names. The list of 10 show the names of real loudspeakers from which IRs have been obtained and their corresponding ID. We will return to the three user-created names shortly.

- ➔ The top section of the pink window frame allows new IRs to be selected.
- ➔ Virtual speakers 1-4 share the same IRs, as do virtual speakers 5-8, 9-12, 13-16, 17-18, 19-20.
- ➔ To change a set of IRs, change the ID for the speaker group, then press the button to activate the new IR buffers.
- ➔ The bottom section of the pink window frame allows IRs for certain loudspeakers to be bypassed.

² Feedback to motorised fader control could be implemented in a future version with access to suitable hardware.

³ During the development phase, filter coefficients were tested as an alternative to convolution, where coefficients were extracted from impulse response data. We however found no CPU or technical advantage to this method, where furthermore the filter method lost any directionally derived phase information captured in the original IRs. Technical explanation can be found in [4].

Adding custom IRs

The users own selection of IRs can also be used. IRs must be at a sampling rate of 44.1 KHz, labelled in the following way, and placed in the same location as the Virtualmonium's own IRs:

Centre (and omni for speakers modelled with single impulse responses):

```
user1_44100_1.wav  
user2_44100_1.wav  
user3_44100_1.wav
```

Side 1a-1b:

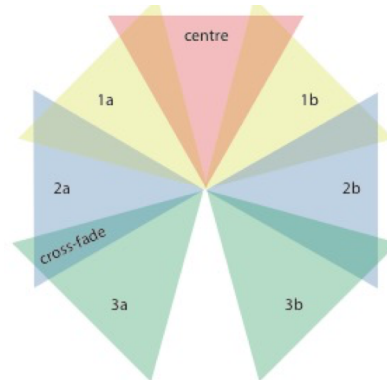
```
user1_44100_7.wav;  
user2_44100_7.wav;  
user3_44100_7.wav;
```

side 2a-2b:

```
user1_44100_13.wav;  
user2_44100_13.wav;  
user3_44100_13.wav;
```

side 3a-3b:

```
user1_44100_19.wav;  
user2_44100_19.wav;  
user3_44100_19.wav;
```



Room models

IRCAMs room model consisting of early and late reflections is used [5]. The room model results in acoustically correct distance related gain differences, low pass filtering and reverberation. Most interestingly, it allows us to rotate speakers (in the yaw settings) and stimulate wall reflections in a similar way as one would do so in a real acousmonium. As a default, a concert hall sized model is used. The actual significance of the room and the reverberation, or the extent to which the room is stimulated and heard, is determined by the amount of each virtual-speaker signal that is fed into the model.

Room model early reflections levels (complex speakers 1-8)

Speakers 1-8 are only sent to the early reflections room model. This window controls the amount of signal that is sent.

Room model reverb levels simple speakers 9-30

This window controls the amount of signal from speakers 9-30 that is sent to the full room model.

Manually on/off distance-delay

The Spat room model includes distance derived time delays. The delay time is calculated based on the speed of sound and the distance of the virtual speaker from the centre of the listening space. These delays are therefore correct with respect to distance. However, as we are modelling virtual speakers (and virtual rooms), our emulation is not subject to the full acoustic and spatial complexity of real rooms and real speakers. The result is, in some situations, that distance time-delay can result in audible comb-filtering and other modulation artefacts. The section *Manually on/off distance-delay* allows user control over which virtual speakers are processed with a delay.

Dynamic response processing, speakers 9-30

Dynamic response modelling allows the user to specify two types of compression and excitation, one solo compression, and a high-pass filter tweeters emulation. Use the popups at the top of the window to select dynamic processing. *TIP*: tweeter emulation makes more sense in the 3-D version.

Dynamic response controls

The *Dynamic response controls* allow compression, excitation and high-pass filtering to be tuned as desired.

Speaker gain balance (stereo diffusion)

This window allows the input gain for each virtual speaker to be changed, in the same way as you would increase the gain of speaker amplifiers in an acousmonium. To avoid user error, this section has only been implemented for stereo, and not for multichannel sources. *TIP:* Increasing the gain for distant loudspeakers in stereo diffusion can be useful for creating ‘big’ sounds far away, that stimulate the room model.

4. Background ambisonics theory, setting up the ambisonics array, and choosing the correct Virtualmonium version

The Virtualmonium is built in higher-order ambisonics. Ambisonics is a method for creating 3-D sound-fields. For detailed information on ambisonics, many resources are available (such as [6] [7]).

For HOA, increasing numbers of directional components in the spatial encoding facilitates an increasingly accurate representation of the spatial scene, as well as an increasing size of area within which the audience will hear the intended result. To decode a certain number of directional components in the spatial encoding requires a certain number of loudspeakers. The greater the number of directional components, the more loudspeakers are required. Although numerical methods may predict source localization for different orders of ambisonics, few studies have tested subjective evaluations [8] [9]. We do however experience that the angular resolution increases with the order of ambisonics. For the Virtualmonium this will mean that the higher the order of ambisonics, the more virtual loudspeakers will be audibly distinguishable. The more speakers there are in the ambisonics array, the higher the order of ambisonics that can be decoded. Therefore, the user should choose the best version of the Virtualmonium that the available ambisonics array can support.

Setting up the ambisonics array

- All loudspeakers should be of the same type and set to the same gain.
- Without working knowledge of ambisonics, the speakers should be arranged in a circle and equally spaced.
- With working knowledge of ambisonics, the speakers can be arranged in other geometries such as rectangles and ovals, the speaker locations measured, and the coordinates entered into VM ARRAY SETUP -> *User defined ambisonics speaker locations*. The decoder will automatically calculate gain and time alinement.
- As with all ambisonics arrays, the more accurate the speaker placement, the more accurate the result.
- Speakers should be numbered as shown in ‘VM ARRAY SETUP -> VM-layout’.
- A test signal should be used to ensure that all speakers are of the same perceptual gain when standing in the centre of the space.

24 channels ambisonics array in 2-D or 3-D?

The 24-channel 2-D ambisonics array supports 11th order ambisonics, while the 24-channel 3-D array supports only 4th order ambisonics. However, if it is possible to set up an accurate 3-D array, then this option is interesting even though of lower resolution. The 3-D array will allow virtual speaker elevation as well as tweeter-grid emulation. Nevertheless, in EAU’s own use, audiences have expressed having experienced a sense of source elevation even when only using the 2-D array. Setting up a 3-D array is trickier than a 2-D array, and only recommended with working knowledge ambisonics and the decoder tolerances of speaker distribution. The 3-D array’s loudspeaker locations will need measuring and these coordinates loaded into the Virtualmonium in VM ARRAY SETUP -> *User defined ambisonics speaker locations*.

How many virtual loudspeakers will the system support?

Determining the number of virtual loudspeakers that the Virtualmonium can support is not straightforward. If all virtual loudspeakers were the same colour and proximity, we would resolve one result; if virtual loudspeakers are a mixture of different colouration and proximity, the number is greater.

Over the 14-channel ambisonics array encoded in 11th order 2-D HOA, if specifying different virtual-speaker distances and speaker responses, we find that the system supports approximately 30 virtual loudspeakers. This is the number of virtual loudspeakers that we could differentiate between *perceptually*. For the 16-channel ambisonics array encoded in 7th order 2-D HOA, the number of perceptually relevant virtual loudspeakers will be less. However, both ambisonics encodings allow the user to feed 30 virtual speakers, and the best way to decide how many to actually use is via critical listening. It is not necessary to use all virtual speakers even though they may be visible on the interface.

5. Ongoing work

The binaural version

The binaural version currently uses the free Hoalibrary binaural decoder (<http://www.mshparisnord.fr/hoalibrary/en/>). Alternative decoders that allow the user to enter their own HRTF set are being investigated.

Decoders and room models

Ambisonics decoders and room models are in a state of development. The methods implemented are the best options currently available for real-time use. In our own ongoing work, we keep abreast of new developments, which will reflect in future versions of the Virtualmonium.

Ambisonic arrays

Alternative ambisonic arrays can be made on request.

Further reading:

- Barrett, N., Jensenius, A R. 2016. The 'Virtualmonium': an instrument for classical sound diffusion over a virtual-loudspeaker orchestra. Proceedings of the International Conference on New Interfaces for Musical Expression NIME'16. Brisbane, Australia.
- Barrett, N. 2016. A 3-D future loudspeaker orchestras emulated in higher-order Ambisonics. Proceeding of the International Computer Music Conference 2016. Utrecht, Netherlands.

Bibliography

- [1] IRCAM. [Online]. Available: <http://forumnet.ircam.fr/product/spat/> Last accessed October 2017..
- [2] F. Zotter, H. Pomberger and M. Noisternig, "Energy-Preserving Ambisonic Decoding," *Acta Acustica United with Acustica*, vol. 98, p. 37–47, 2012.
- [3] J. Harrison, "Diffusion: theories and practices, with particular reference to the BEAST system.," 1999. [Online]. Available: http://econtact.ca/2_4/Beast.htm. [Accessed 2017].
- [4] J. Tylka, R. Sridhar and E. Choueiri, "A Database of Loudspeaker Polar Radiation Measurements," *Audio Engineering Society Convention 139*, October 2015.
- [5] T. Carpentier, M. Noisternig and O. Warusfel, "Hybrid reverberation processor with perceptual control," in *Proc. of the 17th Int. Conference on Digital Audio Effects (DAFx-14)*, Erlangen, Germany., 2014.
- [6] F. Hollerweger, "<http://flo.mur.at/writings/HOA-intro.pdf>," 2008. [Online].
- [7] J. Daniel and S. Moreau, "Further Study of Sound Field Coding with Higher-Order Ambisonics," *116th Convention of the Audio Engineering Society*, 8-11 May 2004.
- [8] M. Frank, F. Zotter and A. Sontacchi, "Localization Experiments Using Different 2-D Ambisonics Decoders," *25th Tonmeistertagung - VDT International Convention*, 2008.
- [9] H. Wierstorf, A. Raake and S. Spors, "Localization in Wave Field Synthesis and higher order Ambisonics at different positions within the listening area," *Proc. DAGA.*, 2013.

- [10] N. J. A. J. Barrett, "The 'Virtualmonium': an instrument for classical sound diffusion over a virtual-loudspeaker orchestra," in *Proceedings of the International Conference on New Interfaces for Musical Expression NIME*, '16. Brisbane, Australia., 2016.