LESON

INSIDE THE SOUND

DeepStereo Monitor

Real-time binaural monitoring for immersive audio production

VST3 - AU - AAX Plugin

macOS - Windows

music UNIT

Concept

DeepStereo Monitor is a binaural monitoring solution specifically designed for professional sound engineers, mix engineers, and immersive audio content creators who require accurate stereophonic monitoring of their multichannel mixes.

Immersive production formats such as Dolby Atmos or MPEG-H require precise and reliable monitoring tools. Standard binaural rendering techniques often introduce coloration or simulated room effects that may bias mixing decisions.

DeepStereo Monitor eliminates this ambiguity by providing a reference binaural representation that reflects the raw multichannel mix content rather than a simulated listening environment. This approach enables reliable mixing decisions both on loudspeakers and through current immersive playback systems.

DeepStereo Monitor establishes a new reference for immersive production. It is designed to meet the needs of audio professionals who require:

- Accurate representation of multichannel mixes
- A monitoring tool prioritizing sonic realism
- Predictable translation to loudspeaker systems

Key features

Universal HRTFs for all users

DeepStereo Monitor is based on a patented "universal HRTF" technology delivering consistent binaural rendering to all listeners without requiring individualized calibration. This approach ensures reliable spatial perception and mix translation, which are essential in immersive workflows.

Native binaural monitoring without coloration

Rather than emulating room acoustics, DeepStereo Monitor prioritizes neutrality and fidelity. The mix is perceived exactly as it sounds, without spectral or spatial coloration altering the signal.

Compatible with all headphones

Platform-independent by design, DeepStereo Monitor operates with any standard stereo headphones, with or without head tracking, offering flexibility for studio, mobile studio, or remote workflows.

Compatibility with all immersive audio formats

DeepStereo Monitor supports major multichannel formats up to 22.2 and provides customizable configurations up to 32 channels. It is fully compatible with immersive export formats such as Dolby Atmos and MPEG-H. Users can reliably monitor complex spatial content from diverse sources without leaving the headphone environment.

Seamless DAW integration (Mac/PC)

DeepStereo Monitor is available as a VST3, AAX, and AU plugin compatible with all major digital audio workstations running on macOS and Windows. Simply insert the plugin on a monitoring bus or any routing chain to instantly access immersive monitoring capabilities without additional hardware.

Advanced monitoring tools

DeepStereo Monitor provides comprehensive monitoring controls:

- Speaker groups (solo & mute): isolate specific channels for targeted mixing decisions
- HRTF profiles: optimized selections for music or cinematic workflows
- Contour modes: three listening comfort options for extended externalization while preserving neutrality
- Sound scene orientation: navigate the spatial field with listening direction control
- Dynamic head tracking support: maintains three-dimensional scene stability according to listener head orientation when using a head tracker

General architecture

DeepStereo Monitor consists of three distinct and complementary components:

- A binaural synthesis engine based on universal transfer functions HRTF profiles: optimized selections for music or cinematic workflows
- A rendering engine incorporating specific processing to manage complex interactions between sources
- A “contour” system designed to enhance externalization and spatial perception without resorting to room modeling

About sound

DeepStereo technology allows for faithful reproduction of the original source prior to binauralization, particularly at the front reference angle, where most sound work is typically performed.

During the mixing and/or mastering phases in a professional setting, high frequencies are particularly necessary for controlling many aspects, both technical and artistic: transients, harmonic distortions, crossfades, sibilance, breath sounds, and saturation.

The presence of fully reproduced high frequencies, unattenuated and unsoftened, is essential in a professional monitoring tool, on every element of the audio chain: preamp, monitoring controller, converter, headphones, speakers.

The presence, clarity, and definition of the high frequencies that can be heard when listening to the DeepStereo Monitor are therefore neither flattering, nor artificial.

Perceptive externalization

Deep Stereo Monitor incorporates a “contour” system designed to enhance spatialization and perception of space. Unlike traditional approaches based on simulating a room or reverberation field, this system does not rely on a physical model of a room.

Although the perception of distance and externalization is physically linked to the presence of a diffuse field, the introduction of a room effect causes signal coloration, which is incompatible with the requirements of fidelity and neutrality specific to monitoring.

The contour system therefore adopts a non-physical, perceptual approach, aiming to reinforce certain cues of externalization without introducing any noticeable acoustic signature. The goal is to maintain the closest possible proximity to the signal being listened to, while improving the stability and clarity of the spatial image in the headphones.

Download, installation & activation

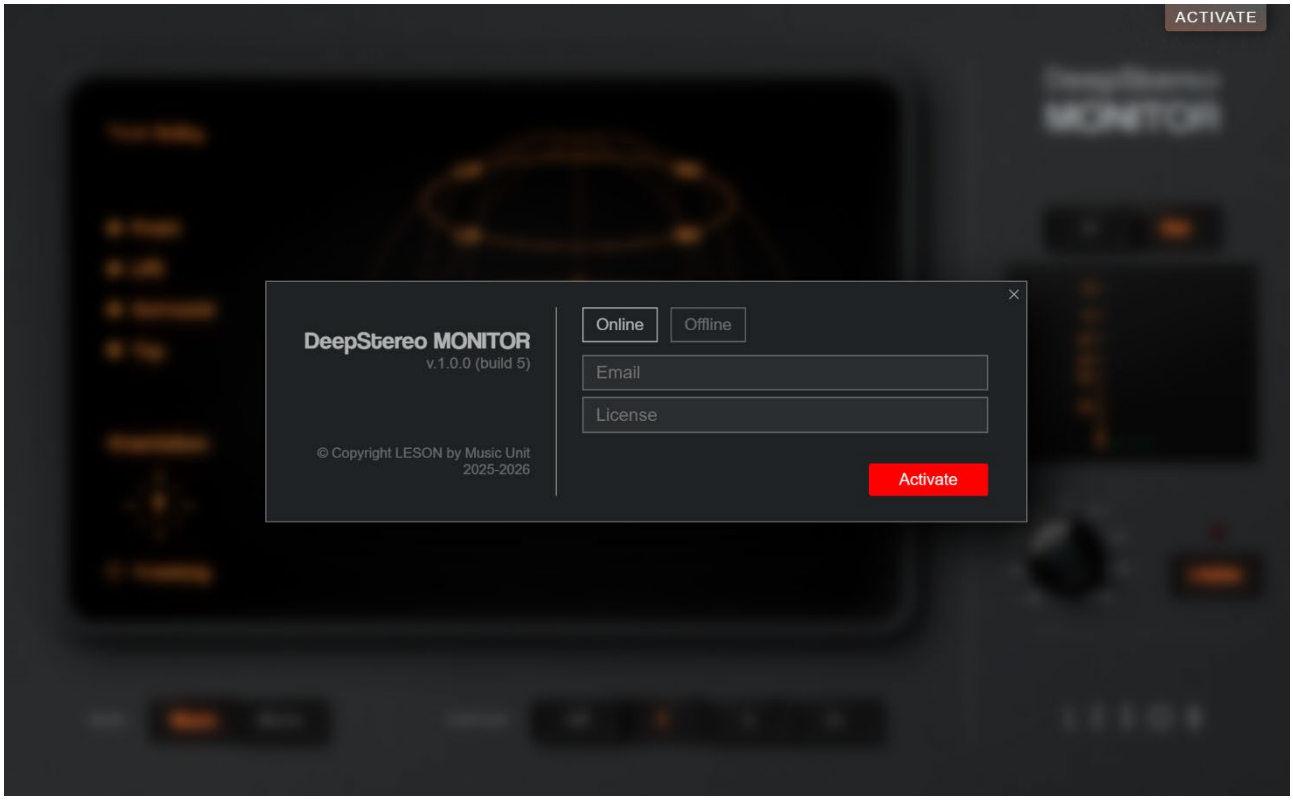
DeepStereo Monitor is available for download from **this Web page**.

Select your platform - macOS or Windows - and download the corresponding installer.

Double-click the installer icon and follow the instructions.

Launch your preferred DAW and insert the DeepStereo Monitor plugin on an audio track in the desired format - VST3, AAX, or AU - depending on your OS and DAW.

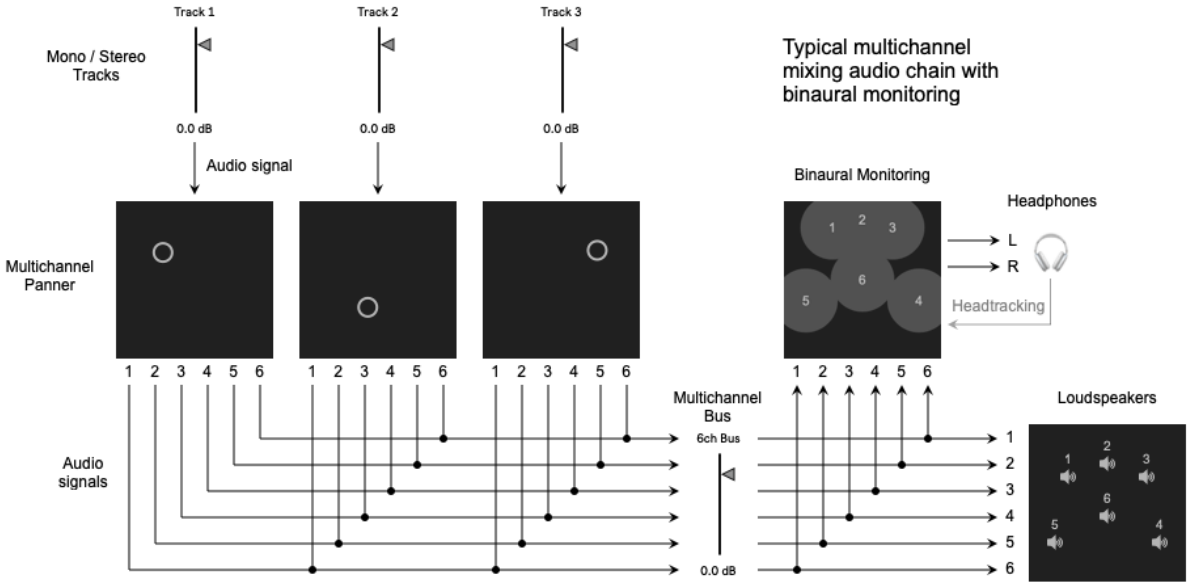
Within the plugin interface, click “ACTIVATE” and enter your email address and the license key received by email upon purchase.



Integration with various DAWs

Each DAW on the market offers a specific mechanism for implementing a multichannel mixing session with binaural monitoring. Rather than listing each of these mechanisms exhaustively, we can refer to the diagram below, which describes a typical audio chain involving main monitoring on speakers and, in parallel, binaural monitoring on a pair of headphones.

Here, we will use a 6-channel mixing session in 5.1 format as an example:



We will typically start by creating an audio bus track - auxiliary bus or master bus - each of whose separate outputs will be routed to a dedicated speaker via the converters of the audio interface used.

On each mono or stereo source track of the mix (vocals, instruments, etc.), we will insert a panning plugin or a panner integrated into the DAW, as appropriate. The panner will be configured to the monitoring bus format - 5.1 in this case - and the panner outputs will be routed to this bus.

To insert binaural monitoring into the audio chain, an auxiliary bus will be created - depending on the DAW - with the number of channels corresponding to the session - 6 in this case. This auxiliary bus will then be fed with the outputs from the main monitoring bus.

In the binaural monitoring plugin, a "virtual speaker" configuration preset corresponding to the format of the mixing session will be selected - 5.1 in this case. The stereo outputs of the binaural monitoring will be routed to specific hardware outputs in order to feed only the headphones.

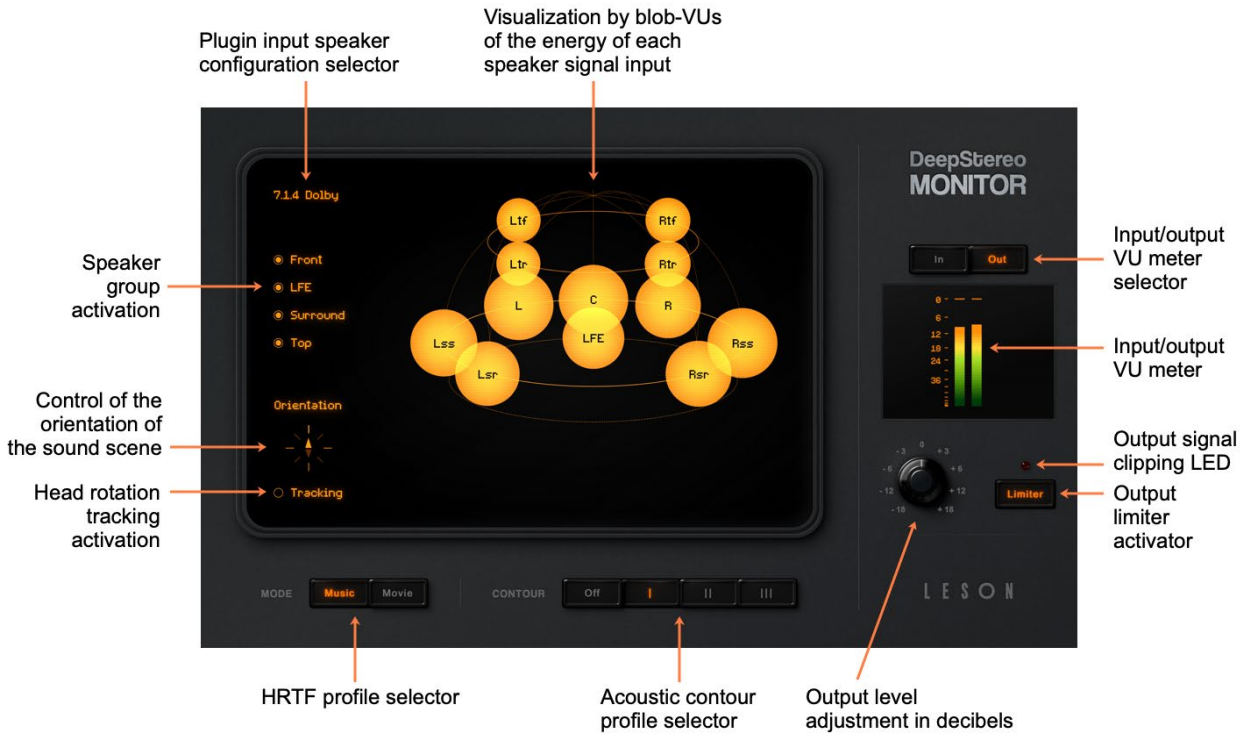
The goal is to create a channel parallel to the speaker monitoring, which "listens" to the main multichannel mix as closely as possible.

By creating two separate audio channels in this way, it is possible to avoid having to work in a multichannel-equipped studio all the time. It also allows you to export the binaural rendering at any time.

User interface

Simplified interface

The DeepStereo Monitor plugin interface is structured into two distinct sections. The central display visualizes the spatial sound scene and includes controls related to speaker configuration selection and editing.



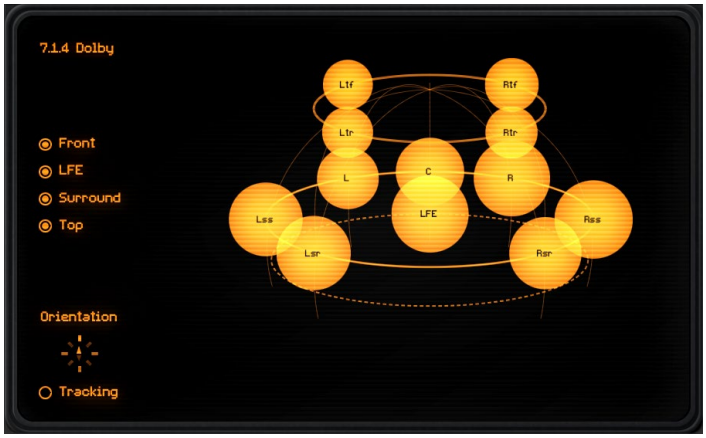
Surrounding the central display are binaural rendering controls such as HRTF profile selection, listening comfort presets, and output settings.

Dynamic GUI scaling

The interface size can be adjusted by click-dragging from any window corner.

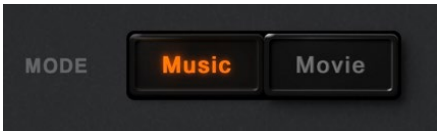
Below a certain size threshold, the interface reduces to essential spatial visualization.

This mechanism preserves visual feedback without compromising other DAW operations.



Binaural rendering parameters

Mode



The Mode selector allows selection of the binaural rendering profile (HRTF) according to content type.

Music mode provides highly transparent spectral response in the frontal plane and is optimized for music-dominant monitoring workflows - in which a large part of the sound information is delivered from the front plane.

Movie mode is optimized for music-for-visuals content with broader immersive distribution, including ambient music, dialogue, and sound effects.

Contour

The Contour selector provides multiple scene externalization profiles to enhance monitoring comfort.

Contour I offers an optimal balance between spatial impression and absence of temporal coloration. *Contour II* and *III* increase the sense of envelopment and are intended for mixes with a lot of room effects and/or reverberations. *Contour Off* corresponds to no externalization.

VU meters selection

The In/Out selector toggles between multichannel input meters (up to 32 channels) and binaural stereo output meters.

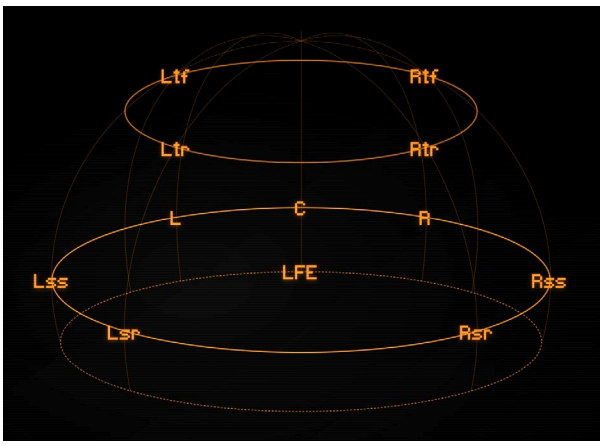
Binaural output management

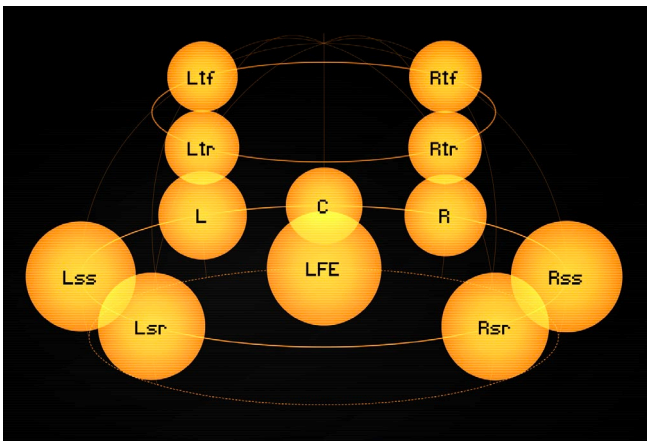
Binaural output level can be adjusted within a ±18 dB range. Double-clicking the gain control restores 0 dB default value.

A high-precision limiter can be engaged post-gain to prevent output clipping, indicated by the red LED.

Sound scene visualization

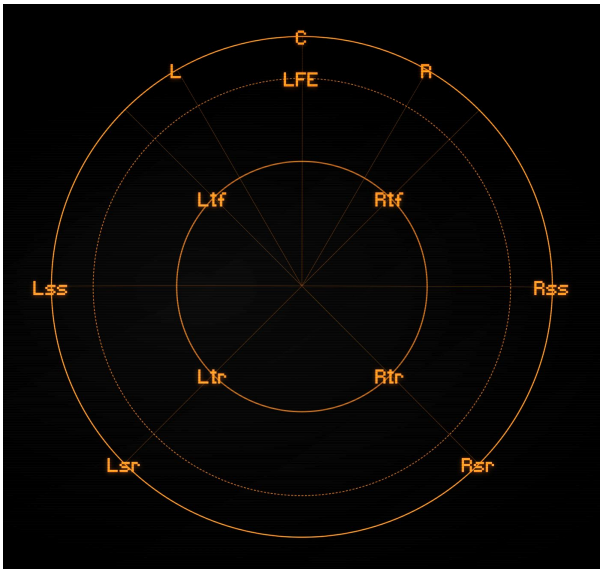
The central display represents the sound scene on a three-dimensional dome. Virtual speaker positions and labels depend on the selected configuration preset (7.1.4 Dolby, 5.1, 7.1, etc.).

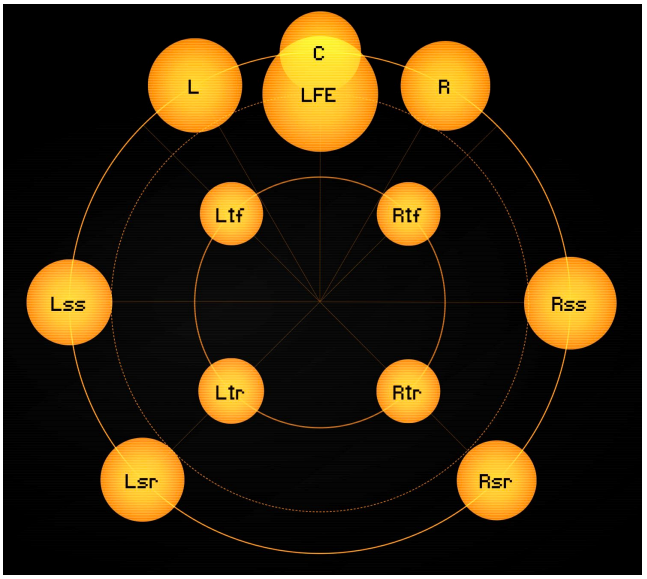




Input signal intensity is represented by dynamically sized blobs, each corresponding to one input channel, enabling rapid identification of spatial activity.

The sound scene can be vertically reoriented by click-dragging in the central area until a top view is reached.



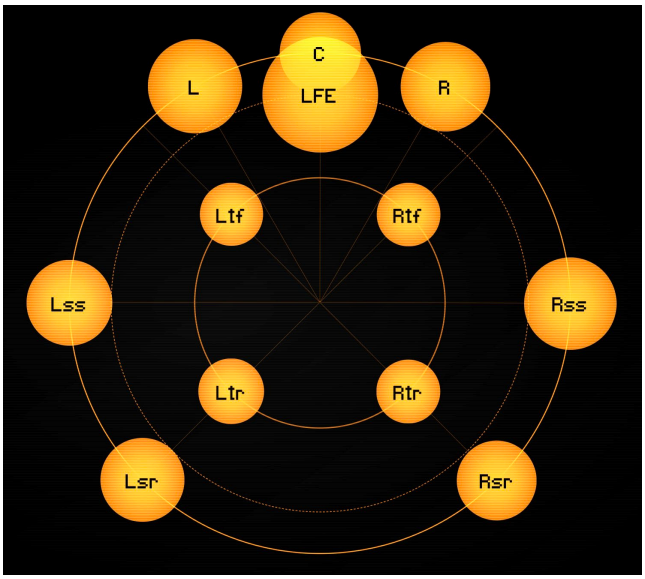


Double-clicking on the central area restores the default view.

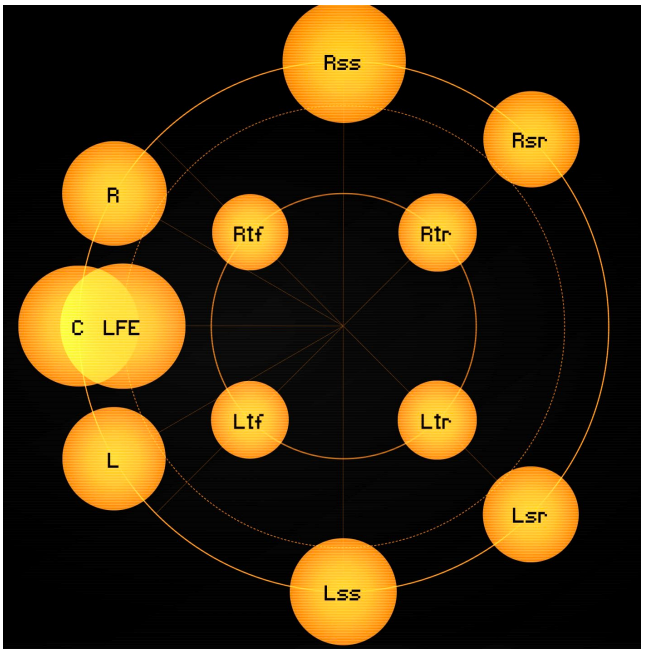
Sound scene orientation

The orientation around the vertical axis ("Yaw") of the sound scene can be adjusted using the Orientation slider. This is useful when it is necessary to check the correct spatial distribution of the sound scene.

For example, while it is naturally difficult to distinguish between sound activity in the front and rear planes, orienting the sound scene by ±90° allows for better comparison of the distribution of sound across these planes by momentarily shifting them onto the lateral auditory axis.



You can click on the axes of the Orientation control to perform a quick rotation.



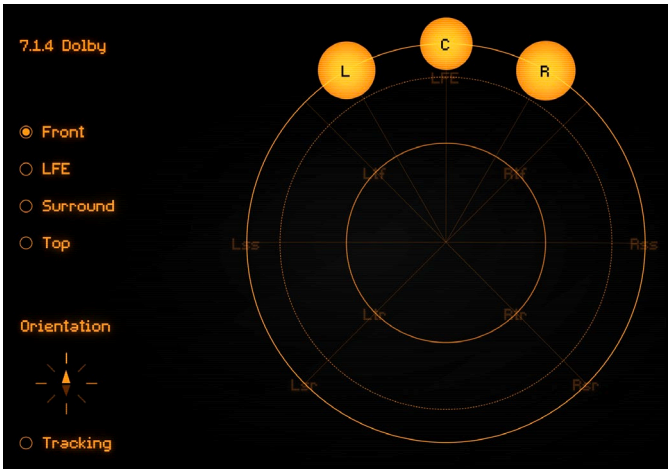
Double-clicking the Orientation control restores default orientation.

Speaker group activation

Within each configuration preset, speakers are organized into simplified spatial groups:

- Front: frontal plane up to ±45° azimuth
- LFE: low-frequency channel
- Surround: lateral plane beyond ±45° azimuth
- Top: elevated plane

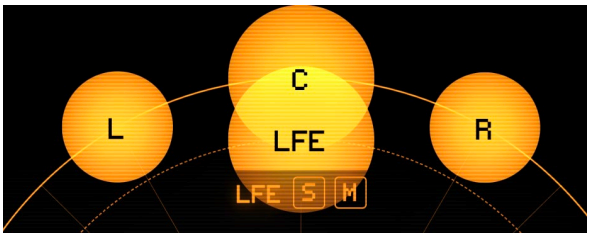
Each group can be activated or deactivated via dedicated switches. This is useful when it is necessary to quickly determine the sound activity in a given group.



Note: Shift-clicking on a switch isolates the corresponding group (solo).

Individual speaker activation

Each speaker can be individually activated, muted, or soloed to quickly evaluate specific channel activity. This is useful when it is necessary to quickly determine or control sound activity in a given audio input channel. Hovering over a speaker enables dedicated Solo [S] or Mute [M] controls.



A shortcut to the group switching to which the speaker belongs can also be used to quickly isolate the signals from all the speakers in that group.

Head tracking

DeepStereo Monitor's binaural rendering compensates for the listener's head movements, provided that a physical head tracking device is used in conjunction with headphones. This allows the position of a sound source to be perceived as fixed relative to the actual orientation of the head, while improving the externalization of the entire sound scene.

The Tracking switch activates the rotation tracking of the head physically associated with the headphones:

DeepStereo Monitor supports the **Supperware 1** head tracker. Supperware 1 is a device that can be easily attached to a headset and generates gyroscopic data via a network - Euler or quaternions - with low latency.

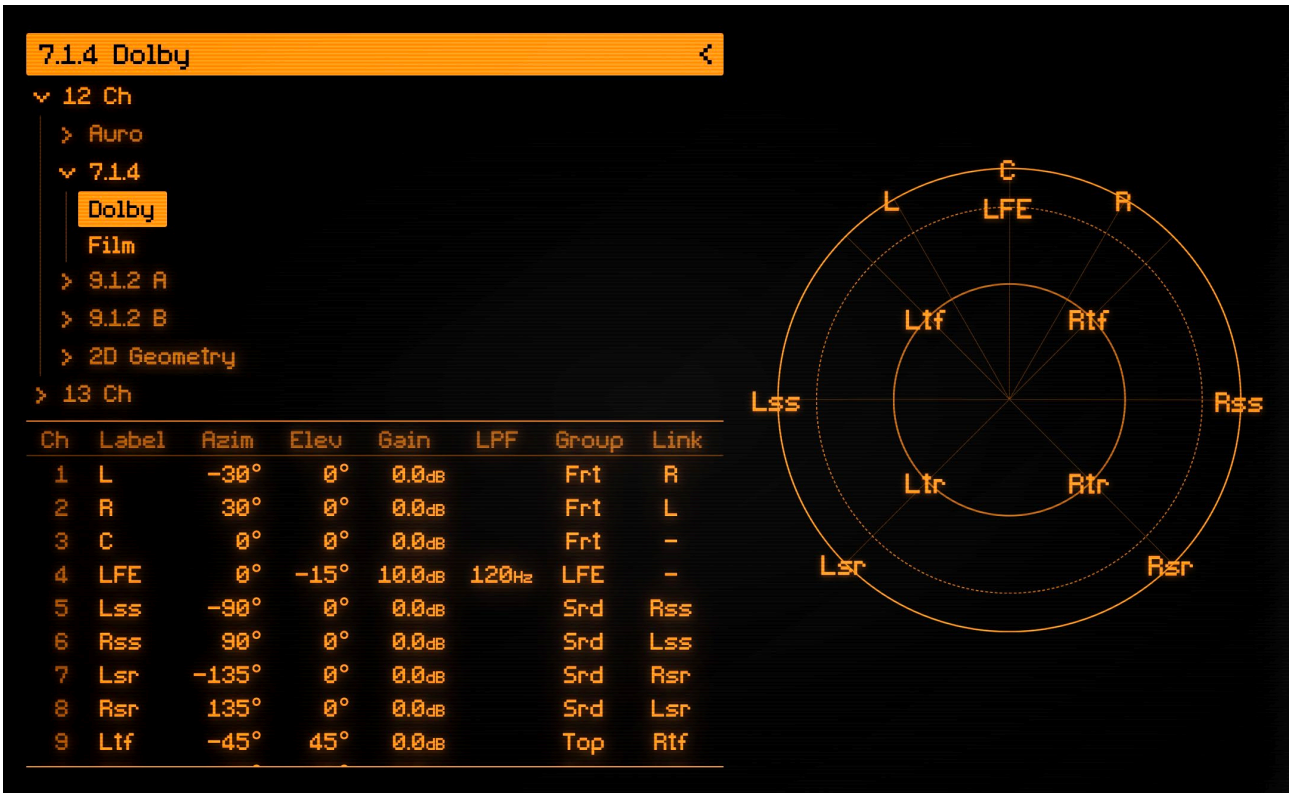
The Supperware 1 settings - available from the Supperware **Bridgehead** application - that allow you to control DeepStereo Monitor head tracking are as follows:

- IP address: local [127.0.0.1]
- UDP port number: 8004
- Open Sound Control message: /[yaw,pitch,roll]
- Output data: yaw (deg), pitch (deg), roll (deg)
- Default values: yaw: 0°, pitch: 0°, roll: 0°

Speaker configuration presets

This section describes the mechanism for applying factory presets for speaker configurations.

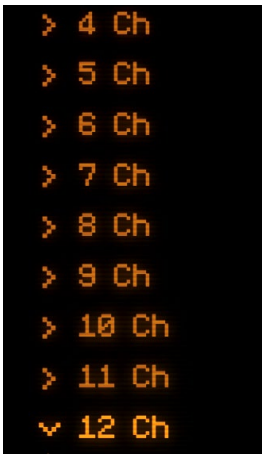
When you click on the label showing the name of the current speaker configuration, the configuration display changes to show a configuration selector and an input channel data editor representing each virtual speaker.



Ch	Label	Azim	Elev	Gain	LPF	Group	Link
1	L	-30°	0°	0.0dB		Frt	R
2	R	30°	0°	0.0dB		Frt	L
3	C	0°	0°	0.0dB		Frt	-
4	LFE	0°	-15°	10.0dB	120Hz	LFE	-
5	Lss	-90°	0°	0.0dB		Srd	Rss
6	Rss	90°	0°	0.0dB		Srd	Lss
7	Lsr	-135°	0°	0.0dB		Srd	Rsr
8	Rsr	135°	0°	0.0dB		Srd	Lsr
9	Ltf	-45°	45°	0.0dB		Top	Rtf

Available channel count

When you insert the plugin on a track, the number of channels on that track determines the maximum number of input channels supported by the DeepStereo Monitor plugin.



Thus, a track in 7.1.4 format offers up to 12 audio channels, while a track in 5.1 format will only offer 6 channels. Depending on the number of channels available on the track, DeepStereo Monitor adapts the available configurations, limiting the possible choices to those that do not exceed this number.

It is therefore possible to select a configuration with fewer channels than the track, for a more personalized use.

DeepStereo Monitor can accommodate up to 32 channels at its input.

Track format detection based on available channels

When DeepStereo Monitor is inserted on a track, it detects the number of channels available on that track and suggests a default configuration corresponding to the most common use for that number of channels.

For example, a 12-channel track, typically used for a 7.1.4 configuration, will cause DeepStereo Monitor to load a 7.1.4 configuration by default. A 6-channel track, typically used for a 5.1 configuration, will cause DeepStereo Monitor to load a 5.1 configuration by default.

Configuration detection only applies when DeepStereo Monitor is inserted and does not affect subsequent changes, such as selecting a different configuration - for example, using a “hexagonal” configuration rather than a 5.1 configuration for a 6-channel track.

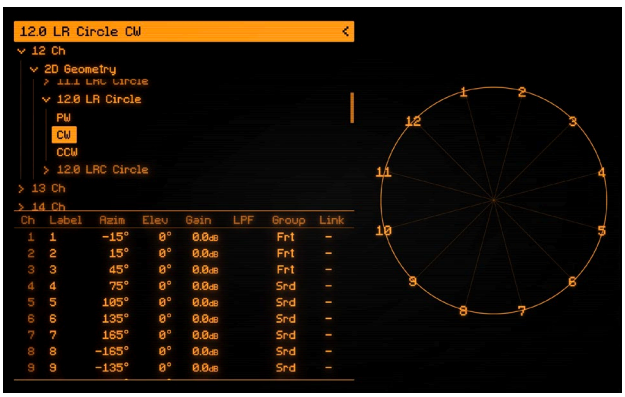
Navigating configuration presets

DeepStereo Monitor offers a large number of available configurations, classified in a systematic order:

Number of channels
- standards (Dolby, Auro, SMPTE, film, etc.)
- 2D geometry (circle, paired)

Click on the tabs to expand or collapse the subsets.

Slide the cursor on the right to scroll through the collection vertically.



Technical concepts

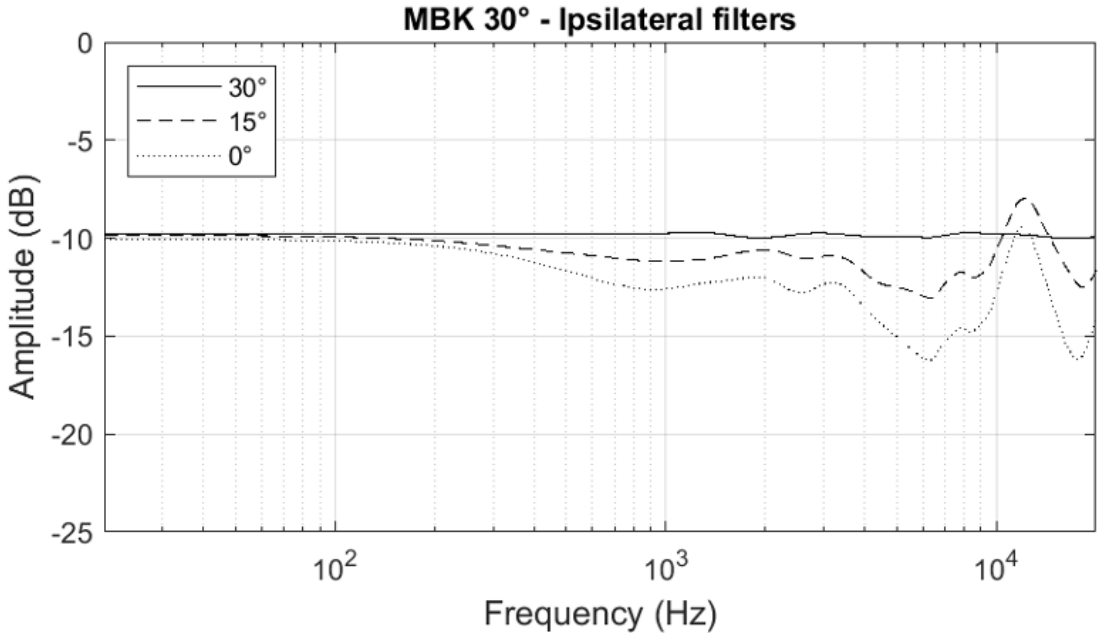
About DeepStereo, innovative binaural technology

The DeepStereo Monitor plugin is based on an innovative binaural synthesis engine, designed with a simple perceptual goal in mind: when a listener hears a signal on speakers, then hears the same signal on headphones, they should feel like they are hearing the same sound content, with the same tonal balance and spatial intent. The listener does not explicitly perceive the absence of transfer functions related to their morphology, which are naturally involved when listening on speakers.

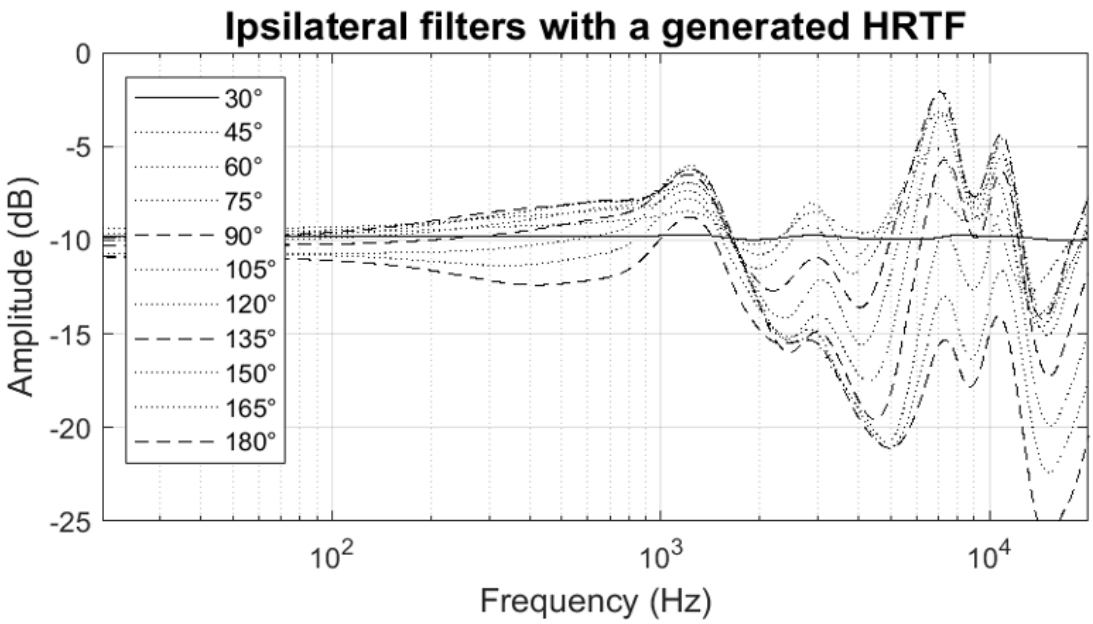
This observation highlights the ability of the human auditory system to naturally compensate for acoustic and morphological phenomena occurring between a sound source and the ear, and to maintain a consistent perception of sound content. The DeepStereo Monitor engine follows this logic: it aims to reproduce a perception equivalent to that obtained on speakers when using headphones, by attempting to eliminate natural psychoacoustic compensation while retaining all the spatial cues necessary for the localization and stability of the sound image.

Universal transfer functions

The transfer functions used by the DeepStereo Monitor's binaural synthesis engine are derived from a set of measurements taken on around 100 subjects in an anechoic chamber. These measurements were averaged to obtain a statistically representative HRTF, then subjected to a perceptual normalization process. The fundamental principle of this normalization is to define a reference angle for which the frequency response of the ipsilateral signal - i.e., the signal received first by the ear closest to the source - is made as flat as possible. This reference angle usually corresponds to the front left and right speakers



This normalization is then consistently compensated for across all transfer function amplitudes - i.e., only frequency responses. Furthermore, the other spatialization indices - interaural time differences (ITD), level differences (ILD), and phase differences (IPD) - are fully preserved.



The result is a coherent, stable, and controllable binaural space based on the preservation of spatial relationships and cues, rather than on an exact reproduction of individual transfer functions, averaged or otherwise.

Binaural rendering engine and management of interactions between sources

The direct application of binaural transfer functions alone is not sufficient to guarantee perceptual neutrality. Indeed, while the filtering applied to the ipsilateral signal can be considered neutral for a given direction, the actual perception always results from the combination of the ipsilateral and contralateral signals.

The DeepStereo Monitor's binaural rendering engine therefore incorporates specific processing designed to compensate for the cumulative influence of contralateral filters, as well as interactions related to common content between sources.

This processing preserves tonal and spatial stability in complex mixing configurations, while maintaining the integrity of the binaural cues necessary for reliable spatial perception.

About

Leson by Music Unit

Leson is Music Unit's “haute couture” sound studio, dedicated to designing innovative audio solutions. Motivated by our long-standing partners - artists, sound engineers, and electronics manufacturers - who want us to translate our expertise into standalone tools that are accessible to everyone, we are developing a portfolio of patented inventions and high-end audio products for both content producers and broadcasters.

LESON

Music Unit

Music Unit is a French music and sound creation studio founded in 2004 that approaches the music and audio industry through creativity and musical innovation as well as technology.

Music Unit brings together artists, producers, and engineers who are masters of emerging audio technologies - spatial audio, AI, VR, AR - and want to share their knowledge and talents. Our mission is to offer audiences the opportunity to enjoy new artistic experiences.

Contact : Music Unit 28 Rue Kléber 93100 Montreuil Grand Paris – France support@musicunit.fr

Credits

Martin Antiphon: DSP development, audio design
Jérôme Boivin: quality assurance, distribution
Éric de Broche des Combes: graphic design
Alexandre Chaigne: product manager, distribution
Julien Chirol: team manager
Pierre Luzy: audio design & fine-tuning, documentation
Manuel Poletti: graphic design, documentation
Kevin Schires: DSP development, UI development, graphic design
Team Leson: product design