IRIS: A VST3 PLUGIN FOR 2D NAVIGATION OF IMPULSE RESPONSE COLLECTIONS

Luna Valentin
CCRMA
Stanford University
CA, USA
luna@ccrma.stanford.edu

Celeste Betancur Gutiérrez
CCRMA
Stanford University
CA, USA
celes@ccrma.stanford.edu

Romain Michon
Emeraude Team
INRIA
Lyon, FR
romain.michon@inria.fr

ABSTRACT

Impulse response (IR) collections are useful in virtual acoustics, sound design, and field-based acoustic research, but they remain difficult to explore as continuous resources in lightweight real-time plugin workflows. This demo paper presents IRIS, a VST3 plugin for arranging, navigating, and auditioning measured or user-defined IR collections in a two-dimensional navigation plane. Each IR is represented as a node whose position can be imported from metadata or assigned manually. During navigation, nearby responses are combined using Gaussian distance-based weighting, while a bounded active set limits the number of simultaneous convolutions. The system also includes smoothing, hysteresis, optional preprocessing, boundary attenuation, OSC control, and coupled multichannel handling. The demo focuses on workflow and audible behavior rather than perceptual validation. A short timing characterization reports practical real-time limits as a function of IR length, buffer size, and active-set size. IRIS is presented as a practical tool for exploratory, analytical, and creative navigation of IR collections rather than as a physically optimal interpolation method.

1. INTRODUCTION

Impulse response (IR) convolution is widely used to reproduce or transform acoustic responses in music production, virtual acoustics, immersive audio, and sound design [1, 2]. Measured IR collections often capture variation across positions, configurations, or acoustic conditions, but they remain difficult to explore as continuous resources in lightweight plugin workflows. Users often have to select individual responses, organize them manually, or build custom routing and crossfading strategies.

Substantial work has addressed dynamic convolution, navigable auralization, spatial room impulse response interpolation, and Ambisonic perspective rendering [3, 4, 5, 6, 7]. IRIS addresses a narrower workflow problem: arranging, navigating, and auditioning measured or user-defined IR collections as two-dimensional fields inside a VST3 plugin.

In IRIS, each IR is represented as a node in a normalized navigation plane. Node positions can be imported from JSON or CSV metadata or assigned manually. A navigation cursor defines the current point of evaluation, and nearby responses are combined using Gaussian distance-based weighting. A bounded active set limits the number of simultaneous convolutions, while smoothing, hysteresis, and boundary attenuation shape navigation behavior. The contribution is practical and system-oriented. IRIS is not presented as physically optimal IR interpolation or perceptually transparent rendering, but as a usable real-time tool for exploratory, analytical, and creative navigation of IR datasets.

2. RELATED WORK AND POSITIONING

IRIS sits between real-time dynamic convolution tools, RIR interpolation methods, and spatial or Ambisonic rendering systems. SPARTA and COMPASS provide real-time spatial-audio plugins for spatialisation, production, visualisation, and manipulation of spatial sound scenes [3]. SPARTA 6DoFconv and MCFX-6DoFconv extend this ecosystem toward real-time six-degree-of-freedom convolution workflows for SOFA-based RIR datasets [8]. RoomZ is especially close to IRIS: it provides a plugin workflow for dynamic RIR convolution auralizations over one- or two-dimensional spaces, using nearest-neighbor or triangle-based interpolation over measured RIR grids [4]. IRIS is therefore not the first plugin for real-time RIR navigation. Its contribution is a different lightweight workflow based on user-arranged or metadata-imported IR nodes, Gaussian weighting, bounded active-set convolution, smoothing, OSC control, coupled multichannel handling, and user-defined boundary attenuation.

RIR interpolation methods often make stronger physical or perceptual assumptions than IRIS, explicitly treating direct sound, early reflections, late reverberation, phase relationships, or directional information [9, 10, 11, 12, 13, 5]. This work shows that simple time-domain blending can introduce temporal smearing, comb filtering, coloration, or unstable source perception. IRIS does not replace perceptually informed spatial RIR interpolation. It implements a simplified strategy directly inside a real-time plugin.

Directional and Ambisonic RIR datasets add constraints because their channels encode a coupled sound field [14, 15, 6, 16]. IRIS treats each multichannel IR as one node and applies the same active-set selection, weights, smoothing, onset alignment, and normalization to all channels. This avoids independent per-channel interpolation, but does not implement Ambisonic perspective translation, direction-of-arrival estimation, or directional early-reflection reconstruction.

3. IRIS SYSTEM

IRIS represents a collection of impulse responses as points in a normalized two-dimensional plane $[0, 1] \times [0, 1]$ (Fig. 1). Each IR h_i has a position (x_i, y_i), imported from metadata or assigned manually. The plane may correspond to measured coordinates, but it can also organize IRs by timbre, microphone position, instrument, or another user-defined logic. A cursor (x_c, y_c) defines the

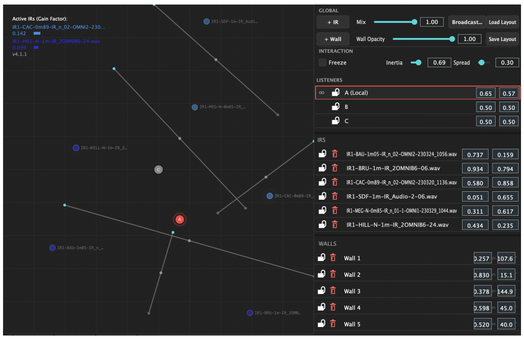

Figure 1: IRIS interface with navigation plane, IR nodes, cursor, and user-defined boundary elements.

current point of evaluation and can be controlled from the interface, DAW automation, MIDI, or OSC.

At each update, IRIS computes the squared distance between the cursor and each IR node,

$$d_i^2 = (x_i - x_c)^2 + (y_i - y_c)^2,$$

and assigns an unnormalized Gaussian weight,

$$w_i = \exp\left(-\frac{d_i^2}{2\sigma^2}\right).$$

The spread σ is controlled by a normalized **Spread** parameter $s \in [0, 1]$, with $\sigma = 0.05 + 1.5s^2$.

IRIS does not convolve all loaded IRs. The top $k_{\min} = 4$ IRs are always included, and additional candidates up to $k_{\max} = 8$ are selected with hysteresis. With $w_{\max} = \max_i w_i$, inactive IRs enter the active set when $w_i \geq 0.10 w_{\max}$ and remain active until $w_i < 0.05 w_{\max}$. Active weights are normalized over the current active set A:

$$\tilde{w}_i = \frac{w_i}{\sum_{j \in A} w_j}.$$

A first-order filter with coefficient $\alpha = 0.25$ smooths active weights at control rate. Cursor inertia optionally smooths abrupt user input before weights are computed.

IRIS can apply optional onset alignment and RMS normalization. These operations may reduce temporal smearing or level jumps, but they can also remove meaningful propagation-delay, distance, or source-level cues. They are therefore workflow options rather than required corrections.

User-defined boundaries provide an additional weighting control. When the line from the cursor to an IR node intersects a boundary segment, that IR weight is multiplied by $(1 - \beta)$, where $\beta \in [0, 1]$ is the boundary opacity. This broadband attenuation can create occlusion-like behavior, but it is not a diffraction, transmission, or wall-proximity model.

The output is computed as

$$y(t) = \sum_{i \in A} \tilde{w}_i \, (x * h_i)(t).$$

Real-time convolution uses JUCE's `dsp::Convolution`, with one convolution engine per active IR. Multichannel IRs are treated as single nodes: active-set selection, weights, smoothing, onset alignment, and normalization are computed once and applied identically to all channels.

4. DEMONSTRATION

The demo presents IRIS as an interactive tool for arranging, navigating, and auditioning IR collections. Users load an IR dataset, inspect the nodes in the navigation plane, move the cursor between responses, and modify the parameters that shape the interpolation.

Table 1: Main interactions demonstrated in IRIS.

Interaction	Demonstrated behavior
IR layout	Import positions or arrange nodes manually.
Navigation	Audition continuous movement between IRs.
Spread	Change the locality of Gaussian weighting.
Smoothing	Compare direct and smoothed cursor motion.
Boundaries	Draw attenuation boundaries in the plane.
Preprocessing	Toggle onset alignment and normalization.
Multichannel IRs	Use shared interpolation state across channels.

The demonstration focuses on workflow and audible behavior rather than perceptual validation. It shows IRIS with measured layouts, manually arranged timbral maps, and hybrid control spaces.

A timing benchmark was run on an Apple M4 Pro at 48 kHz using a standalone build of the plugin processor. With a 512-sample buffer, 16 active 1 s IRs required 2.87 ms on average, 8 active 4 s IRs required 5.60 ms, and 4 active 10 s IRs required 7.10 ms. These conditions remained below the 10.67 ms buffer duration. Longer IRs and larger active sets exceeded the budget. The current implementation is therefore suitable for exploratory listening and studio workflows with moderate buffers, but not for ultra-low-latency monitoring with very long IRs and large active sets.

5. CONCLUSION

IRIS is a lightweight VST3 plugin for navigating impulse-response collections in a two-dimensional field. It combines Gaussian weighting, bounded active-set convolution, smoothing, hysteresis, optional preprocessing, boundary attenuation, OSC control, and coupled multichannel handling.

The contribution is practical rather than physically comprehensive. IRIS gives users a controllable way to explore measured or user-arranged IR collections while keeping the limits of simplified interpolation explicit. Future work will address large-dataset scalability, additional interpolation modes, perceptual tests, and stronger validation for Ambisonic use.

6. REFERENCES

[1] A. Harker and P. A. Tremblay, "The HISSTools impulse response toolbox: Convolution for the masses," in *Proceedings of the International Computer Music Conference (ICMC)*, 2012, pp. 148–155.

[2] W. G. Gardner, "Efficient convolution without input-output delay," *Journal of the Audio Engineering Society*, vol. 43, no. 3, pp. 127–136, 1995.

[3] L. McCormack and A. Politis, "SPARTA & COMPASS: Real-time implementations of linear and parametric spatial audio reproduction and processing methods," in *Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio*, 2019. [Online]. Available: https://aaltodoc.aalto.fi/items/72a60211-d51a-4404-b90d-096ae3970b97

[4] D. Poirier-Quinot, P. Stitt, and B. F. G. Katz, "RoomZ: Spatial panning plugin for dynamic RIR convolution auralisations," in *AES 2023 International Conference on Spatial and Immersive Audio*, 2023. [Online]. Available: https://hal.science/hal-04188949/document

[5] T. McKenzie, N. Meyer-Kahlen, R. Daugintis, L. McCormack, S. J. Schlecht, and V. Pulkki, "Perceptually informed interpolation and rendering of spatial room impulse responses for room transitions," in *Proceedings of the 24th International Congress on Acoustics*, 2022, pp. 1–11. [Online]. Available: https://aaltodoc.aalto.fi/items/84500f3a-6aed-400f-be62-f130736a1631

[6] K. Müller and F. Zotter, "Auralization based on multi-perspective Ambisonic room impulse responses," *Acta Acustica*, vol. 4, no. 6, p. 25, 2020.

[7] J. De Muynke, D. Poirier-Quinot, and B. F. G. Katz, "Effect of interpolation artifacts on perceived stability of nearby sources in a navigable reverberant virtual environment," *Journal of the Audio Engineering Society*, vol. 72, no. 10, pp. 664–678, 2024.

[8] D. Stefani, M. Binelli, A. Farina, and L. Turchet, "Improved real-time six-degrees-of-freedom dynamic auralization through nonuniformly partitioned convolution," *Journal of the Audio Engineering Society*, vol. 73, no. 10, pp. 671–681, 2025.

[9] G. Kearney, C. Masterson, and F. Boland, "Acoustic impulse response interpolation for multichannel systems using dynamic time warping," in *Audio Engineering Society 35th Conference*, Feb. 2009. [Online]. Available: https://pure.york.ac.uk/portal/en/publications/acoustic-impulse-response-interpolation-for-multichannel-systems-

[10] N. Antonello, E. De Sena, M. Moonen, P. A. Naylor, and T. van Waterschoot, "Room impulse response interpolation using a sparse spatio-temporal representation of the sound field," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 10, pp. 1929–1941, 2017.

[11] O. Das, P. Calamia, and S. V. Amengual Garí, "Room impulse response interpolation from a sparse set of measurements using a modal architecture," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 960–964.

[12] F. Katzberg, R. Mazur, M. Maaß, M. Böhme, and A. Mertins, "Spatial interpolation of room impulse responses using compressed sensing," in *Proceedings of the International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2018, pp. 426–430.

[13] J. Zhao, X. Zheng, C. Ritz, and D. Jang, "Interpolating the directional room impulse response for dynamic spatial audio reproduction," *Applied Sciences*, vol. 12, no. 4, p. 2061, 2022.

[14] J. G. Tylka and E. Y. Choueiri, "Comparison of techniques for binaural navigation of higher-order ambisonic soundfields," in *Audio Engineering Society Convention 139*, Oct. 2015, paper 9421. [Online]. Available: https://aes.org/publications/elibrary-page/?id=17977

[15] ——, "Models for evaluating navigational techniques for higher-order ambisonics," *Proceedings of Meetings on Acoustics*, vol. 30, no. 1, p. 050009, 2017.

[16] K. Müller and F. Zotter, "The PerspectiveLiberator: An upmixing 6DoF rendering plugin for single-perspective Ambisonic room impulse responses," in *Proceedings of DAGA 2021*, 2021. [Online]. Available: https://pub.dega-akustik.de/DAGA_2021/data/articles/000320.pdf