



Examining the interrelation behavior of distance metrics for head-related transfer function evaluation: a case study

Shaima'a Doma* , Natálie Brožová , and Janina Fels 

Institute for Hearing Technology and Acoustics, RWTH Aachen University, Kopernikusstrasse 5, 52074 Aachen, Germany

Received 19 August 2022, Accepted 6 June 2023

Abstract – The comparison of head-related transfer functions (HRTFs), e.g., for validating different acquisition methods, requires a meaningful way of quantifying HRTF differences – a problem to which literature offers no standardized approach. This impedes the comparability and interpretability of studies. The present work therefore addresses the lack of proper understanding of the behavior of commonly used distance metrics by applying seven metrics to individually measured and approximated HRTF datasets. Covering a variety of spectral deviations, we perform both intra-individual comparisons (contrasting different levels of spectral detail for the same individual) and inter-individual comparisons (assessing metric reactions to non-individual cue differences). The metrics exhibit selective reactions to distinct spectral alterations. Particularly, the results demonstrate that spectrally localized errors go undetected by five out of seven metrics. Further analysis emphasizes inconsistencies in metric correlation patterns. These observations highlight the need for a multi-dimensional metric, capturing various types of HRTF differences for a proper assessment of errors.

Keywords: Binaural technology, Similarity criterion, Virtual acoustics

1 Introduction

In model-based binaural synthesis, head-related transfer functions (HRTFs) play a crucial role in creating virtual acoustic scenes. These multi-dimensional datasets usually cover a number of sound incidence directions, which allows for synthesizing virtual sound sources at almost arbitrary positions around the listener. For acquiring individualized HRTF sets, a variety of methods can be applied, such as acoustical measurements (e.g., [1–5]), sound field simulations (e.g., [6–10]), spectral reconstruction (e.g., [11, 12]), or adjustment of the magnitude [13] and/or phase [14] of a given generic HRTF set.

For validation, it is necessary to assess the subjective quality of these acquired HRTF. However, as listening experiments are time consuming and furthermore underlie uncertainties in the participants' responses, it is often needed to numerically quantify the differences between transfer functions resulting from the different approaches, as well as the variability within each approach. For example, approximation methods require a similarity criterion for a direct cross-comparison of the resulting spectra with their respective target HRTFs. On observing studies of this kind, the lack of a standardized way of capturing HRTF differences becomes apparent. The studies use a variety of distance

metrics, which, depending on their properties, detect different aspects of spectral dissimilarity. For example, a varying response of different metrics to an angular mismatch between HRTF spectra was examined in [15]. Still, as no single metric is sufficient to predict and describe audible differences, a higher-level metric is striven for that should combine the benefits of different metrics. The main focus of the present work lies in understanding the interrelation between a selected set of metrics. A case study is performed for this purpose, evaluating an HRTF individualization method (principal component reconstruction) and thereby offering observations for different kinds of spectral deviations.

In this context, four questions are of interest: a) how the approximation errors introduced by the individualization method are reflected in distance metric behavior; and b) how these approximation errors scale relative to the error found between generic HRTFs. A comparison of metric behavior in different cases further gives insight into the redundancy between metrics, hence, Pearson correlation coefficients are examined to address the following questions: c) What are the common trends in metric behavior, and are they case-dependent?; and d) can metric behavior be set in relation with anthropometric deviations? Conclusions are drawn regarding the future definition of a generalized overarching metric.

*Corresponding author: shaimaa.doma@akustik.rwth-aachen.de

2 Materials

The following sections introduce the distance metrics considered in the present work as well as the HRTF data used for analysis.

2.1 Distance metric overview

Seven distance metrics with varying properties were selected from or based on literature, covering different categories, as visualized in Figure 1. All of the metrics are of monaural nature, contrasting the transfer function of one ear and a single incidence direction, respectively, with the corresponding transfer function in another HRTF set. Unless otherwise stated, the metrics consider the full frequency range up to 22.050 Hz.

The classic frequency bin-based calculation of the Mean-Squared Error (MSE) metric

$$\text{MSE} = \frac{1}{n\text{Bins}} \cdot \sum_{i=1}^{n\text{Bins}} [|\text{HRTF}_1(i)| - |\text{HRTF}_2(i)|]^2, \quad (1)$$

with $n\text{Bins}$ equal to 129, is extended by a perceptual weighting, introducing the Critical-Band Mean Squared Error (CB) [15]

$$\text{MSE}_{\text{CB}} = \frac{1}{n\text{Bins}} \cdot \sum_{i=1}^{n\text{Bins}} (\alpha_{\text{CB}}(i) \cdot [|\text{HRTF}_1(i)| - |\text{HRTF}_2(i)|])^2, \quad (2)$$

with factors $\alpha_{\text{CB}}(i) = \frac{\alpha_0}{\Delta f_{\text{CB}}(i)}$, $\Delta f_{\text{CB}}(i)$ being the critical band width at the respective frequency, and α_0 the value that allows the sum of $\alpha_{\text{CB}}(i)$ to be equal to one.

The Mel Frequency Cepstral Distortion [16, 17] performs a similar calculation on Mel Frequency Cepstral Coefficients (MFCCs) [18], defined on $n\text{Coeff}$ (here: 24) bands as

$$\text{MSE}_{\text{MFCD}} = \frac{1}{n\text{Coeff}} \cdot \sum_{i=1}^{n\text{Coeff}} [\text{MFCC}_1(i) - \text{MFCC}_2(i)]^2. \quad (3)$$

Metrics based on mean squares capture overall gain differences between spectra; in contrast, a variance calculation emphasizes changes in spectral shape (SS), regardless of a potential gain offset. The latter category is represented here by the Inter-Subject Spectral Difference (ISSD) [13]

$$\text{SS}_{\text{ISSD}} = \frac{20^2}{n\text{Bins}} \cdot \sum_{i=1}^{n\text{Bins}} \left(\lg \left[\frac{\text{DTF}(i)_1}{\text{DTF}(i)_2} \right] - \frac{1}{n\text{Bins}} \cdot \sum_{j=1}^{n\text{Bins}} \lg \left[\frac{\text{DTF}(j)_1}{\text{DTF}(j)_2} \right] \right)^2, \quad (4)$$

with the variance calculated on bin-level on Directional Transfer Functions (DTFs, HRTFs normalized by the diffuse field transfer function) for frequencies up to 13 kHz,

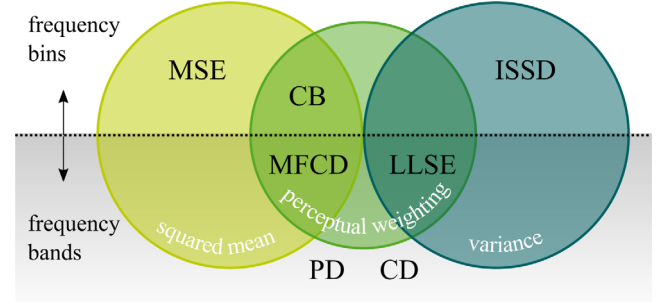


Figure 1. Classification of the selected HRTF distance metrics. Based on metric properties, e.g. the frequency resolution, the main arithmetic operation performed, and the consideration of perceptual properties of the auditory system (or lack thereof), different behavior of the metrics can be presumed.

and by the Loudness Level Spectral Error (LLSE, based on [19])

$$\text{SS}_{\text{LLSE}} = \frac{1}{n\text{ERB}} \cdot \sum_{i=1}^{n\text{ERB}} \left(\Delta L_L(i) - \frac{1}{n\text{ERB}} \cdot \sum_{j=1}^{n\text{ERB}} \Delta L_L(j) \right)^2, \quad (5)$$

with loudness level difference ΔL_L calculated in phon for $n\text{ERB}$ (here: 42) equivalent rectangular bandwidth (ERB) channels, respectively. A direct comparison of broadband spectral energy is provided by the Power Difference (PD)

$$\text{PD} = 10 \cdot \lg \left(\frac{\sum_{i=1}^{n\text{Bins}} \text{HRTF}_1(i)^2}{\sum_{j=1}^{n\text{Bins}} \text{HRTF}_2(j)^2} \right), \quad (6)$$

thus capturing directional level differences, which would be present even if the user chose to normalize the complete HRTF sets, e.g., based on the level difference for frontal incidence. Finally, the Correlation Distance (CD) metric [20] considers phase in addition to magnitude information, as included in the head-related impulse response (HRIR), yielding

$$\text{CD} = \frac{1}{n\text{GT}} \cdot \sum_{i=1}^{n\text{GT}} [1 - \max(\text{corr}(\text{HRIR}_1 * \text{BP}_i, \text{HRIR}_2 * \text{BP}_i))], \quad (7)$$

with the Gammatone (GT) bandpass filters (BP) corresponding to $n\text{GT}$ (here: 42) bands.

Merely based on metric properties, certain observations can be expected. In contrast to envelope-based metrics, bin-based calculations should better detect changes in fine spectral structure. By considering the logarithmic frequency characteristic of the inner ear, either through weighting (MSE_{CB}) or through the choice of bandwidths in a filter bank (SS_{LLSE} , MSE_{MFCD} and CD), an overrating of errors

in the high frequency range is avoided, better accounting for the human auditory perception.

2.2 HRTF reconstruction

As a basis for the present analysis, we used 20 measured HRTF sets from the ITA HRTF database [21]. Only left-ear data were evaluated, which were furthermore limited to elevations between $\pm 45^\circ$, thus avoiding over-representation of incidence directions close to the poles, given the $5^\circ \times 5^\circ$ equiangular sampling grid. This resulted in a total of 1296 incidence directions. The transfer functions were additionally approximated using Principal Component Analysis (PCA) reconstruction, as described in [22]. This individualization method comprises two steps: First, the magnitude spectra are expressed as a weighted sum of PCs (here: 23), the ideal weighting score being a direct output of the PCA. Thus reconstructed data are in the following referred to as the “idealPCA” dataset. The second step is an approximation of these ideal weights, using a weighted sum of anthropometric features. Six head and ear dimensions with lowest bi-variate correlation [23] are chosen for this regression model, leading to the following features: {“h”, “d2”, “d3”, “d6”, “d7”, “d8”}, as defined in the HRTF database. The resulting dataset is in the following referred to as “anthroPCA”. Since PCA is here used on magnitude data, reconstructed spectra initially possess no phase information. A minimum phase [24] is added, yielding the correct shape of the HRIR, with the impulse energy properly distributed over time.

3 Results and discussion

In the following, inter- and intra-individual errors are discussed for different datasets, followed by a global comparison of correlation behavior and a discussion of consequences for metric design.

3.1 Inter-individual error

Listening with non-individual spatial cues is relevant whenever only a generic HRTF set is available. This dataset may belong to a dummy head or another person, with a directivity more or less suitable for the listener. Since inter-individual differences between approximated HRTFs are not of practical relevance, the following evaluation is limited to the 20 “measured” HRTFs from the ITA database. Three representative HRTF sets are selected, to be each compared to the remaining 19 sets, respectively. The selection is based on the relative deviation Δa of the individual’s anthropometric features from the mean of the 20 individuals. This deviation is defined for individual i as

$$\Delta a_i = \frac{1}{n\text{Feat}} \cdot \sum_{j=1}^{n\text{Feat}} \frac{|a_{j,i} - \mu_j|}{\mu_j}, \quad (8)$$

with $n\text{Feat}$ (here: 11) features (a_j) from the ITA database {“w”, “dF”, “h”, “d1” – “d8”}, and μ_j being the mean over

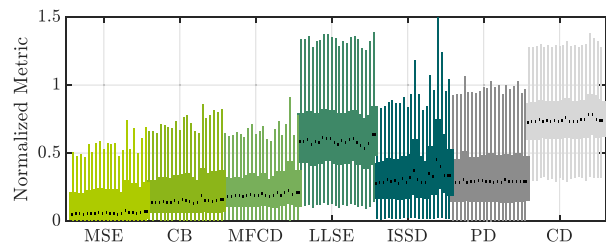


Figure 2. Inter-individual HRTF errors for different reference individuals i , sorted by the deviation of their anthropometric features from the database mean, as expressed by Δa_i (Eq. (8)). The black dots indicate the respective median values.

feature j of the 20 individuals. A closer look at the relation between metric values and Δa , sorting the deviation Δa_i by ascending order, see Figure 2, reveals that not for all metrics, the value range scales with Δa . Note that the distribution of metric values *after* normalization by the global (i.e., including the inter-individual subsets below and intra-individual data from Sect. 3.2) 95th percentile of the respective metrics is shown. The box plots represent data within the respective 25th and 75th percentiles. Besides the rather weak tendency for an increase in median values (black dots), the box plots for each individual metric, respectively, show a substantial overlap.

Nevertheless, representative datasets with the minimum (ID 13), median, i.e., 11th rank (ID 8) and maximum (ID 1) anthropometric deviations are selected for further analysis in Figure 3a. At the top, each of the three selected HRTF sets is contrasted with the other 19 individuals. The black and gray lines correspond to pairs of contra-lateral horizontal plane magnitude spectra at -20° azimuth for ID 13 in direct comparison to IDs 8 and 1, shown in the middle row. Considering the example spectra, the much larger spectral deviations between IDs 13 vs. 8, compared to IDs 13 vs. 1 are striking. The disparity is directly reflected in the corresponding metric values (dark gray line). This rather unexpected behavior can be explained by the fact that the selection of these datasets is based on anthropometric similarity to the *mean* and does not guarantee higher similarity to individual datasets.

Generally, measured HRTFs possess a high resolution of spectral cues, with sharp direction-dependent peaks and notches. Therefore, a mismatch between spectra with such properties seems to affect most metrics, capturing a variety of errors.

3.2 Intra-individual error

Similarly, intra-individual errors are visualized in Figures 3b–3d. Here, data points for 20 subjects \times 1296 directions are available for each of the metrics. The example HRTF spectra are again extracted at -20° , yet from IDs 13 and 1 in the different datasets, respectively. Each of the two approximation steps comprised in PCA individualization introduces a different kind of error to the spectra, with both errors superimposing in the “anthroPCA” dataset, i.e., the final output of the individualization method.

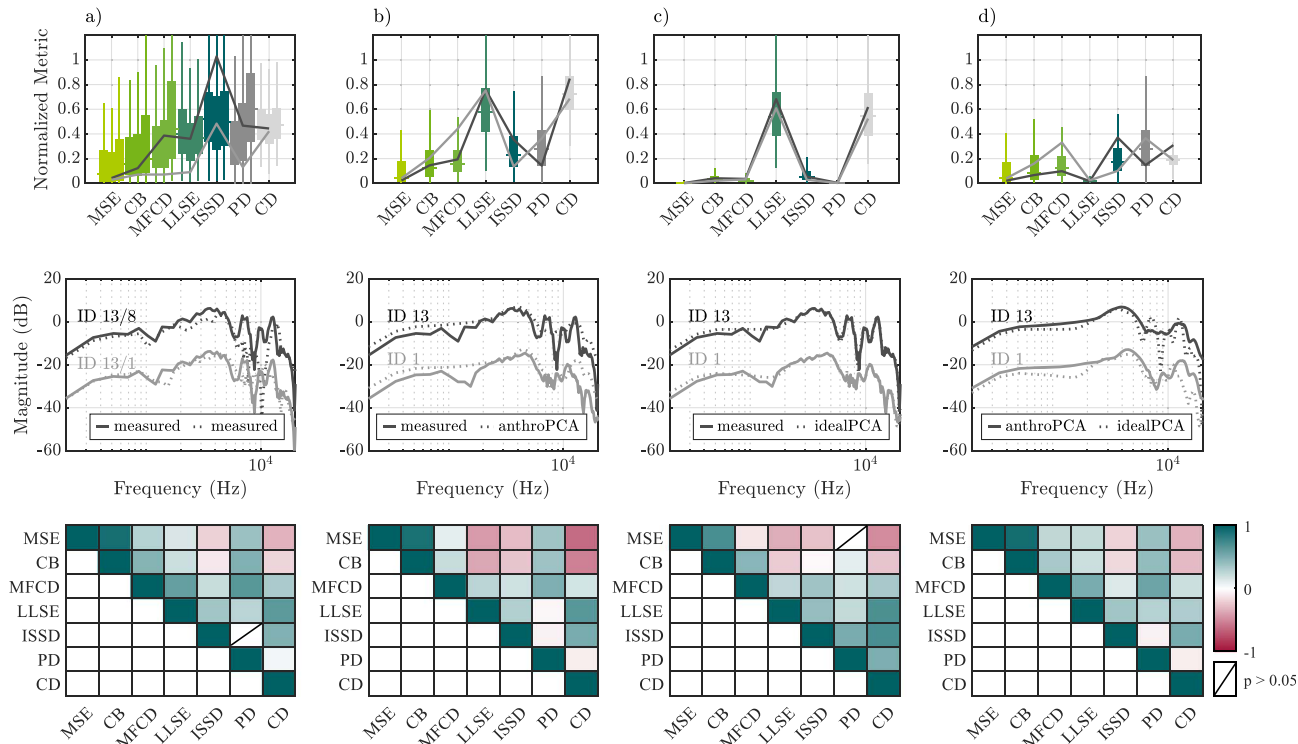


Figure 3. Metric behavior for (a) inter-individual errors within the “measured” dataset (with IDs 13, 8, 1 as reference, from left to right) and for intra-individual errors between (b) the “measured” and “anthroPCA”, (c) the “measured” and “idealPCA”, and (d) the “idealPCA” and “anthroPCA” datasets. The box plots (top) visualize the distribution of metric values, normalized by the 95th percentile of each metric. Two example pairs of transfer functions are depicted in the middle row, with an offset of 20 dB introduced for better visibility. Corresponding (normalized) metric values are indicated by the gray and black lines superimposing the box plots. Pearson correlation coefficients with $p < 0.05$ after Bonferroni correction (bottom) display varying patterns, with most prominent differences present in case (c).

Column (c) displays the error caused by the first approximation of measured spectra by PCs. As can be observed in the middle row, the compared spectra are almost identical, except for smoothing of prominent notches. Such deviations, inherently localized in frequency, seem to be well captured by variance-based metrics, particularly SS_{LLSE} . Squared mean-based metrics – due to their averaging nature – are dominated by the spectral similarity in the remaining frequency regions, thus overseeing the localized, though prominent and very likely audible [25], error. Another strong reaction to the observed smoothing effect is present in the CD metric. It can be inferred that the localized spectral error and ensuing changes in minimum phase response significantly modify the shape of HRIRs, affecting the temporal cross-correlation.

Column (d) shows the error introduced by the regression model, which approximates the ideal PC scores based on anthropometric features of the target listener. The behavior is complementary to the previously discussed case (c); the CD metric consistently shows a much smaller error and SS_{LLSE} almost none, while the remaining metrics exhibit a rise. As visible in the example spectra, this approximation step entails generally flatter magnitude spectra with possible disparities in peak and notch frequencies, compared to the spectra reconstructed with ideal scores. Rather than only dampening individual cue resonances, “non-individual”

components are introduced to the spectra, prompting a reaction of the metrics that previously also detected inter-individual errors in column (a). Nonetheless, since the spectral cues are not very prominent, the metrics react less strongly.

Column (b) evaluates the overall performance of the individualization method in its typical use. Superposition of the two approximation errors (columns (c) and (d)) can be well observed in the resulting distributions. Most values lie below the inter-individual case (a), where measured HRTFs are contrasted. However, two overshooting metrics (SS_{LLSE} and CD) put forward the question of perceptual relevance of the different metrics. Solely based on the five other distance metrics, individualized HRTFs would be preferred, rather than a generic HRTF set. However, the perceptual effect of the different types of errors should still be taken into account for a correct judgment.

3.3 Correlation analysis

Pearson correlation coefficients ($p < 0.05$ after Bonferroni correction) were calculated to assess linear relationships between the metrics. Due to outliers towards high metric values, a logarithmic transform was applied beforehand to approximate a normal distribution. As shown in the bottom row of Figure 3, correlation patterns show some variations between use cases. Common trends lie in moderate to strong

Table 1. Global 50th and 95th percentiles of the metric values of combined inter- and intra-individual HRTFs comparisons. The 95th percentiles are used for normalization in Figures 2 and 3.

Unit	MSE	CB	MFCD	LLSE	ISSD	PD	CD
	–	–	dB ²	phon ²	dB ²	dB	–
50%	0.12	8.7e-6	1.99	12.05	13.72	1.83	0.09
95%	3.57	8.3e-5	12.25	32.29	49.37	6.74	0.20

correlation of MSE with MSE_{CB} , and moderate correlation between CD and variance-based metrics, especially SS_{LLSE} . The latter goes in line with the joint reaction of the two metrics in the box plots of columns (b) and (c).

Largest deviations in correlation behavior are visible in column (c), where the “measured” and “idealPCA” datasets are contrasted. An otherwise moderate correlation between the power difference PD and metrics MSE_{CB} and MSE_{MFCD} is not present in this case, since level differences are barely found between the spectra. Such a deviation between correlation patterns emphasizes the effect of different types of spectral properties on metric behavior and interrelation.

A weak to moderate negative correlation between CD and metrics MSE and MSE_{CB} , especially in column (b), could relate to the level dependency of the linear MSE calculation. Their response to ipsi-lateral errors is boosted, whereas CD tends to be more sensitive to contra-lateral errors, where the first (main) impulse of the HRIR, consisting of a summation of multiple paths around the head, is more temporally spread and of lower level. Later HRIR components, more affected by differences in pinna resonances, therefore tend to be more dominant in CD calculation for contra-lateral incidence.

Another noteworthy observation is the comparatively low correlation within each of the two groups (MSE and SS), respectively, with the exception of the regular MSE and MSE_{CB} . As the correlations referred to are mostly “only” moderate, the observations are not indicative for a lack of additional information provided by any of the metrics within the groups.

3.4 Implications for a generalized metric

A holistic approach to capture HRTF differences should not be restricted to specific cases but should allow for an assessment of arbitrary spectral deviations. This includes, besides the four covered cases, also the comparison of measurements performed in different setups, the evaluation of errors ensuing from spatial interpolation or other HRTF processing methods. In the intended use of the metric, a-priori knowledge would therefore not be given; instead, the interplay of metrics should suggest the presence of one (or more) type(s) of spectral deviations. To enable such a description, the generalized metric should, at least to some degree, retain its multi-dimensional nature.

While much redundancy between metrics can be observed, its case-dependency advises against the elimination of a subset of metrics, as no metric can provide all of the information obtained by another. Instead, an approach

should be chosen that exploits the specific variance of each of the metrics, e.g., based on PCA.

From the observations, it can be assumed that some metrics may react to specific spectral deviations. A classification of distance metrics into subgroups may describe this tendency, however, the reduction of the groups to one “representative” metric each may be counterproductive from a numerical perspective, as much of the variance would be lost. Nonetheless, a distinction should be made between the spectral variance explained by the metrics and the perceptual variance in audibility upon using the examined transfer functions for auralization. The findings do not rule out that fewer metrics may be sufficient for modeling auditory detectability. Nonetheless, it should be emphasized that it may be difficult to draw a direct connection between the spectral descriptors and auditory percepts. The subjectively perceived change is not only dependent on the two spectra being contrasted. Instead, the similarity (or lack thereof) between a given spectrum and an individual’s “learned” spatial cues has a direct effect on their attributing the spectral change to a perceptual property, such as localization or sound coloration. Therefore, only a spatially inclusive analysis, applying knowledge on the listener’s own HRTF set, would allow for generalization.

4 Conclusion

The present work examined the behavior of and correlation between seven distance metrics that are applicable to HRTF comparison. As a case study, metric responses to differences between generic HRTFs as well as to errors introduced by HRTF reconstruction were assessed. It could be observed that metric reactions and interrelation are very dependent on the underlying spectral deviations. While non-individual cues in generic HRTF comparison lead to the largest values, two metrics (SS_{LLSE} and CD) showed the strongest reactions to a loss in peak and notch quality – a spectral error missed by most of the other metrics, especially when the latter were based on squared mean calculations. Although some trends in moderate to strong correlation were consistent in the four examined cases, the correlation patterns showed a rather strong case dependency. The results indicate that a generalized metric for describing HRTF differences with no a-priori knowledge on underlying deviations should be modeled based on information from many (if not all) metrics, rather than reducing the latter to a representative subset. Future work could extend the study by additional metrics (including perceptual auditory models) and multi-variate approaches, such as Factor Analysis. A direct relation of the given metrics to audibility and to more specific anthropometric dissimilarities between individuals could furthermore be assessed.

Data availability statement

The presented distance metric data and the measured and reconstructed HRTFs in use can be downloaded from <https://doi.org/10.18154/RWTH-2023-05668>.

Conflict of interest

Part of this work has been previously presented at DAGA22, Stuttgart, Germany. The authors take complete responsibility for the integrity of the data and the accuracy of the analysis.

Funding

This work was funded by the German Research Foundation (DFG) – Project no. 402811912.

References

1. H. Møller, M.F. Sørensen, D. Hammershøi, C.B. Jensen: Head-related transfer functions of human subjects. *Journal of the Audio Engineering Society* 43, 5 (1995) 300–321.
2. V. Pulkki, M.-V. Laitinen, V. Sivonen: HRTF measurements with a continuously moving loud speaker and swept sines, in 128th Audio Engineering Society Convention, London, UK, May 22–25, 2010.
3. J.-G. Richter, G. Behler, J. Fels: Evaluation of a fast HRTF measurement system, in 140th Audio Engineering Society Convention, Paris, France, June 4–7, 2016.
4. J. He, R. Ranjan, W.-S. Gan: Fast continuous HRTF acquisition with unconstrained movements of human subjects, in Sound 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2016, pp. 321–325.
5. J. Reijniers, B. Partoens, J. Steckel, H. Peremans: HRTF measurement by means of unsupervised head movements with respect to a single fixed speaker. *IEEE Access* 8 (2020) 92287–92300.
6. B.F. Katz: Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation. *Journal of the Acoustical Society of America* 110, 5 (2001) 2440–2448.
7. M. Otani, S. Ise: Fast calculation system specialized for head-related transfer function based on boundary element method. *Journal of the Acoustical Society of America* 119, 5 (2006) 2589–2598.
8. H. Ziegelwanger, P. Majdak, W. Kreuzer: Numerical calculation of listener-specific head-related transfer functions and sound localization: Microphone model and mesh discretization. *Journal of the Acoustical Society of America* 138, 1 (2015) 208–222.
9. T. Xiao, Q. Huo Liu: Finite difference computation of head-related transfer function for human hearing. *Journal of the Acoustical Society of America* 113, 5 (2003) 2434–2441.
10. H. Takemoto, P. Mokhtari, H. Kato, R. Nishimura, K. Iida: Mechanism for generating peaks and notches of head-related transfer functions in the median plane. *Journal of the Acoustical Society of America* 132, 6 (2012) 3832–3841.
11. D.J. Kistler, F.L. Wightman: A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction, *Journal of the Acoustical Society of America* 91, 3 (1992) 1637–1647.
12. B.-S. Xie: Recovery of individual head-related transfer functions from a small set of measurements. *Journal of the Acoustical Society of America* 132, 1 (2012) 282–294.
13. J.C. Middlebrooks: Individual differences in external-ear transfer functions reduced by scaling in frequency. *Journal of the Acoustical Society of America* 106, 3 (1999) 1480–1492.
14. R. Algazi, C. Avendano, R.O. Duda: Estimation of a spherical-head model from anthropometry. *Journal of the Audio Engineering Society* 49, 6 (2001) 472–479.
15. R. Nicol, V. Lemaire, A. Bondu, S. Busson; Looking for a relevant similarity criterion for HRTF clustering: A comparative study, in 120th Audio Engineering Society Convention, Paris, France, May 20–23, 2006.
16. S. Shimada, N. Hayashi, S. Hayashi: A clustering method for sound localization transfer functions. *Journal of the Audio Engineering Society* 42, 7/8 (1994) 577–584.
17. K.-S. Lee, S.-P. Lee: A relevant distance criterion for interpolation of head-related transfer functions. *IEEE Transactions on Audio, Speech, and Language Processing* 19 (2011) 1780–1790.
18. S. Davis, P. Mermelstein: Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 28, 4 (1980) 357–366.
19. J. Huopaniemi, N. Zacharov, M. Karjalainen: Objective and subjective evaluation of head-related transfer function filter design. *Journal of the Audio Engineering Society* 47, 4 (1999) 218–239.
20. B. Xie, C. Zhang, X. Zhong: A cluster and subjective selection-based hrtf customization scheme for improving binaural reproduction of 5.1 channel surround sound, in 134th Audio Engineering Society Convention, Rome, Italy, May 4–7, 2013.
21. R. Bomhardt, M. de la Fuente Klein, J. Fels: A high-resolution head-related transfer function and three-dimensional ear model database. *Proceedings of Meetings on Acoustics* 29, 1 (2016) 050002.
22. S. Hwang, Y. Park: Interpretations on principal components analysis of head-related impulse responses in the median plane. *Journal of the Acoustical Society of America* 123, 4 (2008) EL65–EL71.
23. R. Bomhardt, J. Fels: Individualization of head-related transfer functions by the principle component analysis based on anthropometric measurements. *Proceedings of Meetings on Acoustics* 140, 4 (2016) 3277–3277.
24. A. Kulkarni, S.K. Isabelle, H.S. Colburn: On the minimum-phase approximation of head-related transfer functions, in *Proceedings of 1995 Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, USA, October 15–18, 1995, pp. 84–87.
25. M. Kohmen, R. Bomhardt, J. Fels, M. Vorländer: Just noticeable notch smoothing of head-related transfer functions, in *Fortschritte der Akustik – DAGA 2018*, Munich, Germany, March 19–22, 2018.

Cite this article as: Doma S. Brožová N. & Fels J. 2023. Examining the interrelation behavior of distance metrics for head-related transfer function evaluation: a case study. *Acta Acustica*, 7, 34.