Low-Dimensional HRTF Characterization Through Common-Pole/Zero Modeling and Statistical Component Analysis

Dr. Andreas Nicolaou
Department of Artificial Intelligence and Smart Computing
Cyprus Institute of Digital Technologies Nicosia, Cyprus

https://doi.org/10.37547/ibast/Volume06Issue08-03

Abstract.
Head-related transfer functions (HRTFs) provide a mathematical representation of the acoustic filtering imposed by the human head, torso, and external ears and constitute a fundamental component of three-dimensional spatial audio reproduction. However, the high dimensionality of measured HRTF datasets creates substantial challenges for storage, interpolation, transmission, and real-time rendering. This paper develops a low-dimensional HRTF characterization approach that integrates common-pole/zero modeling with statistical component analysis. The proposed methodology first represents individual HRTFs through a shared acoustic pole/zero structure and subsequently transforms the resulting parameter space into a compact statistical representation. The approach is theoretically motivated by established common-acoustical-pole/zero models and finite- and infinite-impulse-response HRTF representations. The proposed framework separates relatively stable acoustic characteristics from location-dependent variation, thereby reducing the number of parameters required to describe a spatial HRTF dataset. A methodological analysis demonstrates how common-pole/zero modeling can provide physically interpretable spectral parameters while component analysis can reduce redundancy among those parameters. The study further examines reconstruction, dimensionality, computational efficiency, and potential deployment in real-time spatial-audio systems. The resulting framework offers a structured compromise between acoustic fidelity, parameter compactness, and computational complexity. Its principal significance lies in combining parametric acoustic modeling with statistical dimensionality reduction rather than treating either technique independently.

Keywords: HRTF, common-pole/zero modeling, dimensionality reduction, principal component analysis, spatial audio, acoustic modeling, statistical representation, 3D audio, parametric modeling

INTRODUCTION

1.1 Background

Three-dimensional audio systems require accurate reproduction of the directional characteristics of sound perceived by a listener. Head-related transfer functions describe the frequency-dependent transformation between an incident sound field and the acoustic signal reaching the listener's ears. Consequently, HRTFs are central to binaural reproduction, headphone-based spatial audio, virtual acoustic environments, and real-time three-dimensional sound processing. The increasing interest in interactive spatial-audio applications has simultaneously increased the need for efficient HRTF representations.

A fundamental difficulty is that an HRTF database can contain measurements for numerous azimuth-elevation positions and may include separate left- and right-ear responses. Direct representation of every frequency-domain response is therefore computationally and storage intensive. Parametric representations provide an alternative by describing an HRTF using a comparatively small set of model parameters. Pole-zero approaches are particularly relevant because they approximate spectral characteristics using rational transfer functions whose parameters have a direct relationship with the frequency response structure (Blommer and Wakefield, 1997; Haneda et al., 1999).

Earlier research demonstrated that pole-zero approximations can effectively represent HRTFs while controlling approximation error. Blommer and Wakefield (1997) investigated pole-zero approximations using a logarithmic error criterion, establishing an important basis for perceptually and spectrally meaningful parametric representation. Haneda et al. (1999) subsequently developed common-acoustical-pole and zero modeling, emphasizing shared structures across HRTFs. Chen and Hsieh (2000) further considered common acoustical poles and zeros for three-dimensional sound processing, illustrating the relevance of parameterized representations to practical spatial-audio systems.

The dimensionality problem, however, is not completely eliminated by parametric modeling. A large HRTF dataset may still require numerous pole, zero, gain, and location-dependent parameters. Statistical component analysis can therefore provide a second level of compression by identifying dominant patterns in the parameter space. The combination of parametric acoustic modeling and statistical dimensionality reduction is the central focus of this study.

1.2 Problem Statement

Conventional HRTF representations face a trade-off between fidelity and computational efficiency. Direct impulse-response or frequency-response representations preserve substantial information but require relatively large data structures. Finite-impulse-response representations can be computationally attractive, whereas infinite-impulse-response representations can provide compact descriptions of spectral behavior (Kulkarni and Colburn, 1995; Kulkarni and Colburn, 2004). Common-pole/zero modeling introduces additional compactness through shared acoustic structures, but a large number of model coefficients may remain when many spatial positions are considered.

The central research problem is therefore how to obtain a compact HRTF representation without losing the dominant spectral information required for spatial perception. This study addresses the problem by combining common-pole/zero modeling with statistical component analysis.

1.3 Research Relevance and Objectives

The research is relevant to real-time spatial-audio systems where memory, bandwidth, and computational resources constrain the number and complexity of HRTF representations that can be stored or transmitted. Parameterized HRTF rendering has already been investigated for mobile three-dimensional audio applications, demonstrating the practical motivation for efficient representations (Song et al., 2017).

The primary objective is to formulate a low-dimensional representation in which common acoustic structures capture the underlying spectral characteristics and statistical components capture the dominant variation among spatial HRTFs. The secondary objectives are to establish a reconstruction mechanism, analyze dimensionality reduction, and identify the principal trade-offs between compactness and representation accuracy.

1.4 Scope and Significance

The proposed framework focuses on parametric representation rather than the perceptual validation of a particular listener population. It is intended as a general methodological architecture applicable to HRTF databases represented in frequency or transfer-function form. Its significance arises from the integration of two complementary abstractions: physically motivated acoustic parameters and statistically dominant variations.

The broader methodological principle is consistent with contemporary data-driven engineering frameworks in which complex system descriptions are transformed into compact representations before downstream processing. Ramamurthy, Konduru, and Amanmadov (2026), for example, emphasize adaptive representation and reasoning in complex software-engineering systems; although their application domain differs from spatial acoustics, the conceptual principle of reducing complex structured information into a manageable representation supports the architectural rationale of the present framework (Ramamurthy et al., 2026).

LITERATURE REVIEW

HRTF representation has evolved from direct signal-based descriptions toward increasingly compact parametric models. Kulkarni and Colburn (1995) examined finite-impulse-response filter models, demonstrating that HRTFs can be represented using a finite number of filter coefficients. Such representations are attractive because convolution-based implementation is straightforward; however, the number of coefficients can become substantial when high spectral resolution or long impulse responses are required.

The use of infinite-impulse-response structures provides another representation strategy. Kulkarni and Colburn (2004) investigated IIR models of HRTFs, offering a framework in which poles and zeros can represent spectral characteristics more compactly than direct sample-based descriptions. The fundamental advantage is that a frequency response with identifiable resonant structures can potentially be represented with fewer parameters than an equivalent long impulse response.

Blommer and Wakefield (1997) specifically addressed pole-zero approximation using a logarithmic error criterion. Their work is theoretically important for the present framework because an error measure applied to the logarithmic magnitude response can emphasize relative spectral deviations rather than treating all linear-scale errors equally. Such an approach is appropriate when the objective is to preserve the structural characteristics of an HRTF spectrum.

Haneda et al. (1999) introduced common-acoustical-pole and zero modeling of HRTFs. The central concept is particularly relevant because common poles and zeros can represent structures shared among multiple HRTFs, while other parameters describe directional variation. This provides a natural basis for reducing the number of independent parameters in a spatial HRTF dataset.

Chen and Hsieh (2000) extended the common-pole/zero concept toward three-dimensional sound processing. Their work reinforces the practical significance of shared acoustic structures for spatial audio, where multiple directional responses must be represented efficiently. Liu and Hsieh (2001) similarly examined common-acoustic-poles/zeros approximation, further establishing the relevance of shared parametric structures to HRTF approximation.

The literature also indicates that HRTF representation is not restricted to rational acoustic models. Hu et al. (2019) investigated spherical-wavelet modeling, demonstrating an alternative mathematical basis for representing spatially varying HRTFs. Wavelet-based approaches can capture spatial structure efficiently, but they differ conceptually from common-pole/zero methods because their compression mechanism is based on basis-function localization rather than explicit acoustic poles and zeros.

Hu et al. (2014) situated three-dimensional audio technology within a broader system-level context, emphasizing the importance of efficient representations and processing mechanisms in practical 3D audio. Song et al. (2017) specifically considered parameterized HRTFs for real-time rendering on mobile devices, highlighting the relationship between compact representation and computational deployment.

Collectively, the literature establishes two important directions: parametric acoustic modeling and compact mathematical representation. Nevertheless, these directions can be combined more explicitly. Common-pole/zero models provide interpretable acoustic parameters, but their parameter vectors can remain correlated and redundant. Statistical component analysis can exploit this redundancy by identifying a smaller set of dominant parameter combinations.

This creates the principal research gap addressed here: rather than selecting between acoustic parameterization and statistical dimensionality reduction, the proposed framework uses them sequentially. The common-pole/zero stage provides an acoustically meaningful intermediate representation, while the statistical stage provides low-dimensional compression.

The architectural logic is also compatible with broader computational frameworks that transform complex structured information into compact representations before analysis or inference. This principle is relevant to the adaptive representation philosophy discussed by Ramamurthy et al. (2026), although the present study applies it specifically to acoustic signal characterization rather than software engineering (Ramamurthy et al., 2026).

METHODOLOGY

3.1 Proposed Framework

The proposed methodology consists of five principal stages:

HRTF dataset → spectral preprocessing → common-pole/zero estimation → parameter matrix construction → statistical component reduction → low-dimensional reconstruction

Let an HRTF measured at spatial position (s) be represented by

$$
H_s(z)=G_s
\frac{\prod_{m=1}^{M}(1-z_m^{(s)}z^{-1})}
{\prod_{n=1}^{N}(1-p_n^{(s)}z^{-1})},
$$

where (G_s) denotes gain, ($z_m^{(s)}$) represents zero locations, and ($p_n^{(s)}$) represents pole locations.

In a common-pole/zero architecture, a subset of the poles and zeros is shared across spatial positions:

$$
H_s(z)=G_s
\frac{\prod_{m=1}^{M}(1-\bar{z}_m z^{-1})}
{\prod_{n=1}^{N}(1-\bar{p}_n z^{-1})}
\cdot R_s(z),
$$

where ($\bar{p}_n$) and ($\bar{z}_m$) represent common acoustic structures and ($R_s(z)$) captures position-dependent variation. This formulation follows the conceptual direction of common-acoustical-pole/zero modeling developed for HRTFs (Haneda et al., 1999; Chen and Hsieh, 2000).

3.2 HRTF Preprocessing

Before parametric estimation, the HRTF responses should be converted into a consistent representation. For each spatial position, the left- and right-ear transfer functions are independently considered. Magnitude responses may be transformed into a logarithmic scale:

$$
L_s(f)=20\log_{10}|H_s(f)|.
$$

This transformation is consistent with the logarithmic-error perspective used in pole-zero approximation (Blommer and Wakefield, 1997). It makes relative spectral deviations easier to analyze and provides a suitable domain for evaluating reconstruction differences.

Preprocessing should maintain identical frequency sampling and spatial indexing. Inconsistent frequency grids would introduce artificial variance into the statistical stage and could cause the principal components to represent preprocessing artifacts rather than genuine HRTF characteristics.

3.3 Common-Pole/Zero Estimation

The first modeling stage estimates a rational representation for each HRTF while encouraging common acoustic structures across positions. Pole-zero estimation can be formulated as an optimization problem:

$$
\theta_s^*=
\arg\min_{\theta_s}
E(H_s,\hat{H}_s),
$$

where (θ_s) contains gain, pole, and zero parameters and (E) denotes the approximation error.

A logarithmic spectral criterion may be expressed as

$$
E_s=\sum_f
w_f
\left[
L_s(f)-\hat{L}_s(f)
\right]^2,
$$

where (w_f) represents optional frequency weighting. The objective is not simply to minimize numerical error but to preserve the dominant spectral structure. This is particularly relevant because HRTFs contain frequency-dependent peaks and notches that can be poorly represented by a model optimized solely for unweighted linear error.

The common-pole/zero strategy reduces the number of independently estimated parameters. Instead of fitting every HRTF using completely unrelated poles and zeros, the model searches for structures that explain multiple spatial responses. This shared structure is theoretically attractive because directional HRTF responses are not independent signals; they arise from variations of the same underlying acoustic system.

3.4 Parameter Matrix Construction

After estimation, each HRTF is converted into a parameter vector:

$$
x_s=
[
G_s,
\Re(p_1),\Im(p_1),\ldots,
\Re(z_1),\Im(z_1),\ldots
].
$$

The complete dataset forms a parameter matrix

$$
X=
[x_1,x_2,\ldots,x_S]^T.
$$

This matrix is the input to statistical component analysis. Parameter normalization is necessary because gain, pole radius, pole angle, zero radius, and zero angle may have different numerical scales.

3.5 Statistical Component Analysis

Principal component analysis is applied to the normalized parameter matrix. Let the centered matrix be (X_c). Its covariance matrix is

$$
C=\frac{1}{S-1}X_c^TX_c.
$$

Eigenvalue decomposition gives

$$
C=V\Lambda V^T,
$$

where (V) contains eigenvectors and (Λ) contains eigenvalues. The first (K) eigenvectors define the reduced representation:

$$
Y=X_cV_K.
$$

The original parameter vector can then be reconstructed as

$$
\hat{X}=YV_K^T+\mu,
$$

where (μ) is the original parameter mean.

The number of retained components is selected according to a cumulative explained-variance criterion:

$$
\rho_K=
\frac{\sum_{k=1}^{K}\lambda_k}
{\sum_{j=1}^{D}\lambda_j}.
$$

A higher (ρ_K) generally provides more accurate reconstruction but requires greater storage and computational cost. Thus, (K) becomes a controllable parameter defining the trade-off between dimensionality and fidelity.

3.6 Reconstruction and HRTF Rendering

The reconstructed parameter vector is converted back into poles, zeros, and gains. These parameters generate an estimated HRTF:

$$
\hat{H}_s(z)=F(\hat{x}_s),
$$

where (F) denotes the inverse parameter-to-filter transformation.

For a mobile spatial-audio application, only the low-dimensional component vector (Y_s) needs to be stored or transmitted. The common acoustic model and PCA basis can be shared globally, while each spatial position is represented by a small coefficient vector. This approach extends the practical motivation of parameterized HRTF rendering for real-time systems (Song et al., 2017).

3.7 Evaluation Criteria

The proposed approach should be evaluated using four complementary criteria. First, spectral reconstruction error determines how closely the reconstructed HRTF matches the original. Second, dimensionality ratio evaluates the reduction from the original parameter count to the retained component count. Third, computational complexity measures the cost of reconstructing filters and generating binaural signals. Fourth, structural stability assesses whether reconstructed poles and zeros remain suitable for filter implementation.

A useful conceptual compression ratio is

[

CR=\frac{D}{K},

]

where (D) is the original parameter dimensionality and (K) is the retained statistical dimensionality. A high (CR) is desirable only if spectral reconstruction remains acceptable.

The methodological principle is analogous to adaptive representation strategies in complex computational systems, where the objective is not merely to reduce data volume but to preserve information necessary for downstream functionality (Ramamurthy et al., 2026).

RESULTS

Because no measured HRTF dataset or experimental benchmark was supplied, the findings below represent analytical outcomes implied by the proposed framework rather than fabricated numerical results.

The first finding is that common-pole/zero modeling provides an appropriate intermediate representation for dimensionality reduction. Direct PCA on raw HRTF samples would treat every frequency bin as an independent dimension, whereas the parametric stage converts the response into a substantially more structured set of acoustic variables. The approach therefore performs dimensionality reduction twice: first through rational acoustic representation and second through statistical component extraction.

The second finding concerns parameter redundancy. Spatially adjacent HRTFs are expected to exhibit related spectral structures because they correspond to different observations of the same acoustic system. Common-pole/zero modeling explicitly captures shared characteristics, while PCA can subsequently identify correlated changes among the remaining parameters. This combination should produce a more compact representation than applying either mechanism independently.

Third, the proposed framework creates a controllable fidelity-complexity relationship. Increasing the number of poles and zeros improves the ability of the acoustic model to represent fine spectral structures but increases the dimensionality of the parameter matrix. Conversely, retaining fewer principal components reduces storage and computational requirements while increasing reconstruction error. The optimal configuration is therefore application-dependent rather than universally defined.

Fourth, the approach is consistent with the progression of existing HRTF modeling research. FIR models provide direct filter representations (Kulkarni and Colburn, 1995), whereas IIR models and pole-zero structures provide more compact spectral parameterization (Kulkarni and Colburn, 2004). Common-acoustic-pole/zero methods further exploit shared structures across HRTFs (Haneda et al., 1999; Liu and Hsieh, 2001). The proposed method adds a statistical compression layer to this established parametric foundation.

Finally, the methodology is particularly suitable for systems in which HRTFs must be transmitted or reconstructed dynamically. The parameterized rendering objective investigated by Song et al. (2017) demonstrates the practical relevance of compact HRTF descriptions for resource-constrained real-time systems. The proposed low-dimensional coefficient representation provides a natural mechanism for such deployment.

DISCUSSION

The main theoretical contribution of the proposed framework is the integration of acoustic interpretability and statistical compression. Pole-zero modeling provides parameters associated with spectral structures, whereas PCA identifies dominant correlations among those parameters. The resulting representation is therefore more structured than a purely statistical representation and more compact than a purely parametric one.

This distinction is important. PCA alone may achieve strong dimensionality reduction, but its components are mathematical combinations of input dimensions and may not possess direct acoustic interpretation. Conversely, a common-pole/zero model provides meaningful filter parameters but does not necessarily remove all redundancy. Combining the two approaches addresses these complementary weaknesses.

The framework also provides a practical mechanism for adaptive model selection. Applications requiring high-fidelity offline reproduction can retain a larger number of principal components. Interactive applications with strict memory or bandwidth limitations can use fewer components. This creates a scalable representation rather than a fixed model.

The proposed approach is also consistent with earlier work on common acoustic structures. Haneda et al. (1999), Chen and Hsieh (2000), and Liu and Hsieh (2001) demonstrate the relevance of common poles and zeros for reducing HRTF representation complexity. The present framework extends this idea by considering the residual parameter space as a statistical object. Instead of assuming that the common-pole/zero model alone provides sufficient compression, it explicitly measures and exploits residual parameter correlations.

Alternative representations such as spherical wavelets provide another perspective on spatial HRTF compression (Hu et al., 2019). Wavelet representations may be particularly effective when spatial smoothness is dominant, whereas the proposed model is advantageous when spectral structure and shared filter characteristics are important. These approaches should therefore be viewed as complementary rather than mutually exclusive.

Several limitations must be acknowledged. First, PCA is fundamentally a linear dimensionality-reduction method. If HRTF parameter variation is strongly nonlinear, a linear component basis may require more components than a nonlinear representation. Second, pole-zero estimation can be sensitive to initialization, model order, and numerical constraints. Third, reducing dimensionality according to explained variance does not automatically guarantee perceptual equivalence. A component explaining relatively little statistical variance may still correspond to an acoustically important spectral feature.

Another limitation concerns stability. Reconstructed pole parameters must remain inside the appropriate stability region for digital filter implementation. Statistical reconstruction does not inherently guarantee this property. Consequently, stability constraints should be incorporated during parameter normalization, component selection, or post-reconstruction validation.

The conceptual architecture also suggests an opportunity for adaptive modeling. Rather than assigning the same number of components to every spatial region, future implementations could allocate more parameters to directions exhibiting strong spectral variation and fewer parameters to regions with smoother responses. Such adaptive allocation would further reduce representation cost while preserving important HRTF structures. This adaptive principle parallels broader computational approaches in which model capacity is allocated according to structural complexity (Ramamurthy et al., 2026).

From a practical perspective, the framework could be integrated into mobile, virtual-reality, augmented-reality, and interactive 3D-audio systems. The global acoustic model and PCA basis could be stored once, while directional HRTFs would be encoded through compact coefficient vectors. This architecture could reduce memory consumption and transmission overhead while retaining the ability to reconstruct position-specific binaural filters.

CONCLUSION

This study presented a low-dimensional HRTF characterization framework combining common-pole/zero modeling with statistical component analysis. The approach addresses the representation challenge by first converting HRTFs into structured acoustic parameters and subsequently removing statistical redundancy through principal components.

The analysis indicates that the combination is theoretically advantageous because the two stages address different sources of complexity. Common poles and zeros capture shared acoustic structures, while PCA compresses correlated parameter variation. The resulting representation offers a tunable trade-off among fidelity, dimensionality, memory requirements, and computational cost.

The framework builds upon established HRTF modeling research involving FIR, IIR, pole-zero, common-acoustic-pole/zero, and parameterized representations. Its principal contribution is not the replacement of these models but their integration into a two-stage low-dimensional architecture.

Future research should validate the framework using standardized HRTF datasets and quantitative reconstruction experiments. Particular attention should be given to model-order selection, stability-constrained PCA reconstruction, perceptual evaluation, spatial interpolation, and comparison with alternative dimensionality-reduction techniques. Such investigations would establish whether the proposed representation can simultaneously achieve substantial compression and perceptually faithful spatial-audio reproduction.

REFERENCES

1. M. A. Blommer and G. H. Wakefield, “Pole-zero approximations for head-related transfer functions using a logarithmic error criterion,” IEEE Trans. Speech Audio Process., vol. 5, no. 3, pp. 278–287, 1997.

2. M. Chen and S. -F. Hsieh, "Common acoustical-poles/zeros modeling for 3D sound processing," in 2000 IEEE Int. Conf. Acous. Speech Signal Process. Proc. (Cat. No. 00CH37100), Istanbul, Turkey, IEEE, 2000, vol. 2, pp. II785–II788.
3. Y. Haneda, S. Makino, Y. Kaneda, and N. Kitawaki, "Common-acoustical-pole and zero modeling of head-related transfer functions," IEEE Trans. Speech Audio Process., vol. 7, no. 2, pp. 188–196, 1999.
4. R. Hu et al., "Review on three-dimension audio technology," (in Chinese), Shuju Caiji Yu Chuli/J. Data Acquis. Process., vol. 29, pp. 661–676, 2014.
5. S. Hu, J. Trevino, C. Salvador, S. Sakamoto, and Y. J. A. A. Suzuki, "Modeling head-related transfer functions with spherical wavelets," Appl. Acoust., vol. 146, no. 6, pp. 81–88, 2019.
6. A. Kulkarni and H. S. Colburn, "Efficient finite-impulse-response filter models of the head-related transfer function," J. Acoust. Soc. Am., vol. 97, no. 5_Supplement, pp. 3278, 1995.
7. A. Kulkarni and H. S. Colburn, "Infinite-impulse-response models of the head-related transfer function," J. Acoust. Soc. Am., vol. 115, no. 4, pp. 1714–1728, 2004.
8. C. Liu and S. -F. Hsieh, "Common-acoustic-poles/zeros approximation of head-related transfer functions," in 2001 IEEE Int. Conf. Acous. Speech Sig. Process. Proc. (Cat. No. 01CH37221), Salt Lake City, UT, USA, IEEE, 2001, vol. 5, pp. 3341–3344.
9. Y. Song, W. Tu, R. Hu, X. Wang, W. Chen and C. Yang, "An efficient method using the parameterized HRTFs for 3D audio real-time rendering on mobile devices," in Pac. Rim Conf. Multim., Harbin, China, Springer, 2017, pp. 663–673.
10. K. Ramamurthy, R. K. Konduru and N. Amanmadov, "EvoGraphCoder: An Evolutionary Graph-Reasoning Framework for Self-Adaptive Software Engineering," in IEEE Access, vol. 14, pp. 63063-63076, 2026, doi: 10.1109/ACCESS.2026.3686019.
11. Geo Philip, Paulson, Integrated Intelligent Building Energy Management: A Multi-LayerFramework for Renewable Energy, HVAC Optimization, and SmartElectrical Network Coordination. Available at SSRN: https://ssrn.com/abstract=6993209 or http://dx.doi.org/10.2139/ssrn.6993209