A Perceptual Approach to HRTF Personalisation: From Localisation to Spatial Literacy

Dr Iain McGregor
Edinburgh Napier University
[ORCID: 0000-0002-0230-9007]

Introduction

Efforts to personalise Head-Related Transfer Functions (HRTFs) have traditionally focused on anatomical accuracy, modelling the head, ears and torso to predict how sound should reach an individual's ears. While valuable, such approaches often require specialist equipment and yield variable results. They also struggle to accommodate changes in hearing or playback context and rarely provide users with opportunities to build or refine spatial perception.

This paper outlines an alternative, interaction-driven method: a perceptual approach to spatial calibration that learns from how listeners respond, rather than how they are built. Unlike anatomical or database-driven HRTF methods, this system adapts dynamically to how users perceive and respond to sound in space. It creates a calibration method that is accessible, reflexive and capable of evolving over time.

A Task-Based Calibration Method

This method builds a perceptual profile through interaction, rather than prediction.

A sound is rendered using a generic or best-fit HRTF and presented in space. The listener then indicates where they believe the sound originated, either by

- turning their head toward the perceived direction using head-tracked headphones
- or moving a handheld device such as a smartphone in three-dimensional space to point to or place it at the perceived location, capturing both azimuth and depth.

The sound stops immediately once movement begins, avoiding dynamic cues that might alter perception. The system records the response and refines future renderings accordingly.

For example, a listener might hear a chirping bird behind them. They rotate to face the perceived source. The system logs this input and adjusts the spatial model. Over time, this builds a perceptual rendering tuned to the individual listener.

Cue-Specific Probing

Different auditory cues shape our sense of space. These can be specifically tested and weighted.

The stimuli presented can be tailored to emphasise different spatial features, such as,

- interaural time differences (ITD) at low frequencies
- interaural level differences (ILD) at high frequencies
- spectral shaping from pinnae
- onset timing and reverberation trails.

The system learns how each listener interprets these cues and adjusts emphasis or compensation accordingly. This creates a personalised rendering profile rooted in perception, not assumption.

Hearing and Playback Variability

Instead of isolating hearing thresholds or measuring playback conditions, the system integrates them into the calibration process.

Since all judgements are made perceptually, the system automatically accommodates,

- partial hearing loss or asymmetry
- playback system irregularities
- masking effects from ambient sound
- and device-specific frequency roll-offs.

Temporary perceptual shifts, such as during a cold, can also be accounted for by prompting recalibration or monitoring passive indicators.

Passive Indicators of Perceptual Change

The system can respond to perceptual changes without direct user input.

Shifts in the user’s own voice, as captured through a microphone, or changes in how they use volume controls may signal a perceptual threshold change. These signals could trigger an optional recalibration or suggest switching to an alternate profile that better matches current conditions.

Safe Listening and Targeted Adaptation

Rather than encouraging higher overall volume, the system enhances only what is necessary.

When a listener struggles with clarity, most systems simply amplify everything. This approach is more selective, boosting only the bands or cues needed for spatial awareness. As a result, users often reduce total volume rather than increasing it. This supports safe listening practices, particularly in prolonged use or among those with hearing sensitivities.

Designing for Engagement and Inclusion

Spatial awareness is developmental and emotional. Engagement improves learning.

Instead of abstract test tones, sounds can be selected to reflect a listener's interests or environment. For example,

- a dinosaur roar or puppy bark for a child
- the voice of a familiar person
- or sounds from a preferred show, toy or book.

This makes calibration intuitive and emotionally relevant, especially for children and neurodiverse users. The model begins with simple azimuth placement and gradually introduces elevation, distance and complexity.

Perceptual inconsistency is expected and accommodated. The goal is not to achieve laboratory precision but to support ongoing adaptation.

Training Spatial Acuity

The model can enhance perception through gradual exaggeration of spatial cues.

By increasing cue contrast during training, such as exaggerated reverberation or widened stereo separation, the system teaches users to identify fine distinctions. Once sensitivity improves, exaggeration is reduced. The result is lasting perceptual acuity, useful in both everyday listening and safety-critical environments.

Extending to Real-World Use

The perceptual model is formed in controlled conditions but can be applied in situ.

This method is not designed to simulate every acoustic environment. Instead, it builds a core perceptual rendering profile that users can carry into different contexts. Additional layers, such as environmental modelling or real-world testing, can be introduced later depending on user goals.

Complementing Anatomical Models

This method does not replace anatomical estimation. It improves it.

Existing HRTF databases or scans can serve as a starting point. Where available, audiograms can also be used to initialise or constrain rendering, especially in known cases of frequency-specific hearing loss. If audiometric data is unavailable, the system could optionally draw on standardised presbycusis profiles (such as ISO 7029) when a user’s age and gender are known. The perceptual method then tunes the rendering based on how each individual actually hears, refining the model iteratively through interaction. It is compatible with existing workflows and can be layered atop anatomical or audiological approaches.

Equalising Spatial Resolution

Perceptual clarity is often strongest in the horizontal plane. This method aims for balance.

By deliberately incorporating vertical and distance cues, through both stimulus design and perceptual mapping, the system encourages more equal resolution across axes. Combined with ambisonics or object-based rendering, this approach promotes full three-dimensional spatial literacy.

Long-Term Adaptation and Model Evolution

A spatial model should evolve as the listener does.

The system supports user-initiated recalibration but can also adapt passively over time. Changes in hearing, context or equipment can trigger subtle updates. All processing is performed on-device with no cloud dependency. If desired, anonymised data can be shared to support broader model improvements, just as crash logs or user metrics are sometimes shared voluntarily in software.

Spatial Awareness Is Fundamentally Human

The system is designed for perceptual inclusion, not just for technical users.

Everyone engages with spatial information daily, in darkness, in motion and across different sensory modes. For many, this involves hearing. For others, it may involve touch, proprioception or visual cues. While this system focuses on auditory spatial calibration, it is grounded in a broader respect for perceptual diversity and adaptability.

It does not assume ability, literacy, language or anatomy. It meets people where they are and grows with them.

Future Directions

The system could be extended into several domains
• adaptive spatial rendering in AR, VR or gaming
• non-invasive audiological profiling and training
• personalisation of assistive devices or hearing aids
• accessible design in public and educational sound systems
• or tools for musical, spatial or auditory learning.

Because the model is perceptual, ethical and user-owned, it can be widely applied without licensing barriers or proprietary restrictions.

Licensing Note

Dr Iain McGregor, Edinburgh Napier University, 2025
ORCID: 0000-0002-0230-9007

This release supports open research, inclusive design and responsible development. Please cite this work if used in academic, technical or creative outputs.

Closing Note

Spatial audio need not aim for anatomical perfection. It can aim instead for perceptual meaning, inclusive clarity and interactive learning. Systems that adapt to the listener, and help the listener adapt in return, may ultimately offer more realism. Not by predicting how we hear, but by listening to what we hear and responding accordingly.

That path remains open.