IRCAM HEar

FLUX:: Immersive

2/6/23

Table of contents

1 Introduction

Product Page | Shop Page

ircam Tools by FLUX::

HEar

Binaural Encoding Tool

HEar allows faithful reproduction of a stereo or surround mix with a pair of conventional stereo headphones. It relies on proven technology to model the various phenomena that occur when playing back audio material through a loudspeaker system.

This allows monitoring a full surround mix in situations when a surround-capable environment is not available or practical. Another typical use of HEar is doing precise checking of a mix, which is convenient with headphones as these provide a ‘surgical’ and very detailed, microscope-like rendering of the audio.

It can also prove very useful in a project studio context, and whenever noise isolation is a concern, as it helps achieving a more realistic sound environment.

2 Concepts

2.1 Auditory scene perception

Our perception of sound mainly relies on our ability to identify and characterize a number of sound sources, depending on the spatial parameters of these sources, such as apparent position and size.

The duplex localization theory developed by Lord Rayleigh in 1907, claims two factors are predominant to characterize perception, namely the differences in arrival time (ITD) and intensity (ILD) between sounds reaching our ears. Perceived sound variations between ears are mainly attributed to the head obstructing sound waves and therefore forcing them to travel around the head in order to reach the opposite-facing ear.

Subsequent studies have confirmed and refined this theory which has prevailed ever since its introduction more than a century ago.

2.2 Localisation

The ITD and ILD localization indexes are derived from measurements of the transfer function between the sound source's origin, taken at a certain incidence, and the listener's eardrums. The transfer function summarize the transformations the sound goes through before reaching the listener, including diffraction, diffusion and reflection on the listener's body and head.
These measurements are commonly referred to as HRTF (Head-Related Transfer Function).

2.3 Binaural technology

Binaural technology encompasses methods for recording, processing, synthesizing and reproducing sound that are specifically designed to preserve tridimensional localisation properties.
In order to mimic the impression of a sound originating from a given incidence, it is sufficient to filter a mono signal, which on its own lacks any kind of spatial information, with both left and right HRTF filters. This constitutes the foundation of binaural synthesis.
> Please note that the resulting signal is only meant to be listened to with headphones, and isn't designed for a conventional stereo loudspeaker setup.

2.4 Virtual head

This plugin relies on HRTF filter measurements made using a KEMAR (Knowles Electronics Manikin For Acoustic Research) dummy head and torso simulation. This type of manikin was conceived during the 1970's for conducting acoustics experiments using a model with anthropometric dimensions equivalent to that of an average human listener.

Courtesy of G.R.A.S. Sound & Vibration
http://www.gras.dk/

2.5 Virtual speakers

The audio input is routed internally to virtual speakers, through a routing matrix. These represent the emulated loudpseaker setup configuration.

3 Controls

3.1 (1) Routing Matrix

The routing matrix gives an overview of the mapping between the plugin's inputs from the DAW track to the virtual speaker internal outputs. The virtual speaker outputs are down-mixed to stereo using a virtual speaker processing algorithm.

Please take note that the plugin's output to the DAW track itself is always stereo as the binaural processing is intended exclusively for use with headphones.

The meters above the fi rst row indicate the source levels of individual input channels.

User controls

3.2 (2) Speaker Mode

Specifies which virtual speaker configuration should be emulated. Available modes depend on the configuration of the track the plugin is inserted into, and comprise of one or more of the following:

- 5.0
- 5.1
- 7.1
- 8.0

3.3 (3) Space Preset

Selects between different spaces with subtly different colorations (Preset 1..3) or completely neutral (No Effect)

3.4 (4) Speaker Width

Controls the width between virtual speakers, expressed in degrees. The default is 60°, which corresponds to the recommended setting. This allows to narrow or broaden the stereo image.

3.5 (5) Angle Shift

Controls the angle between the listener and the centre of the virtual speakers. The default is 0°, which corresponds to the ideal listener position, giving a balanced image between channels.

3.6 (6) Setup Menu

Advanced settings to override default behavior, typically when using hosts that do not conform to the standards.

3.6.1 I/O

Override automatic track I/O specifications. HEar automatically adjusts its I/O configuration based on what the hosts reports to the plugin. Some hosts such as Logic do not report this correctly or do not support asymmetric I/O configurations In this case you have to do this manually and select amongst a number of choices of symmetric (N-to-N) and asymmetric I/O (N to stereo).

3.6.2 Options

These are best left at their default in most cases, but can be changed if required:

- Disable processing during bypass: stops processing completely during bypass. Allows to conserve CPU when using many instances and a lot of bypass on/off automation, such as film or sound effects mixing. Default is off (enabled).
- Use Multi-Thread Automation: dedicate a separate thread for automation. Useful when heavy automation is present in the project to get rid of possible audio dropouts. Default is off (processing and automation share the same thread).
- Try to avoid latency as possible: minimize latency by employing minimal buffering, possibly at the expense of a little CPU overhead. Default is on.
- Report latency: report plugin latency, if any, to the host. Some hosts have difficulty coping with large latency values, in this case you can force the plugin to report zero, but you'll have to manually compensate for this for tracks to remain synced. Default is on (report true latency).

4 Credits

Design of digital signal processing algorithms and implementation in Max: Jean-Marc Jot (Espaces Nouveaux / Ircam).
Objective and perceptual characterization of room acoustical quality: Jean-Pascal Jullien, Olivier Warusfel, Eckhard Kahle.

Additional contributions by Gerard Assayag, Georges Bloch, Martine Marin, Véronique Larcher, Guillaume Vandernoot and Khoa-Van Nguyen.

C++ development code : Thibaut Carpentier , Remy Muller; additional contributions: Gael Martinet.

Thanks to Xavier Chabot, Eric Daubresse, Gerhard Eckel, Serge Lemouton, Gilbert Nouno, Laurent Pottier, Manuel Poletti, Leslie Stuck, and Zack Settel for instructive discussions and advice.

Ircam R&D Director : Hugues Vinet

Collection Manager: Frederick Rousseau

IRCAMTOOLS
Collection Manager for IRCAM: Frederick Rousseau
Collection Manager for Flux:: Gael Martinet

IRCAMTOOLS SPAT, VERB

FLUX::
Head software engineering: Gael Martinet
Developers: Gael Martinet, Samuel Tracol, Siegfried Hand
Designer: Nicolas Philippot

Manual: Lorcan Mc Donagh/Thibaut Carpentier

5 References

L. Rayleigh. On our perception of sound direction. Philosophical magazine, XIII :214–232, 1907.

R. S. Woodworth, (1938) Experimental Psychology. New York: Holt, Rinehart, Winston.

J. Middlebrooks and D. Green. Sound localization by human listeners. Annu Rev Psychol, 42 :135–159, 1991.

E.M. Wenzel, M. Arruda, D.J. Kistler, and F.L. Wightman. Localization using non-individualized head-related transfer functions. J. Acoust. Soc. Am,. 94, pages 111–123, 1993.

J. Blauert. Spatial Hearing : The Psychophysics of Human Sound Localization. MIT Press, Cambridge, MA, 1983.

D. R. Begault, 3-D Sound for Virtual Reality and Multimedia, Cambridge, MA: Academic Press Professional, 1994.

V. Larcher. Techniques de spatialisation des sons pour la réalité virtuelle. PhD thesis, Université Paris 6, 2001.

W.G. Gardner and K. Martin. Hrtf measurements on a kemar dummy-head microphone. Technical Report 280, MIT Media Lab Perceptual Computing, 1994.

M.D. Burkhard, R.M. Sachs, “KEMAR the Knowles Electronics Manikin for Acoustic Research”. Report No. 20032-1. Industrial Research Products, Inc., Elk Village, Illinois, (November 1972).

F. Rumsey. Spatial Audio. Focal Press, 2001.

6 Specifications

6.1 Availability

Ircam HEar is available in:

AU / VST / VST3 / AAX Native/ *AAX AudioSuite*

* *AAX Native & AAX AudioSuite in Pro Tools 11 and later*

6.2 Processing

Ircam HEar provides :

- Up to 16 channels Input/Output in VST/VST3/AU/AAX.
- 64-bits internal floating point processing.
- Sampling rate up to 384 kHz.

6.3 Hardware Requirements

A graphic card fully supporting OpenGL 2.0 is required.

- macOS : OpenGL 2.0 required – Mac Pro 1.1 & Mac Pro 2.1 are not supported.
- Windows : If your computer has an ATi or NVidia graphics card, please assure the latest graphic drivers from the ATi or NVidia website are installed.

6.4 Software License Requirements

In order to use the software an iLok.com user account is required (the iLok USB Smart Key is not required).

6.5 Compatibility

All major native formats are supported

6.5.1 Windows – 10, in 64 bits only.

- VST (2.4)
- VST3 (3.1)
- AAX Native*
- AAX AudioSuite*

6.5.2 macOS (Intel and ARM)

All versions from Sierra (10.12) to latest. (Compatible with previous versions but not supported)

- VST (2.4)
- VST3 (3.1)
- AU
- AAX Native*
- AAX AudioSuite*

* *AAX Native & AAX AudioSuite in Pro Tools 11 and later*