Effect of Using Individually Simulated HRTFs on the Outcome of Tournament Selection Procedures in a Virtual Environment.

Michael Oehler **Minh Voong** **Marlon Regener** **Maurício do V. M. da Costa**

Music Technology & Digital Musicology Lab (MTDML),
Osnabrück University, Germany,
michael.oehler@uos.de

ABSTRACT

Individualized head related transfer functions (HRTF) play an increasingly important role in the field of virtual acoustics, especially for the perceived immersion in virtual environments. A central question in this context is often how exactly the individual fit is realized best. In the current exploratory study, a perception based tournament task is combined with numerically simulated HRTFs. HRTF selection performed in this way is relatively time-saving and does not require much technical effort. The two main objectives are to analyze the impact of numerically simulated individual HRTFs on the outcome of a tournament task and to investigate whether this correlates with the performance in a localization task in which the participants have to spatially localize sounds using the different HRTFs from the tournament task. The study is divided into 3 consecutive parts. First, the participants' individual HRTFs are simulated based on a 3D scan, then each participant takes part in the tournament task to identify the best fitting HRTF, and then the best (and worst) fitting HRTFs are validated in a localization task along with the simulated HRTF. The results of the tournament task show that the numerically simulated individual HRTFs had little effect on the outcome. In the localization task, however, the simulated HRTFs produced the best results, on average.

1. INTRODUCTION

Individualized head related transfer functions (HRTF) play an increasingly important role in the field of virtual acoustics, especially for the acoustic layer in virtual environments (VE). Various studies have shown that well-fitting HRTFs have a positive effect on the perceived immersion in virtual (VR), augmented (AR) or mixed (MR) reality applications [1]. This is especially relevant for music-related applications, where the acoustic domain is of particular importance. However, shortcomings in the determination of the individually appropriate HRTF may result in little or no advantage over generic HRTFs [2]. This depends on many factors, such as the specific measurement procedure [2], the type of validation [3], or the inclusion of additional sensory modalities [4].

A central question in this context is often how exactly the individual fit is realized best. A detailed overview of the different approaches to customizing HRTFs is provided in [5]. Acoustic measurements, the probably most obvious method, are usually very time and resource consuming, although new approaches have now been developed to rapidly measure HRTFs in ordinary home environments. For a recent review on different approaches in this area see [6]. Another technique is numerical simulation, e.g., using Fast-Multipole-accelerated Boundary Element Method (FM-BEM) [7], Finite Difference Time Domain Method (FDTD) [8] or Finite Element Method (FEM) [9]. Comparisons between numerical simulations and acoustic measurements have been around for a long time (e.g. [10]), but have been increasingly discussed again in recent years.

Besides the approach of indirect individualization based on anthropometric data [11-13], individualization based on perceptual feedback seems to be promising for practical everyday use, since HRTF selection performed in this way is relatively time-saving and possible without significant technical effort. Several selection methods have been tested for this purpose [14-17]; a good compromise between accuracy and time required is provided by the Swiss tournament system, for example [18,19]. In corresponding tournament tasks, participants are usually asked to evaluate the difference between two HRTFs based on certain perceptual criteria such as first impression, envelopment, and externalization.

In previous studies, we tested whether a corresponding tournament task could be used to identify individual HRTFs for each person who would also perform best in a spatial localization task [18, 20]. Although better results tended to be found in the localization task for the "winning" HRTFs of the tournament task, the effect sizes were rather small. This could be partly due to the fact that the pre-selection of the HRTFs in the tournament task was realized via a random selection from common HRTF libraries (CIPIC, LISTEN, etc.) and thus this specific selection may have been less well suited to individual participants.

Therefore, in the current study, a tournament task is developed including each participant's individually simulated HRTF in addition to a random selection of profiles from common databases. Although there are numerous sources of potential bias in the numerical simulation of HRTFs that distort the resulting HRTF compared to the participant's actual HRTF (e.g. when creating/scanning 3D models or post-processing meshes), it is assumed that an HRTF that is not equal to, but close to, the participant's own HRTF will score better than the random HRTFs from the databases. Should this be confirmed, simulated HRTFs could then be included in the compilation of individual HRTF profiles of tournament sets in future studies, e.g., multiple simulated models of an individual that differ in certain parameters that, however, still need to be determined.

Accordingly, the two main objectives of the present study are (a) to analyze the impact of numerically simulated individual HRTFs on the outcome of a tournament task that uses perceptual categories such as externalization, envelopment, etc., and (b) to investigate whether this correlates with the performance in a localization task in which the participants have to spatially localize sounds using the different HRTFs from the tournament task.

Another general objective is to orient the methods used in the study towards a future application-oriented suitability for everyday use. This means, for example, that the scans of the 3D models for the numerical simulation should not be performed with expensive, difficult-to-access and usually permanently installed laboratory scanners, but with solutions comparable to a scan with (e.g., LIDAR) scanners of future smartphone generations. It is assumed that the loss of precision compared to laboratory scanners and conditions can then (partly) be compensated by the proposed tournament procedure with several variants of the individually simulated HRTF. However, this aspect goes beyond the scope of the current study.

Furthermore, the listening experiment is to be realized with bone conducting headphones (BCH) and not with "regular" headphones (RH) placed in, on or above the ear. Especially for MR or AR applications that rely on the perception of both natural and virtual sounds, the use of RH is less suitable [21]. Since BCH transmits sound via vibrations and the speakers are placed behind the ear or at the temple, i.e., they radiate sound directly into the inner ear [22], they are more suitable for transmitting the acoustic signal in AR or MR applications. Of course, this could be a critical factor that biases the results in both tasks and therefore needs to be carefully controlled. On the other hand, previous studies have shown that there is no significant difference between the use of RH and BCH in comparable tournament and localization tasks [20]. Although there are currently no commercial MR or AR products on the market that use BCH or comparable technologies (e.g., cartilage conduction), the potential benefits are obvious.

2. METHOD

The study is divided into 3 consecutive parts. First, the participants' individual HRTFs are simulated, then each participant takes part in the tournament task to identify the best fitting HRTF, and then the best fitting HRTF is validated in a localization task along with the HRTF ranked last and the simulated HRTF for comparison.

2.1 Participants

A total of 10 participants took part in the study (3 female, 7 male, M = 25.6, SD = 4.54). All participants had to go through both the tournament task and the localization task and all of them had normal hearing.

2.2 Numerical Simulation of HRTFs

The numerical simulations were based on the following phases: (a) 3D scanning, in which a scan of each participant is taken, resulting on a 3D mesh; (b) pre-processing, where the meshes are digitally treated and prepared for the numerical calculations; (c) numerical calculation, in which the transfer function between several points in space was computed for the faces indicated on the participant's mesh as left and right ears; and (d) post-processing, in which the results were converted into a single file containing the HRIRs for each participant. Such steps are detailed in the following.

2.2.1 3D Scanning

In order to get as close as possible to a future application-oriented suitability for everyday use, the 3D-scans of the participants were performed using the consumer-friendly and relatively affordable POP 3D Scanner by Revopoint[1]. Different approaches were tested to get the best possible scans with the available hardware. A free-roaming handheld scan offers the most promising results so far (which is encouraging for a scan provided by a smartphone will almost always be handheld). Further automation and routines for the scanning process are planned for future research. The participants were scanned from the utmost tip of the head to the top of the breast. Detailed scanning was required in the area of the ears while the remaining parts only needed to be roughly captured in order to get the data required for the HRTF simulation. The meshed models were then exported as Polygon File Format (*.ply) and partly edited in post-processing to smooth out rough textures and align the model in the right direction.

2.2.2 Numerical Calculation

Once the 3D meshes were generated, a pre-processing procedure was performed in Blender,[2] a free open-source software for 3D modeling. There, unwanted details were

[1] See https://de.shop.revopoint3d.com (last viewed: Mar. 30, 22)
[2] Available from http://www.blender.org (last viewed: Feb. 03, 22)

smoothed out and artifacts related to the scanning were corrected. Different resolutions were adopted for different regions of the mesh, as to reduce the computational burden. The ears were re-meshed to have a resolution of 1.3 mm, which seemed to be compatible with the details provided by the scanner; the region near the ears and the rest of the head had 3 mm and 6 mm resolutions, respectively; shoulders had 16 mm, and the upper section of the torso had a 33 mm resolution. Figure 1 depicts a participant's mesh, in which one can see the final result of the pre-processing procedure just described.

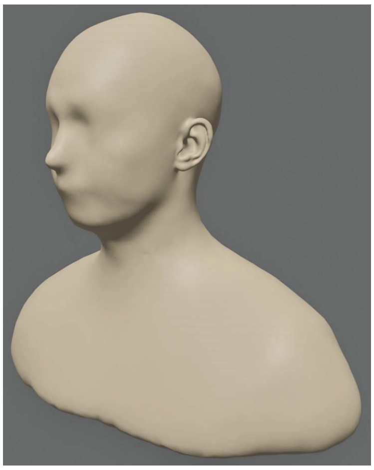

Figure 1. Snapshot of a participant's mesh after pre-processing.

After the meshes were pre-processed, the numerical calculations were conducted using the MESH2HRTF [23,24], which is an open-source library based on the fast-multipole BEM solver. The BEM simulations were performed for the frequency spectrum ranging from 100 Hz to 16 kHz with a 100 Hz step, and the solutions were sampled at 1550 collocated spatial directions distributed at the surface of a sphere with 1.2 m of radius centered at the participant's head. The final HRIRs were then resampled to a sampling rate of 44.1 kHz.

2.3 Tournament Task

A tournament procedure was chosen on the one hand to ensure comparability with previous studies, e.g. [18,20], and on the other hand because it allows several HRTFs to be included in the comparison at the same time in a relatively time-saving manner. The tournament task is mostly identical to the implementation in [18], where participants had to compare different HRTFs in a virtual acoustic environment. In the current study, the experiment itself was also conducted in a VE. In other words, participants can not only hear the sound sources, but also see them. The environment consists of the urban soundscape of a city park with different sound sources. The focus was on high ecological validity by adding objects with their corresponding sounds that fit into an urban park, e.g., car sounds, bird sounds, and a drone flying in a circle above the participant.

The Swiss tournament format, in which equal participants compete against each other, has proven useful. This format is widely used in chess and other sports and offers a good compromise between the time-consuming round robin format and a knock-out system, where promising HRTFs might be sorted out too early due to unforeseen circumstances [18].

Care was taken to ensure that subjects did not use one HRTF for a disproportionately long or short time in each pairwise comparison. For this purpose, the respective time spans were recorded and, in addition, irregularities were noted in the experimental protocol.

2.3.1 Apparatus

To ensure an immersive experience and allowing unrestricted exploration of the environment, the HTC VIVE Pro Eye [25] (including wireless adapter) was used for the study. The VR system's built-in on-ear headphones were not used, as the audio signal was played back through the bone-conducting TREKZ Titanium headphones[3] via Bluetooth. A VIVE controller is used as a virtual pointing tool to make judgments. The room size in which participants can move around is about nine square meters.

For the creation of the virtual environment, Unity [26] is used as a software framework. Its modularity is an important feature, as it makes it easy to implement plugins and additional features in existing projects. In this case, the SOFAlizer plugin [27] adds an auditory spatialization function to Unity's sound engine, which can be used to perform interpolations in relation to the distance of the measurement directions of the HRTF. Another feature is the option to preload up to ten user-defined HRTFs in memory and swap them on-the-fly. This makes it possible for the participants to perceive and evaluate the acoustic environment in real time with different HRTFs. On the other hand, audio signal processing features in Unity are limited to basic spatialization functions, as the focus of the Unity engine is more on visual components. Further processing such as phase shifting or other functions are not offered and rely heavily on third-party solutions.

2.3.2 Test Procedure

The participants are placed in the middle of a virtual urban city park scene, where they can move freely and are solely limited to the space in reality. A virtual display floating in the environment shows the study instructions and also allows the input of data during the study. After recording the general data, such as age and experience with VR environments, the tournament phase begins. Six HRTFs, mostly from previous studies [18,20] and proven to be suitable for localization tasks in several other studies [28,29], compete against each other. For this study the HRTF of the KEMAR artificial head [30] and the

[3] Discontinued product; actual product portfolio in https://shokz.com/ (date last viewed: Feb. 03, 22).

participant's own HRTF are included, increasing the total number of HRTFs to eight (see Tab. 1). The audio of the scene is enabled as soon as the task begins.

HRTF0	LISTEN_1014	Resolution Elevation	
HRTF1	CIPIC_058	Kemar	50 locations
HRTF2	LISTEN_1022	CIPIC	50 locations
HRTF3	KEMAR head	LISTEN	50 locations
HRTF4	LISTEN_1028	Resolution Azimuth	
HRTF5	Personal HRTF	Kemar	approx 5° increments
HRTF6	LISTEN_1049	CIPIC	approx 5° increments
HRTF7	CIPIC_124	LISTEN	15° increments
TestSignal		Measurement Distance	
Kemar	ML Sequence	Kemar	1.4m
CIPIC	Golay-Code	CIPIC	1m
LISTEN	Sweep	LISTEN	1.95m

Table 1. Used HRTFs in the tournament task [31-33].

The tournament consists of six rounds with four matches each. In each match two HRTFs are competing against each other. The participant is asked to rate the difference between these two HRTFs using different criteria, such as *first impression*, *envelopment*, and *externalization*, for three different sound sources: cars, birds, and the drone. Envelopment describes the degree to which the user feels acoustically embedded in the scene, while externalization indicates the degree to which the participant perceives the sound as an external source, e.g., the absence of voice-of-god-artifacts. The different increments on each slider range from 'A far better than B', 'A better than B' and 'A slightly better than B' to 'B far better than A' (see Fig. 2). During the match, the participant can switch between both HRTFs seamlessly to compare both. To ensure that an HRTF is not accidentally skipped, the system checks whether both HRTFs are activated at one point.

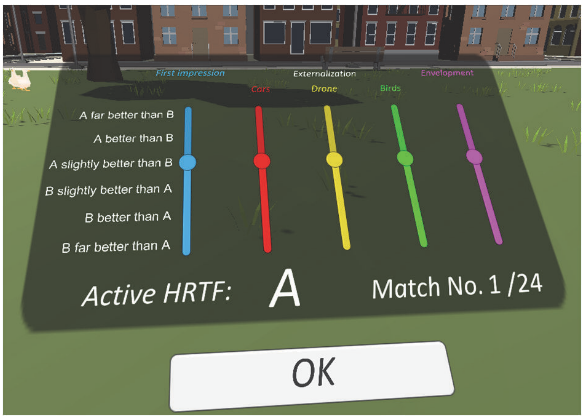

Figure 2. Screenshot of the display in the virtual environment of the tournament task.

After all matches have been played, the evaluation begins, in which the winners of the games are determined. Depending on the judgment, the correspondent HRTF gets a point from one (slightly better) to three (far better). While the "first impression" criterion is weighted with a factor of 1.5, the other criteria are weighted with 1. By comparing the sum of the points, the winner of the match is determined. In case both HRTFs have the same score, the *first impression* criteria decides the winner. The scoreboard is sorted by the number of winning matches. After that, the score for the criteria *first impression*, *externalization*, and *envelopment* is taken into account.

In the next round, the matches are paired based on the scoreboard to determine the next matches. The first place is matched against second place, the third place against the fourth place and so on. After six rounds with a total of 24 matches, the HRTF ranked first is considered the winner of the tournament and also the HRTF best suited for the participant.

2.4 Localization Task

After the tournament task, the winning HRTF, i.e., the HRTF ranked in first place, and the HRTF ranked last are used in a localization task. The personal simulated HRTF is always included in the test set, so a total of three HRTFs are tested in this task. If the personal HRTF ranks first or last, the HRTF that is next in rank to the personal HRTF is included.

The virtual environment is switched to an anechoic chamber resembling a real perception lab where loudspeaker models are arranged in a sphere around the user. Three hundred speakers are arranged in five planes and increments of 6 degrees horizontally and 14 degrees vertically. In addition to the virtual display, the representation of a dial in the loudspeaker array and a blue stripe on the front wall of the chamber facilitates the participant's orientation in the room and faster determination of the frontal direction (see Fig. 3). Before the test, the whole loudspeaker array adjusts its height to the participants head position, including an offset of 4 cm between the eye and the nose.

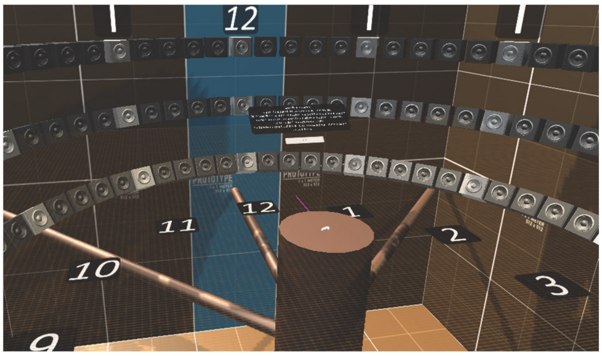

Figure 3. Screenshot of the virtual environment in the localization task.

For each HRTF, the participant is asked to determine 30 directions (see Fig. 4). The directions are given and their order is random. Also, no direction occurs more than once to make sure every direction is evaluated. The directions on the horizontal plane are identical for each layer to ensure a better comparison during evaluation. In addition, gaps have been added (front right and back left) to create a non-symmetrical distribution of the stimuli.

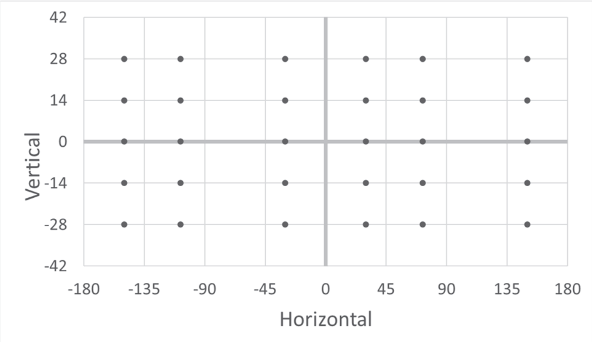

Figure 4. Distribution of the stimuli in the localization task (in degrees).

2.4.1 Test Procedure

The participant is asked to stay in the center of the loudspeaker array in the (virtual) anechoic chamber. During the trial, the sound will be muted in case they move away from the center and the distance from the center exceeds 50 cm. This is also visually implemented, as the floor is a transparent net with a circular platform in the center (see Fig. 3).

In the localization task, the participant has to judge from which loudspeaker the stimuli are originating from. A burst comprising pink noise pulses of 280 ms in length interleaved with silences of the same length is used as stimulus. To make their judgment, the participant marks the corresponding speaker with the VR controller. To avoid the user turning around and looking at the speaker to determine the origin of the sound, the stimulus is muted as soon as the difference between the participant's line of sight and the frontal direction exceeds 15°.

2.5 Ethical Approval

The study was conducted in accordance with the Declaration of Helsinki, with ethical approval obtained from Osnabrück University Ethics Committee (approval 4/71043.5). Anonymity of participants and confidentiality of their data were ensured. Participants were informed about the objectives and the procedure of the study as well as about their right to withdraw from the study at any time without adducing reasons or experiencing any negative consequences. All participants provided informed consent before participation in the study.

3. RESULTS

3.1 Tournament Task

As can be seen in Table 2, HRTF 0, 1, and 2 are overrepresented in the first two ranks, with HRTF 2 taking first place in four cases. The individual HRTFs were ranked in first place in two cases; the remaining HRTFs were ranked in the latter places, being second to last place for only one participant. While the positions in the midfield vary, i.e. different HRTFs can be found there depending on the participant, HRTF 3 (KEMAR HRTF), the only artificial head in the study, ranks last place in almost all cases. There is one exception where the HRTF of the KEMAR head is first and the personal HRTF is second to last.

Participant	HRTF							
	0	1	2	3	4	5	6	7
1	3	6	2	8	5	1	7	4
2	1	7	3	8	2	5	4	6
3	2	5	1	8	3	4	6	7
4	1	2	4	7	6	5	4	3
5	3	2	1	8	4	6	7	5
6	4	5	8	1	3	7	2	6
7	5	2	4	8	7	1	3	6
8	2	5	1	8	4	6	3	7
9	7	1	2	8	5	6	4	3
10	2	3	1	8	5	6	7	4

Table 2. Ranking of each HRTF in the tournament task (1 = first place, 8 = last place)

3.2 Localization Task

For each stimulus, the horizontal and vertical deviation between the stimulus direction and the participant's estimate of where the stimulus is coming from is calculated. Since the angular deviation is given only in positive values, the maximum deviation in the horizontal plane is 180°, and 56° in the vertical plane. As each participant had to determine the position of stimuli from combinations of 5 vertical and 6 horizontal positions, there were a total of 30 judgments.

Participant	HRTF with max. deviation in degree and [HRTF #]	HRTF with min. deviation in degree and [HRTF #]
1	52,8 [3]	36,8 [5]
2	47,2 [0]	29,4 [3]
3	39,6 [5]	31,6 [3]
4	48,6 [3]	26,4 [5]
5	56,8 [5]	43,0 [3]
6	46,2 [3]	35,8 [2]
7	24,8 [3]	17,6 [5]
8	34,0 [3]	30,4 [5]
9	55,0 [1]	49,8 [5]
10	48,8 [3]	30,0 [5]

Table 3. Average horizontal deviation for the worst HRTF (maximum deviation) and the best HRTF (minimum deviation) for the localization task. The individual numerically simulated HRTF is highlighted in grey. The specific HRTF is given in square brackets.

In a first evaluation, only the horizontal plane without elevation was considered, i.e., the mean of the horizontal deviations of all 30 judgments of each person was determined. The average deviation for the HRTF that had performed best in the localization test in each case is 33.1° for all 10 participants, and the average deviation for the HRTF that had performed worst in the localization test in each case is 45.4°. As in [20], deviations are consistently lowest around 90° and 270° and increase toward 180° and

360°. This also partly explains the relatively high deviation values on average. Moreover, a possible front-to-back confusion [34,35] is not taken into account in the calculation. It is noticeable that the winning HRTF from the tournament task was in no case the HRTF with which the best localization result was obtained, but the own (numerically simulated) HRTF is in 6 out of 10 cases the HRTF with which the best localization results are achieved (see Tab. 3). Although a chi-square test showed no significant results ($\chi^2(2, 10) = 3.800$, $p = .15$), probably mainly due to the small number of participants, the simulated HRTFs seem to have a positive effect in the localization task, in contrast to the tournament task.

In a second evaluation, the elevation was taken into account. For this purpose, the horizontal and vertical deviations were simply added for each stimulus. The average deviation for the HRTF that had performed best in the localization test in each case is 49.4°, and the average deviation for the HRTF that had performed worst in the localization test in each case is 60.4°. The results are almost identical as if only the horizontal deviation is considered. For only one participant (#10), the best performing HRTF in the localization test differed in this second evaluation, for whom their own (numerically simulated) HRTF was overperformed by the winning HRTF determined in the tournament task.

4. CONCLUSIONS

The results for the tournament task show that the numerically simulated individual HRTFs had little influence on the outcome, i.e., for the parameters *first impression*, *envelopment*, and *externalization* no advantage could be found for the participants' own HRTFs. Besides the limitation of explanatory power due to the relatively small number of participants in this explorative study, relevant issues could be (a) a possible bias due to the random selection of HRTFs from different databases, (b) the specific tournament mode (swiss style tournament), (c) the rating categories, (d) the visual/acoustic experimental environment within the VE or (e) problems due to the use of BCH. Since it has been shown in [18] and [20] that there is a significant correlation between tournament score and localization performance in a similar procedure with respect to (a), (b), and (c), the specific design of the acoustic environment seems to be important. For example, compared to [18] and [20], participants were free to move around the VE (including motion tracking) and the acoustic components of the soundscape were altered (motion path of car sounds, bird sounds present/absent, etc.). Combined with the use of BCH and the resulting poorer transmission of the higher frequencies, this may have led to the contradictory results. In a follow-up study, it would be beneficial to control these parameters in more detail, or to include them as independent variables. It could be particularly promising to use newer more advanced versions of BCH. The successor models Trekz Open Run and the recently released Open Run Pro have a significantly better transmission function, since higher frequency ranges are not only transmitted via bone conduction, but also partly hybrid as airborne sound. This presumably improves the added value of the individual HRTF in the scenario described here, while at the same time retaining the advantage of non-covered ears, which is particularly relevant for AR applications. In addition, it might be useful to extend the tournament task beyond the evaluation of a soundscape to other stimuli (speech, music, etc.).

However, the results of the localization task show that the individual numerically simulated HRTFs achieved the best results most of the time in the localization test. Although the positive influence of individual HRTFs on various perceptual capacities, in particular sound source localization tasks, has been frequently studied [1,10,11, 36,37], it is noteworthy that this also seems to hold for numerically simulated HRTFs created with a methodological approach oriented towards everyday usability. The overall relatively large angular deviations can probably also be attributed, at least in part, to the weaknesses of the BCH used and can most likely be improved by using newer models with more advanced procedures (see previous section). Another important point that should be addressed in future studies is the optimization of the scanning procedure as well as the post processing. For the intended suitability for everyday use, the scanning process must become more error-tolerant on the one hand, and it would be desirable to further automate post-processing on the other.

As an outlook, numerical HRTF simulation seems to be a promising approach for use in (acoustic) VE. Since the scanning process is always error-prone, several models of an individual could be simulated that differ in certain parameters. The best variant could then be determined in a perception-based tournament task. First, however, the tournament procedure needs to be improved because, although the simulated HRTFs provided a significant advantage in the localization task in the current study, they had little effect on the outcome of the tournament task.

Acknowledgments

The work is funded by the Volkswagen Foundation (VolkswagenStiftung) Germany within the project “New Concepts for Music and Media Technology” (grant no. 96 881).

5. REFERENCES

[1] Jenny, C. and Reuter, C., “Usability of individualized head-related transfer functions in virtual reality: Empirical study with perceptual attributes in sagittal plane sound localization,” JMIR Serious Games, 2020, 8(3), e17576.

[2] Armstrong, C., Thresh, L., Murphy, D., and Kearney, G., “A perceptual evaluation of individual and non-individual HRTFs: A case study of the SADIE II database,” Applied Sciences, 2018, 8(11), pp. 2029.

[3] Jenny, C. and Reuter, C., “Can I trust my ears in VR? Literature review of head-related transfer functions and valuation methods with descriptive attributes in virtual reality,” International Journal of Virtual Reality, 21(2), pp. 29-43.

[4] Berger, C. C., Gonzalez-Franco, M., Tajadura-Jiménez, A., Florencio, D. and Zhang, Z., “Generic HRTFs may be good enough in virtual reality. Improving source localization through cross-modal plasticity,” Frontiers in neuroscience, 2018, 12, pp. 21.

[5] Guezenoc, C. and Seguier, R., “HRTF individualization: A survey,” arXiv preprint arXiv:2003.06183.

[6] Li, S. and Peissig, J., “Measurement of head-related transfer functions: A review,” Applied Sciences, 2020, 10(14), pp. 5014.

[7] Gumerov, N. A., Duraiswami, R., and Zotkin, D. N., “Fast multipole accelerated boundary elements for numerical computation of the head related transfer function,” In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07, Honolulu, 2007, Vol. 1, pp. 165-168.

[8] Prepeliță, S., Geronazzo, M., Avanzini, F., and Savioja, L., “Influence of voxelization on finite difference time domain simulations of head-related transfer functions,” The Journal of the Acoustical Society of America, 2016, 139(5), pp. 2489-2504.

[9] Huttunen, T., Seppälä, E. T., Kirkeby, O., Kärkkäinen, A., and Kärkkäinen, L., “Simulation of the transfer function for a head-and-torso model over the entire audible frequency range,” Journal of Computational Acoustics, 2007, 15(04), pp. 429-448.

[10] Mokhtari, P., Nishimura, R., and Takemoto, H., “Toward HRTF personalization: an auditory-perceptual evaluation of simulated and measured HRTFs,” In Proceedings of the International Community for Auditory Display, 2008.

[11] Middlebrooks, J. C., Macpherson, E. A., and Onsan, Z. A., “Psychophysical customization of directional transfer functions for virtual sound localization,” The Journal of the Acoustical Society of America, 2000, 108(6), pp. 3088-3091.

[12] Zotkin, D. N., Duraiswami, R., and Davis, L. S., Customizable auditory displays. Georgia Institute of Technology, 2002.

[13] Li, L. and Huang, Q., “HRTF personalization modeling based on RBF neural network,” In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, May 2013, pp. 3707-3710.

[14] Iida, K., Ishii, Y., and Nishioka, S., “Personalization of head-related transfer functions in the median plane based on the anthropometry of the listener’s pinnae,” The Journal of the Acoustical Society of America, 2014, 136(1), pp. 317– 333.

[15] Seeber, B. U. and Fastl, H., “Subjective selection of non-individual head-related transfer functions,” in Proceedings of the International Conference on Auditory Display, Boston, 2003, pp. 259-262.

[16] D. N. Zotkin, R. Duraiswami and L. S. Davis, “Rendering localized spatial audio in a virtual auditory space,” in IEEE Transactions on Multimedia, 2004, 6(2), pp. 553–564.

[17] Shukla, R., Stewart, R., and Sandler, M., “User HRTF Selection for 3D Auditory Mixed Reality,” In Proceedings of the Sound and Music Computing Conference, 2021, pp. 84-91.

[18] Voong, T. M. and Oehler, M., “Tournament Formats as Method for Determining Best-fitting HRTF Profiles for Individuals wearing Bone Conduction Headphones,” In Proceedings of the 23rd International Congress on Acoustics integrating 4th EAA Euroregio, 2019, pp. 4841-4847.

[19] Iwaya, Y. “Individualization of head-related transfer functions with tournament-style listening test: Listening with other’s ears,” Acoust. Sc. Tech., 2016, 27(6), pp. 340-343.

[20] Voong, T. M., Reuter, C., and Oehler, M., “Influence of individual HRTF preference on localization accuracy–a comparison between regular and bone conducting headphones,” In Proceedings of the Audio Engineering Society Convention 148, Vienna, May 2020, Audio Engineering Society.

[21] Stanley, R. M., “Measurement and validation of bone-conduction adjustment functions in virtual 3D audio displays,” Doctoral dissertation, 2009, Georgia Institute of Technology.

[22] Chang-Geun, K. Lee, and P. Spencer, “Effectiveness of Advanced Bone Conduction Earphones for People Who Enjoy Outdoor Activities,” In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2011, 55(1), pp. 1788–1792.

[23] Ziegelwanger, H., Kreuzer, W., and Majdak, P., “Mesh2HRTF: Open-source software package for the numerical calculation of head-related transfer functions,” in Proceedings of the 22nd International Congress on Sound and Vibration, Florence, 2015, IT.

[24] Ziegelwanger, H., Majdak, P., and Kreuzer, W., “Numerical calculation of listener-specific head-related transfer functions and sound localization: Microphone model and mesh discretization,” The Journal of the Acoustical Society of America, 2015, 138, pp. 208-222.

[25] VIVE Pro Eye Overview | VIVE United States. https://www.vive.com/us/product/vive-pro-eye/overview/

[26] Unity Real-Time Development Platform | 3D, 2D VR & AR Engine. https://unity.com//

[27] M. P. Jenny C. and C. Reuter, "SOFA Native Spatializer Plugin for Unity - Exchangeable HRTFs in Virtual Reality," in Proceedings of the 144th Convention of the Audio Engineering Society, Milan, 2018.

[28] Roginska, A., Wakefield, G. H., and Santoro, T. S., "User selected HRTFs: Reduced complexity and improved perception," in Undersea Human System Integration Symposium, Providence, RI, USA, 2010, pp. 1–14.

[29] Shukla, R., Stewart, R., Roginska, A., and Sandler, M., eds. "User Selection of Optiman HRTF Sets via Holistic Comparative Evaluation," International Conference on Audio for Virtual and Augmented Reality 2018, Redmond, WA, USA.

[30] Gardner, W. G., and Martin, K. D., "HRTF measurements of a KEMAR," Journal of the Acoustical Society of America, 1995, 97, pp. 3907-3908.

[31] Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C., "The cipic hrtf database," in Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575), 2001, pp. 99–102.

[32] Warusfel, O., "Listen HRTF Database," May 2003. http://recherche.ircam.fr/equipes/salles/listen/index.html.

[33] Miller, J. D., Godfroy-Cooper, M., and Wenzel, E. M., "Using Published HRTFS with Slab3D: Metric-Based Database Selection and Phenomena Observed," presented at the 20th International Conference on Auditory Display (ICAD), New York, 2014, pp. 22-25.

[34] Cho, S. J., Ovcharenko, A., and Chong, U., "Front-Back Confusion Resolution in 3D Sound Localization with HRTF Databases," in International Forum on Strategic Technology, 2006, pp. 239-243.

[35] Park, M., Choi, S., Kim, S., and Bae, K., "Improvement of front-back sound localization characteristics in headphone-based 3D sound generation," in The 7th International Conference on Advanced Communication Technology - ICACT 2005, pp. 273-276.

[36] Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L., "Localization using non individualized head-related transfer functions," in Journal of the Acoustical Society of America 1993, 94, pp.111-123.

[37] Oberem J., Richter J. G., Setzer D., Seibold J., Koch I., and Fels J., "Experiments on localization accuracy with non-individual and individual HRTFs comparing static and dynamic reproduction methods", bioRxiv, 2020.