

From Waveform to Wavefield: Observable Field Optimization for Music Reproduction

Xuanzhen Zhu

Moodify Sound Laboratory, Rongjing Wenchuan (Shenzhen) Technology Co., Ltd., Shenzhen, China

Manuscript type: Open Technology Note - Canon and research program

Abstract-Digital audio is not the acoustic event itself; it is a condition from which an event may be reproduced. This note defines Moodify's transition from waveform-only processing to wavefield observation and optimization. A candidate field includes a target space, reproduction chain, propagation conditions, listener position, and evidence scope. Moodify HEAR observes the source and reproduced field; Moodify Judge determines whether a candidate improves the listening result while preserving identity, safety, and recoverability. The north-star formulation is $F^* = \arg \max J(\text{Render}(\text{Audio}, F))$, refined in production as constrained robust optimization across target playback conditions. The note separates measured, simulated, and inferred evidence; rejects claims that a stereo file alone reconstructs a physical three-dimensional sound field; and proposes a Wavefield Profile, conditional renders, Judge records, and rollback as the persistent output.

Keywords-wavefield; audio intelligence; spatial reproduction; constrained optimization; auditory evidence; human review

1 Introduction

Most audio software treats a released file as the primary object of intervention. Equalization, dynamics, source separation, restoration, spatial processing, and mastering alter the waveform or generate a derived waveform. This view is useful but incomplete: a digital file does not become an acoustic event until a reproduction chain converts it into pressure variation in a physical or modeled space.

Moodify therefore distinguishes *waveform* from *wavefield*. A waveform is an encoded source. A wavefield is the conditioned event produced by a source, rendering process, transducer configuration, propagation environment, and listener position. The technical objective changes from finding a better file to finding a field in which the work is reproduced more effectively.

The contribution of this note is architectural rather than a claim of solved three-dimensional acoustics. It defines the optimization target, evidence boundary, persistent artifacts, and a staged research program. Current Moodify production capability remains file-domain analysis and stereo projection; calibrated physical-field capture is future work.

2 Problem Formulation

Let A denote immutable source audio and F a candidate wavefield description. Let $R(A, F)$ be a rendering or physical reproduction under field F , and let J be Moodify Judge. The north-star objective is

$$F^* = \arg \max_F J(\text{Render}(\text{Audio}, F)) \quad (1)$$

Equation (1) is deliberately compact. It states that Moodify searches an acoustic possibility space rather than applying a fixed preset. It does not imply that J is already a calibrated universal scalar, nor that F can be identified from a stereo file alone.

2.1 Production Objective

For production, the optimization must be robust across declared playback conditions c drawn from a target set C . It also depends on source identity and creative intent I :

$$F^* = \arg \max_{(F \text{ in } \Omega)} E_c[J(\text{Render}(A, F; c), A, I)] \quad (2)$$

The feasible set is constrained by identity preservation, loudness-matched comparison, reproduction safety, evidence completeness, reproducibility, uncertainty, and human authority. If no feasible candidate demonstrates positive net gain over the original playback baseline, the correct result is `KEEP_ORIGINAL`.

3 Definitions and Evidence Classes

Term	Operational meaning
Waveform	Immutable source samples and provenance.
Wavefield	Target space, reproduction chain, propagation conditions, and observation positions.
Render	Simulated rendering or calibrated physical reproduction.
HEAR	Observation of source evidence or reproduced-field evidence.
Judge	Scoped authority for accept, retry, adjust, original, or human escalation.
F^*	Best supported field inside the stated target and evidence scope.

Evidence is labeled **measured**, **simulated**, or **inferred**. A visualization cannot promote simulated or inferred content into measurement. The label follows the artifact into every report and Judge decision.

4 System Architecture

The Wavefield loop preserves the source while allowing the field and rendering conditions to evolve. The same Core owns observations, profiles, verification, and Judge authority across Studio, API, CLI, and App interfaces.

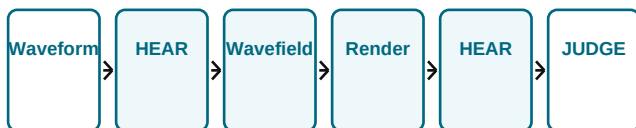


Figure 1: Moodify wavefield observation and constrained selection loop.

4.1 Two-Judge Invariant

Judge Before establishes what happened in the source, whether intervention is justified, and which traits must be protected. Judge After compares source and candidate under aligned conditions. A renderer, studio, model, or service provider may produce a candidate but cannot certify its own success.

4.2 Replaceable Realization Providers

Simulation engines, binaural renderers, loudspeaker systems, rooms, measurement arrays, and external studios are field-realization providers. They may be replaced without creating a second Core or an independent judgment authority.

5 Observability

Moodify's claim is not that every acoustic property is recoverable. Its claim is that previously opaque relationships can be represented with explicit limits. A source file supports time-frequency evolution, dynamics, phase, channel correlation, and stereo amplitude-density observations. A calibrated reproduction setup adds transfer responses and position-specific pressure measurements.

Observable	Supported conclusion	Unsupported conclusion
Time-frequency energy	Spectral evolution in the source	Complete room pressure field
L/R phase and density	Stereo projection and phase risk	Stable image at arbitrary positions
Binaural or array capture	Measured field at sampled positions	Facts at unsampled positions
Propagation simulation	Result under a stated model	Equivalent physical measurement

5.1 Physical Truth Boundary

A mono or stereo file cannot establish a real three-dimensional pressure field. Such a field requires a calibrated playback chain, room description, microphones or a binaural head or array, positions, timing, and environmental metadata. The system must refuse a three-dimensional measurement claim when those prerequisites are absent.

6 Wavefield Search

The search space can include room scale, early reflection timing, decay structure, diffusion, direct-to-reverberant ratio, channel or object layout, listener position, transducer response, and content-dependent rendering parameters. Search may be Bayesian, evolutionary, gradient-free, or human-guided; the method is secondary to governed evidence and repeatable evaluation.

6.1 Multi-objective Judge

Judge should initially produce a structured decision rather than a single aesthetic score. Candidate quality may include identity fidelity, clarity, spatial stability, dynamic integrity, translation, safety, and listening preference. Scalarization weights remain human-governed until calibrated evidence supports a narrower machine authority.

$$J = \{\text{identity, clarity, space, dynamics, translation, safety, uncertainty}\} \quad (3)$$

6.2 Robustness Across Conditions

A field that succeeds only at one chair in one expensive room may still be useful, but its scope must be explicit. A general-purpose profile should be evaluated across target devices, positions, and environments. Expected benefit and worst-case regression should both be reported.

7 Studio-in-the-loop

A physical studio closes the gap between source-domain analysis and real acoustic occurrence. The experiment should compare AI-generated and recorded works under the same calibrated chain, matched loudness, fixed listener positions, and blind review. Multi-position measurements should be synchronized with the digital source and room state.

Stage	Required record
Source	Digest, codec, sample rate, channels, rights, intent
Playback	DAC, amplifier, loudspeakers, gain, calibration
Space	Geometry, treatment, temperature, positions
Capture	Microphone or array, orientation, clock, response
Review	Blind protocol, loudness match, reviewer, uncertainty

The first experiment should seek reproducible differences, not confirm a preferred story. A null or contradictory result is valid evidence and may return the physical-field program to research status.

8 Persistent Artifacts

A single stereo WAV cannot reproduce the same physical field on arbitrary equipment. Moodify therefore preserves a compound result rather than an opaque enhanced file.

Artifact	Purpose
Wavefield Profile	Versioned field target, parameters, applicable conditions
Conditional renders	Headphone, stereo, multichannel, object, or named-studio delivery
Observation evidence	Measured, simulated, and inferred evidence with uncertainty
Judge record	Before/after rationale, protected traits, review, final choice
Rollback	Original source, bypass path, and KEEP_ORIGINAL

9 Implementation Status

The repository currently contains file-domain wave-spectral analysis, loudness and dynamics measurement, phase and stereo metrics, evidence records, playback profiles, reversible Mix Graph concepts, and rule-based technical judgment with human escalation. A mature learning-based Judge-1 is not implemented. Wavefield schema, calibrated studio integration, physical capture, conditional render contracts, and cross-environment Judge calibration are not implemented.

10 Open Research Program

The staged program is: (1) define wavefield vocabulary and observation schema; (2) establish a deterministic Wavefield Profile and render contract; (3) validate a stereo projection baseline and binaural reference renderer; (4) conduct a calibrated studio-in-the-loop experiment; and (5) calibrate field selection and translation judgments.

11 Discussion

From waveform to wavefield is a change in the unit of audio engineering. The source remains essential, but value emerges in the controlled relation between source, reproduction, space, observation, and listener. This relation makes the system more demanding: target conditions must be declared, evidence provenance must survive rendering, and claims must remain scoped.

The approach also changes what openness means. Open code without open theory exposes implementation but hides the reasons that shape it. Moodify Open Technology Notes therefore publish the question, thesis, definitions, formula, method, evidence boundary, failure condition, status, and revision history alongside code.

12 Conclusion

Moodify seeks the field in which a musical work is best reproduced under explicit conditions. The compact objective $F^* = \arg \max J(\text{Render}(\text{Audio}, F))$ provides a durable research direction, while the production formulation constrains optimization by identity, safety, evidence, robustness, and human authority. The immediate task is not to claim solved spatial intelligence, but to build the vocabulary, contracts, experiments, and Judge calibration required to make wavefield selection observable and accountable.

Conflict of Interest

The author is the founder of Rongjing Wenchuan (Shenzhen) Technology Co., Ltd. and the designer of the Moodify project. The architecture is commercially relevant. No independent production-scale validation is claimed in this note.

Data and Code Availability

Moodify code, governance, and technical theory are developed in public at <https://github.com/huliye24/moodify-ai>. Capabilities marked proposed or research are not represented as production implementations.

AI-Assistance Disclosure

A generative language model assisted with manuscript structuring, drafting, and typesetting. The author supplied the technical direction and remains responsible for claims, equations, boundaries, and interpretations.

References

- [1] Moodify, "Wavefield Core Principle - From Waveform to Wavefield," Open Technology Note source, 2026.
- [2] ITU-R, "Algorithms to measure audio programme loudness and true-peak audio level," Recommendation ITU-R BS.1770-5, 2023.
- [3] ITU-R, "Method for the subjective assessment of intermediate quality level of audio systems," Recommendation ITU-R BS.1534-3, 2015.
- [4] ISO, "Acoustics - Measurement of room acoustic parameters," ISO 3382 series.
- [5] AES, "AES standard for file exchange - Spatial acoustic data file format," AES69.
- [6] W3C, "PROV-O: The PROV Ontology," W3C Recommendation, 2013.
- [7] Moodify, "Judge Core Principle," public canonical documentation, 2026.