Audio Engineering Society

Convention Paper 10674

Presented at the 155th Convention
2023 October 25–27, New York, USA

LoCOMo: A Low-Cost Open-Source Head Motorization Kit

Nils Poschadel, Stephan Preihs, and Jürgen Peissig

Institute of Communications Technology, Leibniz University Hannover, Germany

Correspondence should be addressed to Nils Poschadel (poschadel@ikt.uni-hannover.de)

ABSTRACT

When analyzing, simulating or auralizing human spatial hearing, it is increasingly demanded to consider not only different directions of sound incidence but also the respective head-above-torso orientation (HATO). However, incorporating this additional degree of freedom presents new challenges, e. g., when generating head-related transfer function (HRTF) datasets with adequate resolution since this requires automation/motorization to vary HATO during measurements with a head and torso simulator (HATS). In this paper, the release of the LoCOMo kit, a low-cost open-source head motorization hard- and software kit for the KEMAR HATS, is covered. This motorization kit allows to generate HRTF datasets of the KEMAR with variable HATO in an automated procedure. Since this is an externally mounted kit, the acoustic influence of the motorization on measured HRTFs is investigated in addition to its positioning precision. The results show that both the positioning precision and the acoustic impact of the motorization on the measured HRTFs should allow for its usage in various applications.

1 Introduction

Head related transfer functions (HRTFs) describe the sound incidence from a source to the left and right ear, thus incorporating fundamental cues for human spatial hearing [1]. The role of the torso in the analysis of HRTFs has been extensively studied in the past [2, 3, 4, 5, 6]. For instance, Algazi et al. [4] demonstrated that the torso influences HRTFs through shadowing as well as reflection with the latter resulting in a comb-filtering effect.

High-resolution HRTF datasets usually employ a fixed head-above-torso orientation (HATO) with various sound incidence directions, which is widely used in auralizations [7, 8, 9, 10]. However, this fixed HATO does not accurately represent real-life situations where the head and torso are typically moved independently. Brinkmann et al. [11] highlighted the audible impact of correct HATO on auralization, showing noticeable differences between constant and variable HATO for different source positions and audio content [11]. Furthermore, Algazi et al. [2] showed that the torso also affects sound localization.

Due to the relevance of HATO, there are already a few data sets and/or appartures to perform investigations with variable/correct HATO. Brinkmann et al. [12] e. g. have published a high resolution and full-spherical HRTF database for different HATO using the FABIAN dummy head [13]. While the dataset provides high quality data, it is based on a very specific, self-constructed head and torso simulator (HATS), which makes it difficult to reproduce and/or extend.

Comparable datasets with variable HATO for more common commercial HATS such as the Knowles Electronics Manikin for Acoustic Research (KEMAR) do

not exist in a similar quality to the best of our knowledge at the date of publication of this paper.

If such a data set is to be created for a HATS such as the KEMAR, motorization of the head rotation is essential for both timing and precision reasons. While there is already a motorization kit published for the KEMAR, namely the *Two!Ears: Motorization of the KEMAR head* [14], some listed components like the servo controller are no longer available one-to-one and parts of the KEMAR have to be replaced during assembly.

Therefore, we decided to develop our own low-cost open-source head motorization kit for the KEMAR HATS, called the LoCOMo kit. The goal was to develop a low-cost non-invasive and easy to build as well as simple to use motorization kit with off-the-shelf components. Therefore, our design incorporates an external toothed belt drive and an Arduino-controlled stepper motor, accompanied by 3D printed parts, a basic electronic circuit, and a UDP-based MATLAB interface for control. By making the motorization kit publicly available to the audio scientific community, we aim to encourage further research in this field [15].

Given the externally mounted construction of the LoCOMo kit, it is important to investigate any potential acoustic influence on measured HRTFs in order to assess the suitability of the kit for the different conceivable applications. Similar investigations have been conducted for headphones and head-mounted displays (HMDs) in the past, examining changes in binaural cues or spectral properties [16, 17, 18, 19] as well as coloration and localization accuracy [20, 21, 17].

Poirier-Quinot et al. e. g. investigated the impact of wearing HMDs (Microsoft HoloLens and Meta Quest 2) on localization accuracy [21]. Although wearing HMDs had a significant impact on the participant's localization performance, it was concluded that the differences were sufficiently small such that it could be considered as an HMD-free condition in all but the most demanding AR auditory localization studies [21]. Ahrens et al. [22] found similar small detrimental effects in localization accuracy when wearing HMDs, especially for lateral sound source locations.

Genovese et al. [18] objectively investigated the acoustic perturbations caused by two different HMDs (Microsoft HoloLens and Metavision Meta 2) on HRTFs. They found non-negligbile distortions, which were mainly present at the contralateral anterior quadrant and ipsilateral posteroir and in the frequency range of 3 kHz to 8 kHz. In a similar procedure, Pörschmann et al. [16] analyzed the influence of wearing different headgear on measured HRTFs. In doing so, they analyzed both the influence on the spectrum and on binaural cues using a Neumann KU100 and a HEAD acoustics HMMS II.3 dummy head, either equipped with a bicycle helmet, a baseball cap, an Oculus Rift HMD, or AKG K1000 headphones. In their investigations, the spectral differences to the reference were maximal for the AKG K1000 and lowest for the Oculus Rift and the baseball cap, with the strongest deviations being found for contralateral sound incidence directions. The AKG K1000 also had the highest deviation in interaural time differences (ITDs) and interaural level differences (ILDs). For the Oculus Rift, the ITDs and ILDs were mainly affected for frontal directions, while only a very weak influence of the bicycle helmet and the baseball cap were found [16].

In Section 3, we conduct a comparable objective investigation on the acoustic influence of the LoCOMo kit by examining both binaural and spectral properties. Before delving into that analysis, we provide a brief introduction to the LoCOMo kit in Section 2 and analyze its precision.

2 The LoCOMo kit

The LoCOMo kit is based on an Arduino microcontroller, a stepper motor driving a toothed belt together with 3D printed parts for mounting. In addition, two limit switches are incorporated, which on the one hand are used for the calibration process and on the other hand hard limit the range of motion and serve as a safety mechanism to avoid collisions. Given the provided implementation of the microcontroller and the MATLAB API, the LoCOMo kit allows for head rotation from $-90°$ to $90°$ in $1°$ steps, the latter being an implementation choice with finer resolutions being theoretically possible. The motor itself has a step resolution of 200 steps per $360°$ with eight substeps, allowing for a theoretical resolution of $0.225°$. Depending on the gear ratio from motor to neck, even finer resolutions are possible. To facilitate usage, an OLED display integrated into the control box gives instructions on how to use the kit and shows information like the current position or the IP and port for the UDP interface during usage.

Photos of the LoCOMo kit mounted on the KEMAR as well as a view inside the control box with the lid

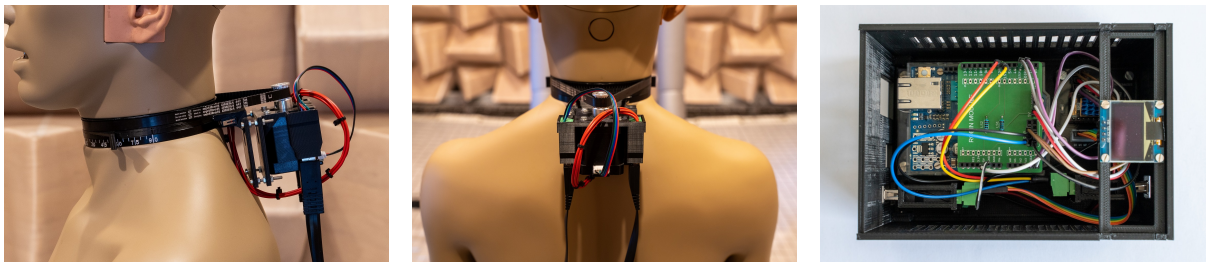


Fig. 1: Photos of the LoCOMo kit being equipped on the KEMAR dummy head from the left and back as well as a view inside the control box.

removed are shown in Fig. 1. For further details, the reader is referred to `https://go.lu-h.de/locomo` [15].

2.1 Precision evaluation

In this section, the precision of the LoCOMo kit is investigated. If maximum precision is needed, the kit supports are so-called precision mode, where a position is always approached from the same side, reducing inaccuracies induced e. g. by non-optimal tensioning of the toothed belt. However, the regular, out-of-the box setup, should be sufficiently precise for almost all applications. The following examinations were nevertheless performed in precision mode and with default speed.

To measure the precision of the LoCOMo kit, head orientation was measured using an OptiTrack motion capture system with two rigid body markers screwed to the ear mounts of the KEMAR. A maximum measurement error of the system of 0.067° was identified by 100 repeated measurements in each of the 181 positions (from −90° to 90° in 1° steps). The distribution of the absolute errors when approaching 1000 random head orientations is shown in Fig. 3. The LoCOMo kit achieved a median absolute error of 0.06° which is in the order of magnitude of the measurement error. The 75 % quartile is 0.087° and the maximum measured error is 0.19°. Thus, with a precision of less than 0.2°, the kit should offer sufficient accuracy for almost all applications.

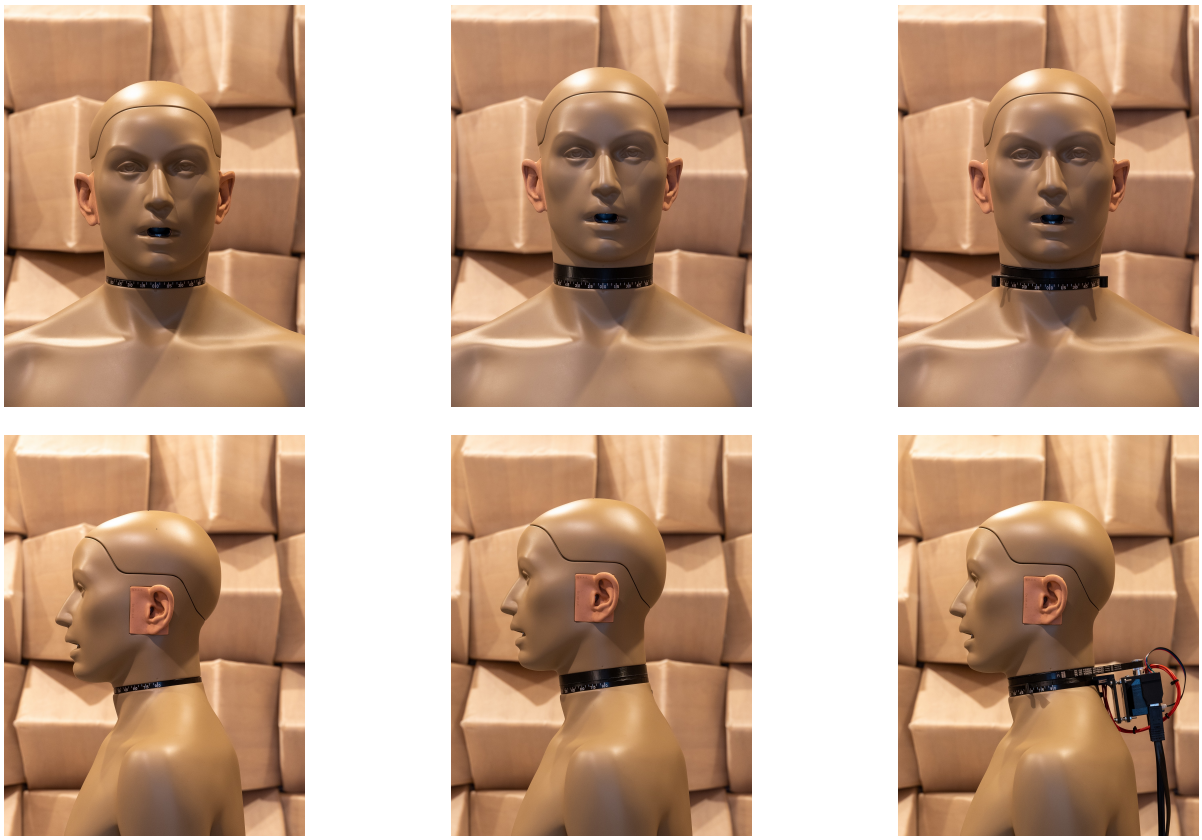
3 Influence of the LoCOMo kit on the HRTFs

As mentioned in the introduction, it can be assumed that this external motorization has an acoustic influence on the HRTFs. In this section, we therefore examine the differences between the original HRTFs (*Orig*) and those recorded with the LoCOMo kit (*LoCOMo*). In addition, to eliminate the differences in the HRTFs only due to the neck extension of about 16 mm, we compare the HRTFs with those in which only a spacer of the corresponding length was applied (*Spacer*). Attaching a neck extension to the KEMAR is quite a common procedure and corresponding spacers are also sold by the original equipment manufacturer.

In the following, ϕ denotes the azimuth angle of a sound source relative to the torso of a HATS. Thereby, azimuth angles of $\phi = 0°, 180°, 90°, 270°$ denote sources in front of and behind, and to the left and right of the torso of the HATS, respectively. Furthermore, θ denotes the elevation of the source with $\theta = 0°$ at the horizontal plane and $\theta = 90°$ at the zenith. The HATO is denoted as ϕ_{HATO}, so that a HATO of $\phi_{\mathrm{HATO}} = 45$ corresponds to a counterclockwise rotation of the head above the torso of 45°.

A HRTF for a sound source at (ϕ, θ), a HATO of ϕ_{HATO} at frequency f is therefore denoted as $\mathrm{HRTF}(\phi, \theta, \phi_{\mathrm{HATO}}, f)$, with arguments possibly being omitted for better readability, depending on the context.

Due to the large amount of time required for measurements without motorization but with variable HATO, we limited ourselves to a few critical elevations. These were chosen with $\theta \in \{-40°, -20°, 0°, 20°, 40°\}$ so that both reflection and shadowing effects were considered. For azimuth ϕ and HATO ϕ_{HATO} a resolution of 10° was chosen. A special case forms the investigation of the broadband ILDs and ITDs in the horizontal plane. For this purpose, HRTFs with an azimuth resolution of 1° were measured at a HATO of 0°.



(a) (b) (c)

Fig. 2: The different investigated configurations of the KEMAR 45BC-12 HATS (a) without modifications (*Orig*), (b) with a neck extension of 16 mm (*Spacer*), and (c) equipped with the LoCOMo kit (*LoCOMo*) shown from the front (top) and side (bottom).

3.1 HRTF measurement setup

The measurements were carried out in the anechoic chamber of the Institute of Communications Technology at Leibniz University Hannover. The room has dimensions of 3.9 m×3.4 m×2.6 m with a cut-off frequency of about 100 Hz. The measurements were conducted using a Genelec 8331A, which has a flat on-axis frequency response from 45 Hz to 37 kHz (±6 dB), the GRAS KEMAR 45BC-12 positioned on a Norsonic No265 turntable, and an Audio Precision APx525 audio analyzer with APx500 v5.0.3 measurement software. The loudspeaker and dummy head were positioned such that the acoustic center of the loudspeaker was in line with the ear level of the HATS. All measurements were conducted at a distance of 1 m. To measure the omnidirectional impulse response of the loudspeaker used for the free-field equalization, a GRAS 40HL microphone was used and positioned at the acoustic center of the HATS with the HATS removed. The impulse responses were measured with a two seconds long exponential sweep with a frequency range between 20 Hz and 20 kHz at a sampling frequency of 48 kHz. The excitation signal was leveled to 85 dBSPL for a sinusoidal tone at 1 kHz measured with the KEMAR equipped with the LoCOMo kit.

Fig. 4 illustrates the measurement setup featuring the KEMAR equipped with the LoCOMo kit and the loudspeaker positioned at an elevation angle of $\theta = 0°$.

In a subsequent postprocessing all HRIRs were first windowed using the same hybrid rectangular window with a 10-sample Hann on- and offset with a flat section in between. This was followed by a spectral division

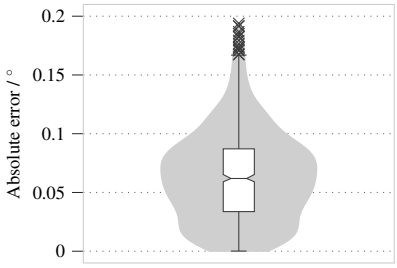


Fig. 3: Distribution of absolute errors of the head motorization shown as violin and box plot. The boxes are drawn from the first to the third quartile. The horizontal line shows the median. The whiskers go from the lowest data still within 1.5 IQR of the lower quartile to the highest data within 1.5 IQR of the upper quartile. The outliers (outside of 1.5 IQR) are marked as crosses.

with regularization between 20 Hz and 20 kHz of the HRTFs with the omnidirectional reference measurement. Windowing and spectral division were conducted using methods from the ITA toolbox [23]. The HRIRs have a final length of 256 samples (5.33 ms) to exclude reflections back from the loudspeaker.

Fig. 5 exemplarily shows HRTFs of the right ear for frontal $(\phi,\theta) = (0°,0°)$ and contralateral $(\phi,\theta) = (270°,0°)$ direction with a HATO of 0°.

While great care was taken during the measurements, it is important to acknowledge several factors that may have influenced the results. Certain influences, such as the truss arcs, are difficult to quantify but are expected to have minor impact as they were consistent across all configurations. On the other hand, inaccuracies arising from non-exact positioning of the dummy head and loudspeaker are likely to be more significant. Specifically, in the Orig configuration, the vertical alignment had to be adjusted due to the absence of a neck extension.

3.2 Binaural cues

For a first comparison, broadband ILDs and ITDs were calculated from the HRIRs measured in the horizontal plane with 0° HATO. The ITDs were calculated from the original phase HRIRs using the threshold onset method with ten-times oversampling with an onset threshold of −10 dB to the overall peak value [16, 24].

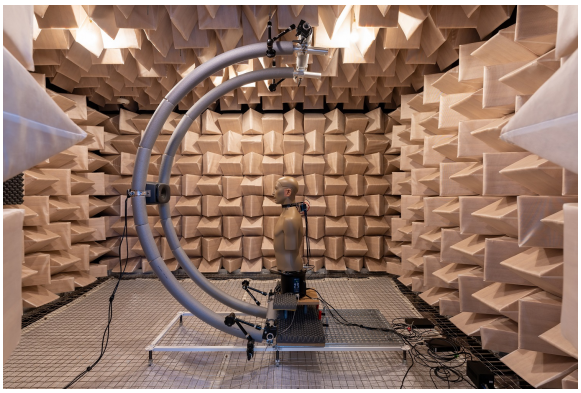


Fig. 4: HRTF measurement setup with the loudspeaker positioned at 1 m distance to the acoustic center of the KEMAR head.

Broadband ILDs were estimated as RMS level differences between left and right ear [11].

As depicted in Fig. 6a, there are only very small differences in the ITDs between the different configurations. The average and maximum ITD differences of the LoCOMo configuration to the Spacer configuration are 1.5 µs and 8.3 µs, respectively and therefore below the ITD difference threshold of 10 µs [1, pp. 152]. The average distance in ITD to the Orig configuration of 4.5 µs is also below the threshold, while the maximum difference is slightly above with 18.8 µs and thus above the difference threshold.

As can be seen in Fig. 6b, the ILDs of the different configurations also match very well. The average ILD differences of the LoCOMo configuration to the Spacer and Orig configurations of 0.28 dB and 0.53 dB, respectively, are below the difference threshold of 0.6 dB [1, pp. 160]. The maximum deviations, however, are above the assumed threshold with 1.03 dB for the Spacer and 1.51 dB for the Orig configuration.

3.3 Spectral structure

To analyze the spectral deviations of the LoCOMo configuration to the Orig and Spacer case, spectral differences were averaged over all $n_\Omega = 3420$ measurements according to (1):

$$\Delta G_f(f) := \frac{1}{n_\Omega} \sum_{\omega\in\Omega} \left| 20\log_{10} \frac{|\mathrm{HRTF}_{\mathrm{LoCOMo}}(\omega,f)|}{|\mathrm{HRTF}_{\mathrm{Test}}(\omega,f)|} \right|. \quad (1)$$

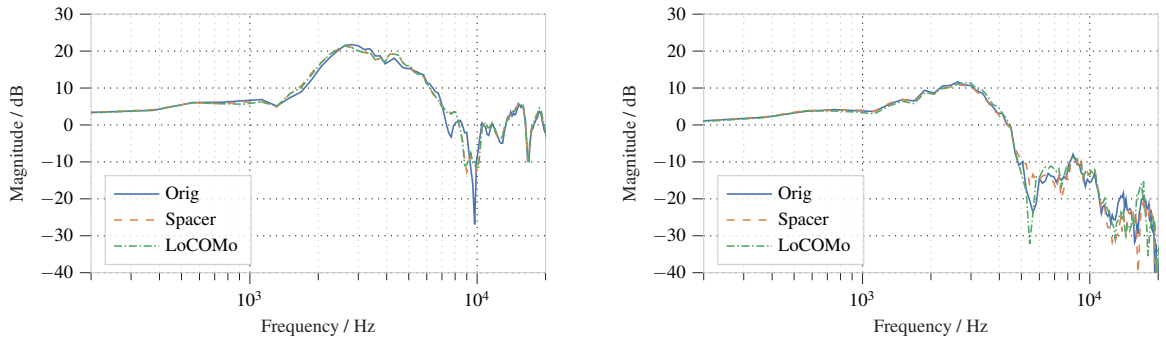


Fig. 5: HRTFs of the right ear for the Orig, Spacer, and LoCOMo configuration for a sound source at (a) frontal and (b) contralateral direction with 0° HATO.

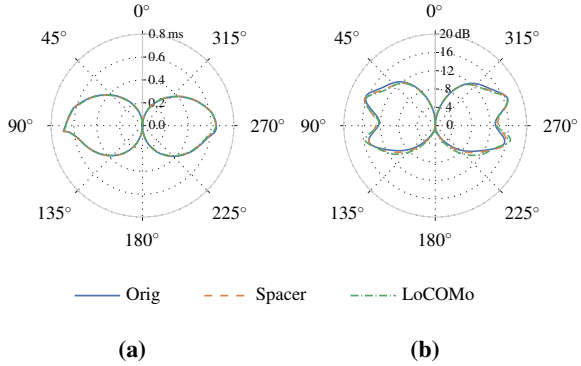


Fig. 6: Broadband ITDs (a) and ILDs (b) for the different configurations in the horizontal plane with 0° HATO.

Thereby, $\Omega := \{(\phi, \theta, \phi_{\text{HATO}})\}$ denotes the set of all combinations of sound source position (ϕ, θ) and HATO ϕ_{HATO} investigated here, f is the temporal frequency in Hz, $\text{HRTF}_{\text{LoCOMo}}$ the reference HRTF corresponding to the LoCOMo configuration and $\text{HRTF}_{\text{Test}}$ the HRTF of one of other two configurations with $\text{Test} \in \{\text{Orig}, \text{Spacer}\}$.

Fig. 7 shows the corresponding results for the right ear. For both configurations and all frequencies, the differences are rather small (less than 3.2 dB) and therefore comparable or even smaller than those of different headgear [16]. Below 2.5 kHz the differences are generally less than 1 dB on average, which can be probably attributed to the larger wavelength in relation to the size of the motorization. The largest differences become noticeable at approximately 6 kHz. However, for each frequency, the differences compared to the Spacer configuration are considerably smaller, while both curves exhibit a similar shape. The difference between Orig and Spacer is expected to be mainly due to the neck extension, indicating that the rest of the apparatus therefore amounts to less than 2.3 dB at around 17 kHz.

In addition to the frequency-dependent deviations, the spatial distribution of the spectral differences were also investigated. In doing so, the HRTFs in the different configurations were analyzed in a Gammatone filter bank $C(f, f_c)$ with $n_{f_c} = 41$ center frequencies f_c in a frequency range of 20 Hz to 20 kHz, following a similar procedure as in [25, 11].

Averaging the spectral differences in each filter band $\Delta\widetilde{\text{HRTF}}(f_c)$ with

$$\Delta\widetilde{\text{HRTF}}(f_c) := 20 \log_{10} \frac{\widetilde{\text{HRTF}}_{\text{LoCOMo}}(f_c)}{\widetilde{\text{HRTF}}_{\text{Test}}(f_c)}$$

and

$$\widetilde{\text{HRTF}}(f_c) := \int_{20}^{20000} C(f, f_c) \, |\text{HRTF}(f)| \, \mathrm{d}f$$

results in the single-value error measure ΔG_{sp} defined as

$$\Delta G_{\text{sp}} := \frac{1}{n_{f_c}} \sum_{f_c} \left| \Delta\widetilde{\text{HRTF}}(f_c) \right| . \tag{2}$$

Please note that the dependency on source position and HATO was omitted for better readability and thus $\Delta G_{\text{sp}} = \Delta G_{\text{sp}}(\phi, \theta, \phi_{\text{HATO}})$.

Fig. 8 shows the spectral differences of LoCOMo in comparison to the Orig and Spacer condition averaged across frequency according to (2) for the right ear without applying any interpolation. Overall, differences are rather small (less than 2.2 dB) with higher values being observed for combinations of azimuth and HATO resulting in contralateral positions, especially for non-negative elevations and in comparison to the Spacer configuration. Table 1 provides the maximum spectral differences for each investigated elevation.

In both Fig. 8 and Table 1 it can also be seen that the deviations increase with decreasing elevation, which is probably due to shadowing effects caused by the kit. This can also be seen from the fact that the differences increase in the azimuth range from about 180° to 270° for almost all HATOs, which are just the constellations where the kit is located between the sound source and the ear.

Fig. 9 shows the same data as Fig. 8, though this time arranged individually as a function of azimuth and elevation for each HATO between $-60°$ and 60°. Again, it can be seen that the largest average spectral differences are found for constellations where sound source, motor and ear are approximately aligned, such as for ϕ_{HATO}, $\theta = -20$, and $\phi = 210$, in both the Orig and Spacer configuration.

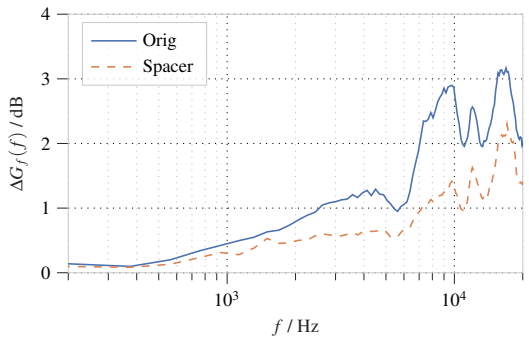


Fig. 7: Spectral differences of the LoCOMo setup compared to the Orig and Spacer configuration of the right ear averaged over all sound source positions and HATOs.

4 Conclusion

A low-cost open source head motorization kit for the KEMAR HATS was presented and analyzed. With a

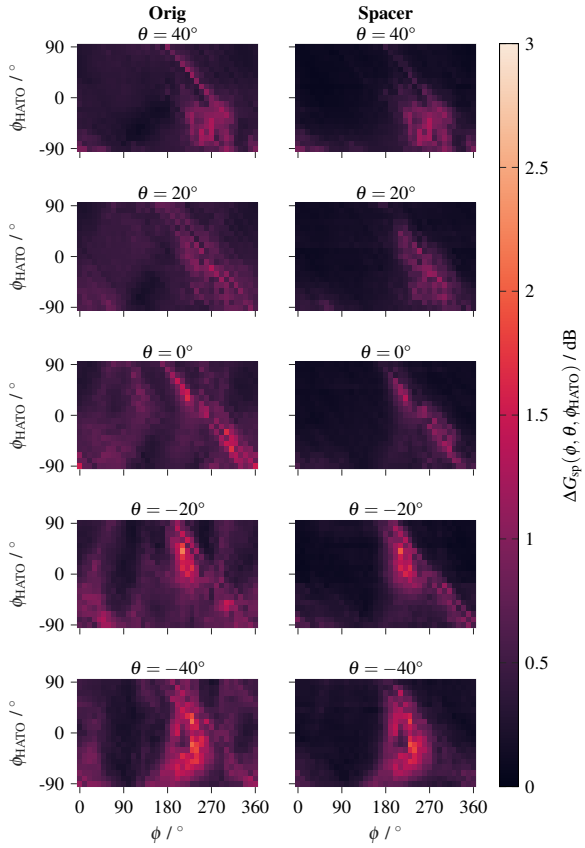


Fig. 8: Spectral differences $\Delta G_{\mathrm{sp}}(\phi, \theta, \phi_{\mathrm{HATO}})$ of the right ear averaged across frequency according to (2) in comparison to the Orig (left) and Spacer (right) configuration and the different elevations.

Table 1: Maximum spectral distances of the LoCOMo setup compared to Orig (a) and Spacer (b) configuration for each elevation investigated.

(a) Orig.

θ / °	$\arg\max_{\phi,\phi_{\mathrm{HATO}}} \Delta G_{\mathrm{sp}}$		$\max_{\phi,\phi_{\mathrm{HATO}}} \Delta G_{\mathrm{sp}}$
	ϕ / °	ϕ_{HATO} / °	ΔG_{sp} / dB
40	270	-40	1.3
20	270	-20	1.3
0	310	-60	1.7
-20	210	40	2.2
-40	230	20	2.1

(b) Spacer.

θ / °	$\arg\max_{\phi,\phi_{\mathrm{HATO}}} \Delta G_{\mathrm{sp}}$		$\max_{\phi,\phi_{\mathrm{HATO}}} \Delta G_{\mathrm{sp}}$
	ϕ / °	ϕ_{HATO} / °	ΔG_{sp} / dB
40	270	-40	1.3
20	290	-40	1.2
0	220	30	1.3
-20	210	40	2.0
-40	230	-20	1.9

precision of less than 0.2°, the kit offers high precision which should allow for usage in almost all applications. In addition to the mechanical accuracy of the positioning, the acoustic influence of the motorization on the measured HRTFs was investigated by a detailed quantitative comparison of broadband binaural cues and the fine spectral structure for five different elevations and at a 10°-resolution of azimuth and HATO. In general, it can be said that the LoCOMo kit has an acoustic influence, which, however, is mainly constrained to contralateral constellations. Based on the rather small differences identified, especially to the HRTFs of a KEMAR with equivalent neck extension, we believe that the use of the motorization is feasible for automated acquisition of HRTF datasets of high quality and comparability. However, the detailed description allows potential users to evaluate the suitability of the motorization for their specific application themselves.

5 Acknowledgments

This work was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) – Project number 517437545. The authors would like to thank Anton Hilbig for his help with the microcontroller programming, circuit design, and documentation.

References

[1] Blauert, J., *Spatial Hearing: The Psychophysics of Human Sound Localization*, The MIT Press, 1996, doi:10.7551/mitpress/6391.001.0001.

[2] Algazi, V. R., Avendano, C., and Duda, R. O., "Elevation localization and head-related transfer function analysis at low frequencies," *The Journal of the Acoustical Society of America*, 109(3), pp. 1110–1122, 2001, doi:10.1121/1.1349185.

[3] Gardner, M. B., "Some monaural and binaural facets of median plane localization," *The Journal of the Acoustical Society of America*, 54(6), pp. 1489–1495, 1973, doi:10.1121/1.1914447.

[4] Algazi, V. R., Duda, R. O., Duralswami, R., Gumerov, N. A., and Tang, Z., "Approximating the head-related transfer function using simple geometric models of the head and torso," *The Journal of the Acoustical Society of America*, 112(5), pp. 2053–2064, 2002, doi:10.1121/1.1508780.

[5] Merimaa, J., *Analysis, synthesis, and perception of spatial sound: Binaural localization modeling and multichannel loudspeaker reproduction*, Dissertation, Helsinki University of Technology, Espoo, 2006.

[6] Guldenschuh, M., Sontacchi, A., Zotter, F., and Höldrich, R., "HRTF modeling in due consideration variable torso reflections," *The Journal of the Acoustical Society of America*, 123(5), p. 3080, 2008, doi:10.1121/1.2932888.

[7] Gardner, W. G. and Martin, K. D., "HRTF measurements of a KEMAR," *The Journal of the Acoustical Society of America*, 97(6), pp. 3907–3908, 1995, doi:10.1121/1.412407.

[8] Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C., "The CIPIC HRTF database," *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, pp. 99–102, 2001, doi:10.1109/ASPAA.2001.969552.

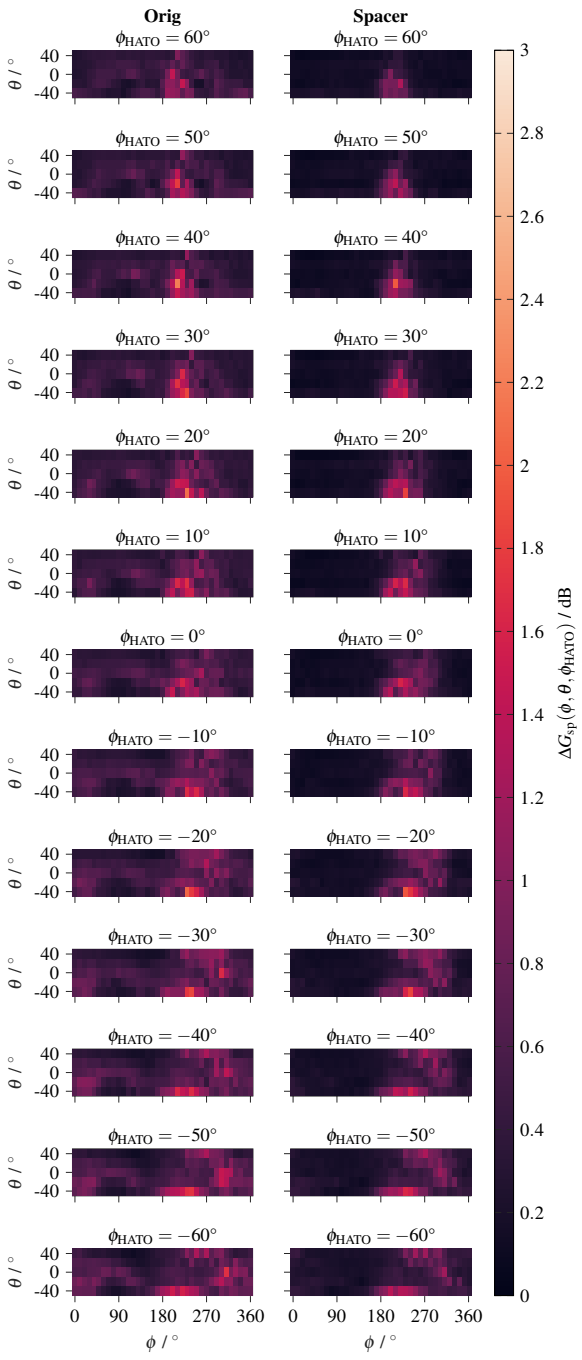


Fig. 9: Spectral differences $\Delta G_{\mathrm{sp}}(\phi, \theta, \phi_{\mathrm{HATO}})$ of the right ear averaged across frequency according to (2) in comparison to the Orig (left) and Spacer (right) configuration and the different HATOs.

[9] Wierstorf, H., Geier, M., Raake, A., and Spors, S., "A Free Database of Head-Related Impulse Response Measurements in the Horizontal Plane with Multiple Distances," Zenodo, 2016, doi:10.5281/ZENODO.55418.

[10] Braren, H. S. and Fels, J., "A High-Resolution Head-Related Transfer Function Data Set and 3D-Scan of KEMAR," RWTH Aachen University, 2020, doi:10.18154/RWTH-2020-11307.

[11] Brinkmann, F., Roden, R., Lindau, A., and Weinzierl, S., "Audibility and Interpolation of Head-Above-Torso Orientation in Binaural Technology," *IEEE Journal of Selected Topics in Signal Processing*, 9(5), pp. 931–942, 2015, doi:10.1109/JSTSP.2015.2414905.

[12] Brinkmann, F., Lindau, A., Weinzierl, S., van de Par, S., Müller-Trapet, M., Opdam, R., and Vorländer, M., "A High Resolution and Full-Spherical Head-Related Transfer Function Database for Different Head-Above-Torso Orientations," *Journal of the Audio Engineering Society*, 65(10), pp. 841–848, 2017, doi:10.17743/jaes.2017.0033.

[13] Lindau, A., Hohn, T., and Weinzierl, S., "Binaural resynthesis for comparative studies of acoustical environments," in *Audio Engineering Society Convention 122*, Vienna, 2007.

[14] "The Two!Ears Auditory Model documentation: Motorization of a KEMAR head," Accessed on: 31.05.2023 [Online], 2016, Available: `http://docs.twoears.eu/en/latest/robo/head-motorization/`.

[15] Poschadel, N., "LoCOMo: A Low-Cost Open-Source Head Motorization Kit for the KEMAR Head and Torso Simulator," GitLab, `https://go.lu-h.de/locomo`, 2023.

[16] Pörschmann, C., Arend, J. M., and Gillioz, R., "How wearing headgear affects measured head-related transfer functions," in *Proceedings of the EAA Spatial Audio Signal Processing Symposium*, pp. 49–54, 2019.

[17] Cuevas-Rodríguez, M., Alon, D. L., Clapp, S. W., Robinson, P., and Mehra, R., "Evaluation of the effect of head-mounted display on individualized head-related transfer functions," in M. Ochmann,

M. Vorländer, and J. Fels, editors, *Proceedings of the 23rd International Congress on Acoustics*, pp. 2635–2642, Deutsche Gesellschaft für Akustik, Berlin, Germany, 2019.

[18] Genovese, A., Zalles, G., Reardon, G., and Roginska, A., "Acoustic perturbations in HRTFs measured on Mixed Reality Headsets," in *International Conference on Audio for Virtual and Augmented Reality 2018*, Redmond, WA, USA, 2018.

[19] Gupta, R., Ranjan, R., He, J., and Gan, W.-S., "Investigation of effect of VR/AR headgear on Head related transfer functions for natural listening," in *International Conference on Audio for Virtual and Augmented Reality 2018*, Redmond, WA, USA, 2018.

[20] Lladó, P., Mckenzie, T., Meyer-Kahlen, N., and Schlecht, S. J., "Predicting Perceptual Transparency of Head-Worn Devices," *Journal of the Audio Engineering Society*, 70(7/8), pp. 585–600, 2022, doi:10.17743/jaes.2022.0024.

[21] Poirier-Quinot, D. and Lawless, M. S., "Impact of wearing a head-mounted display on localization accuracy of real sound sources," *Acta Acustica*, 7, p. 3, 2023, doi:10.1051/aacus/2022055.

[22] Ahrens, A., Lund, K. D., Marschall, M., and Dau, T., "Sound source localization with varying amount of visual information in virtual reality," *PLOS ONE*, 14(3), p. e0214603, 2019, doi:10.1371/journal.pone.0214603.

[23] Berzborn, M., Bomhardt, R., Klein, J., Richter, J.-G., and Vorländer, M., "The ITA-Toolbox: An Open Source MATLAB Toolbox for Acoustic Measurements and Signal Processing," in *Fortschritte der Akustik - DAGA 2017*, Kiel, 2017.

[24] Katz, B. F. G. and Noisternig, M., "A comparative study of Interaural Time Delay estimation methods," *The Journal of the Acoustical Society of America*, 135(6), pp. 3530–3540, 2014, doi:10.1121/1.4875714.

[25] Schärer, Z. and Lindau, A., "Evaluation of Equalization Methods for Binaural Signals," in *Audio Engineering Society Convention 126*, Munich, 2009.