Curvature-adaptive mesh grading for numerical approximation of head-related transfer functions

Timon Palm[1], Sebastian Koch[1], Fabian Brinkmann[2], Marc Alexa[1]
[1] *Computer Graphics, Technical University of Berlin, Germany*
[2] *Audio Communication Group, Technical University of Berlin, Germany*
Correspondence: timon.palm@campus.tu-berlin.de, fabian.brinkmann@tu-berlin.de

Abstract
Head-related transfer functions (HRTFs) describe the direction dependent free field sound propagation from a point source to the listener's ears and are an important tool for audio in virtual and augmented reality. Traditionally, HRTFs are measured acoustically with the listener being positioned in the center of a spherical loudspeaker array. Alternatively, they can be approximated by numerical simulation, for example applying the boundary element method (BEM) to a piece-wise linear representation (a surface mesh) of the listener's head. Numerical approximation may be more economical, particularly in combination with methods for the synthesis of head geometry. To decrease the computation time of the BEM, it has been suggested to gradually decrease the resolution of the mesh with increasing distance from the ear for which the HRTF is approximated. We improve this approach by also considering the curvature of the geometry. The resulting graded meshes lead to faster simulation with the same or better accuracy in the HRTF compared to previous work. Our software is available at https://cg-tub.github.io/hrtf_mesh_grading.

Introduction
Head-related transfer functions (HRTFs) describe the directional sound diffraction of an incident sound wave at the entrance of a listener's ear. The HRTF holds binaural and spectral cues that the brain can exploit to estimate the location of sound sources. Using the boundary element method (BEM), the HRTF can be simulated on a geometrical model (mesh) of the head and pinnae by solving the acoustic wave equation. The surface mesh of the head is in that context the scattering object Ω with its surface boundary Γ, Ω_e defining the exterior domain outside the scattering object, and Γ^* defining the area of an idealized microphone at the blocked ear channel entrance. Loudspeakers with position x^* are emitting spherical waves that create the incident sound field $\phi_{inc}(x^*)$. They are commonly arranged in a spherical sampling grid around the surface mesh.

Two types of errors have to be considered in the numerical approximation:

1. **geometrical error**: occurs due to discretization of the surface mesh.
2. **numerical error**: occurs due to the discretization of the sound field on the surface mesh.

To minimize the geometrical error, head meshes can be obtained with a high accuracy and resolution, resulting in 200,000 elements or more. BEM simulations using such meshes, however, would be very time and memory consuming. The number of elements is thus often reduced by means of remeshing, while carefully controlling geometrical and numerical errors. The next sections introduce the investigated remeshing algorithms, detail the BEM simulations, and discuss the resulting simulated HRTFs.

Remeshing Algorithms
Remeshing is commonly realized in a two-step procedure. In the first step, a sizing field is assigned to every vertex $\mathbf{x}_i = [x_i, y_i, z_i]^T$ that defines the desired edge length $L(\mathbf{x}_i)$. In the second step, the following four operations are iterated to adjust the edge length according to the sizing field:

1. **split edge**: Split edges $(\mathbf{x}_1, \mathbf{x}_2)$ if they are longer than $4/5\, min(L(\mathbf{x}_1), L(\mathbf{x}_2))$.
2. **collapse edge**: Collapse edges $(\mathbf{x}_1, \mathbf{x}_2)$ if they are shorter than $4/3\, min(L(\mathbf{x}_1), L(\mathbf{x}_2))$.
3. **vertex valence**: Flip edges if the valence of neighboring vertices exceeds six, i.e., if more than six edges are connected to one vertex.
4. **vertex position**: Adjust the position of vertices by tangential smoothing to approach equilateral triangles.

An illustration of these steps can be found in [1, Fig. 3]. In the simplest case, a uniform remeshing can be applied where all edges in the mesh have the same length. However, this is not very efficient and results in meshes with many elements. The remeshing approaches introduced in the following thus aim at reducing the mesh resolution for elements that are far away from the receiver, i.e., the ear channel for which the HRTF is simulated. Note that they only differ in the way the sizing field is computed. In all cases, the algorithm deliberately violates the rule of having at least six edges per wave length. It has been shown that it has a minor influence on the simulated HRTF [1, 2].

Distance-based Mesh Grading
The purely distance-based mesh grading proposed by Ziegelwanger et al. [1] calculates the sizing field according to the normalized distance between $\mathbf{x}_i$ and the midpoint of Γ^*. This reduced the number of elements to about 10 % and accelerated the computation by about

80 % while keeping the relative error to a reference introduced by the mesh grading below 10%.

Curvature-based Mesh Grading

Dunyach et al. [3] suggested a mesh grading that uses the local curvature to compute the sizing field. This results in meshes with many small triangles and curved regions and few larger triangles in flat regions. Thus creating visually appealing and smooth meshes with a minimum number of elements. This approach is widely used in the field of Computer Graphics, but has not yet been applied to Numerical Acoustics. It increases the edge lengths at regions with low curvature while keeping the geometrical error below a user specified tolerance ε using

$$L(\mathbf{x}_i) = \sqrt{6\varepsilon/\kappa_i - 3\varepsilon^2}, \tag{1}$$

where κ_i the local absolute maximum curvature computed from the mean curvature H and the Gaussain curvature K

$$\kappa_i = H_i + \sqrt{H_i^2 - K_i}\,, \tag{2}$$

with

$$H_i = \frac{1}{2}\|\Delta \mathbf{x}_i\|\,, \tag{3}$$

and

$$K_i = \frac{1}{A_i}\left(2\pi - \sum_{j\in\nu(j)} \theta_j\right)\,. \tag{4}$$

θ are the angles of the incident triangles around vertex $\mathbf{x}_i$ and A_i its Voronoi area. To further constraint the edge length, $L(\mathbf{x}_i)$ is usually clipped to user defined minimum and maximum edge lengths l_{min} and l_{max}.

Hybrid Mesh Grading

We suggest a hybrid mesh grading approach combining distance and curvature-based criteria for computing the sizing field. In particular, it aims at reducing the geometrical error of the ear of interest. We therefore substituted κ_i in Eq. (1) by

$$\tilde{\kappa}_i = \frac{\kappa_i}{\alpha \cdot d_{\mathbf{x}_i}} \tag{5}$$

where $d_{\mathbf{x}_i}$ is the distance from the point x_i to the ear channel entrance and a scaling factor α.

To obtain a robust and computationally efficient estimation of the ear channel entrance, we assume that the interaural center (the midpoint between the two ear channel entrances) is at the origin of coordinates, that the interaural axis (the axis through the centers of the ear channel entrances) coincides with the y-axis of a right-handed Cartesian coordinate system, and that the left ear has a positive y-coordinate (head looking in positive x-direction). Following these assumptions, the x and z coordinates are zero at the ear channel entrances and the y coordinate $\gamma_{l,r}$ was heuristically determined as

$$\gamma_r = \min(\Omega_y) + [\max(\Omega_y) - \min(\Omega_y)] \cdot 0.15$$
$$\gamma_l = \max(\Omega_y) - [\max(\Omega_y) - \min(\Omega_y)] \cdot 0.15$$

	target		actual			$\#N$	$\#E$
	l_{min}	l_{max}	l_{min}	l_{max}	l_{avg}		
uniform	3.5	3.5	2.1	4.7	3.3	18,743	40,939
	3.0	3.0	1.8	4.0	2.8	25,655	54,763
	2.0	2.0	1.2	2.7	1.8	60,231	123,914
distance	0.7	10.0	0.4	10.6	2.8	13,255	29,962
	0.7	9.0	0.4	9.6	2.5	17,333	38,117
	0.7	7.5	0.4	8.0	2.4	21,682	46,817
	0.7	5.9	0.4	6.2	2.1	29,776	63,005
	0.7	4.8	0.4	5.1	1.9	38,610	80,673
	0.7	4.1	0.4	4.2	1.8	47,378	98,209
hybrid	0.5	15.0	0.6	15.7	3.0	13,231	29,913
	0.5	10.0	0.5	12.1	2.7	16,455	36,362
	0.5	6.0	0.5	7.7	2.5	20,545	44,541
	0.5	3.6	0.5	4.8	2.4	28,913	61,278
	0.5	2.7	0.5	3.6	2.1	41,431	86,314
	0.5	2.3	0.5	3.0	1.9	54,636	112,724

Table 1: Parameters used for remeshing and mesh grading. l gives the target and actual edge lengths in mm, $\#N$ the number of vertices and $\#E$ the number of elements. Actual values were averaged across all subjects.

with Ω_y giving the y-coordinate of all mesh elements and l, r denoting the left and right ear, respectively. The distance than reduces to

$$d_{\mathbf{x}_i} = \sqrt{x_i^2 + z_i^2 + \beta(y_i - \gamma_{l,r})^2}\,. \tag{6}$$

We heuristically determined the factors $\beta = 10$ and

$$\alpha = \begin{cases} 7.5, & z_i < \tilde{z} \text{ and } \kappa_i > 100 \\ 1.5, & \text{else} \end{cases}\,, \tag{7}$$

where the case $\alpha = 7.5$ was designed to increase the edge length at regions below the horizontal plane where the meshes have sharp edges—for example because a mesh was cut to remove parts of the torso—using

$$\tilde{z} = \min(\Omega_z) + [\max(\Omega_z) - \min(\Omega_z)] \cdot 0.22\,. \tag{8}$$

To further reduce the number of elements on the opposite ear, we set $L(x_i) = l_{max}$ if $y_i \geq \gamma_r$ (simulating left ear HRTFs) and $y_i \leq \gamma_l$ (simulating right ear HRTFs). l_{max} is a user specified maximum edge length.

Our implementation is based on the PMP library [4].

HRTF Simulation

The grading algorithms were evaluated using the surface meshes of the first twenty subjects from the HUTUBS HRTF database [5, 6]. For each subject, 15 differently remeshed/graded versions were created with parameters according to Table 1. Uniformly remeshed versions were included as baseline meshes. To obtain a high resolution reference, the raw meshes from the HUTUBS database were remeshed with edge lengths of 0.8 mm in the ear region and 1.5 mm outside the ear region (fulfilling the

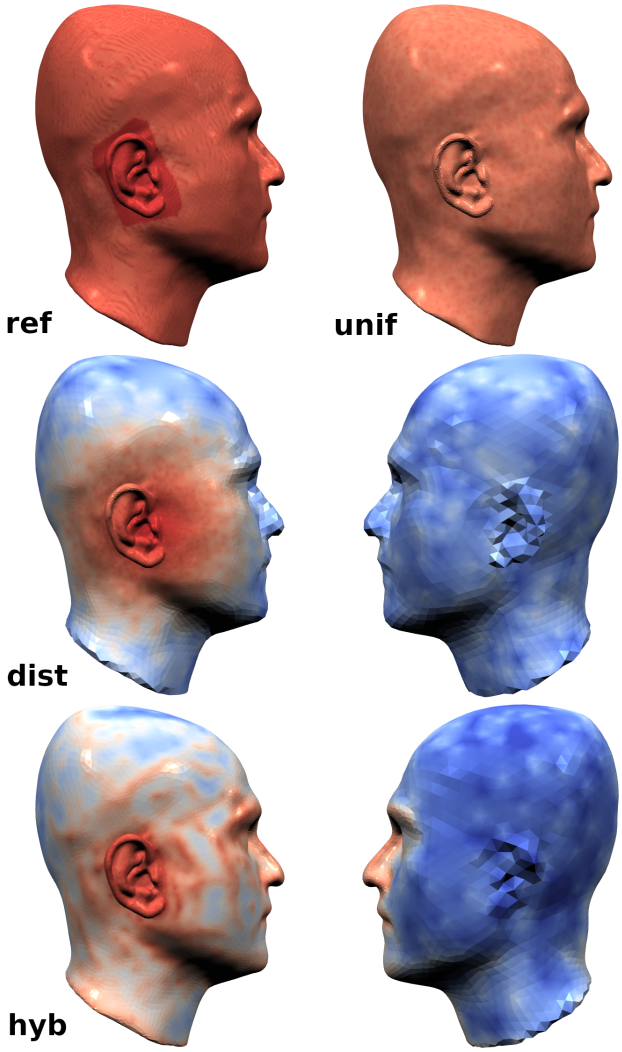

Figure 1: Final meshes with color coded edge lengths (red: small, blue: large): The reference mesh (ref) with ∼133 000 elements, the uniform mesh (unif) with ∼24 000 elements (AEL: 3.0 mm), the distance-based mesh (dist) with ∼19 000 elements (l_{min} = 0.7 mm, l_{max} = 7.5 mm) and our hybrid adaptive HRTF mesh grading (hyb) with ∼20 000 elements (l_{min} = 0.5 mm, l_{max} = 6.0 mm).

criterion of at least six boundary elements per wavelength [2]). The reference meshes contained 65,038 vertices and 133,528 elements on average. Examples for the final meshes are shown in Fig. 1

The BEM calculation for frequencies between 100 and 20,000 Hz in steps of 100 Hz was done using the Multilevel Fast-Multipole (ML-FMM) BEM from Mesh2HRTF [7]. The HRTFs were evaluated at a 1730 points Lebedev grid with a radius of 1.5 m. Acoustic reciprocity was exploited to speed up the measurement by assigning a volume velocity boundary condition to a single element positioned at the entrance of the blocked ear channel.

Error Measure

In analogy to Ziegelwanger et al. [1], the simulated HRTFs were compared to the reference using the error measure

$$e = \frac{\|p_{test} - p_{ref}\|_L}{\|p_{ref}\|_L} \tag{9}$$

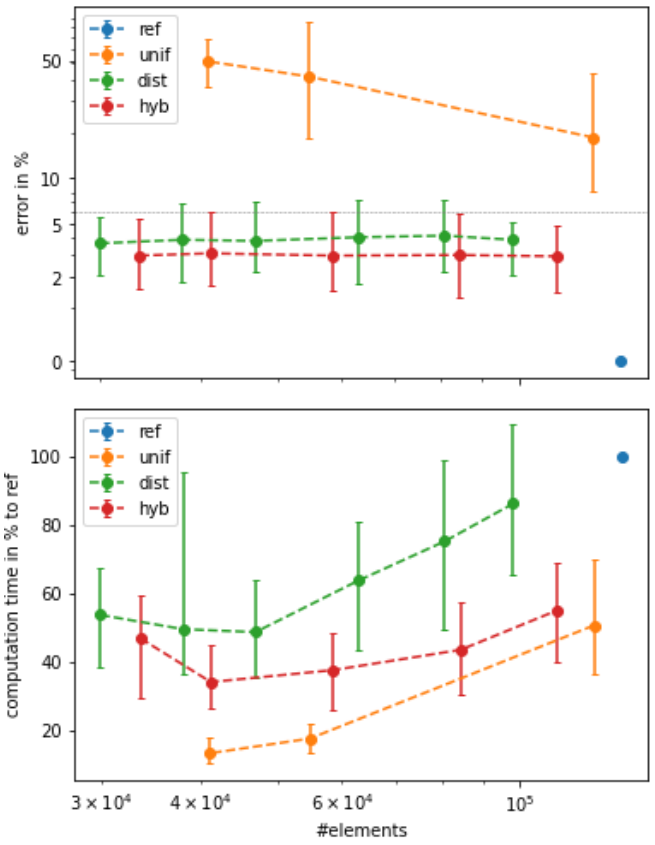

Figure 2: The relative error with L_2-Norm over all 20 subjects: The HRTFs of the adaptive methods had a significantly lower error compared to the uniform remeshing (orange). HRTFs of the curvature adaptive HRTF mesh grading (red) are 1-2% more accurate than the distance-based mesh grading (green).

where p_{test} and p_{ref} denotes the complex-valued sound pressure for the test and reference meshes, respectively. For $|.|_L$, the L_2 norm was used

$$\|f\|_2 = \sqrt{\sum_i \sum_j |x_{ij}|^2 * w_j} \tag{10}$$

where x_{ij} denotes the pressure difference at frequency i and source position j of the Lebedev grid. w_j is a normalized weight with $\sum_j w_j = 1$ for area corrected numerical integration.

Results

Fig. 2 shows the error and computation time in percent relative to the reference. It can be seen that the uniform remeshing is clearly outperformed by the adaptive mesh grading methods, which exhibit relative errors below 6%. (defined as the just noticeable difference by Ziegelwanger et al. [1]). In addition, our proposed hybrid mesh grading produces a 1-2% lower relative error than the distance-based approach on average, which equals to an improvement of about 25%. It seems that both adaptive techniques do not improve with more elements, nor degrade with fewer elements. Figure 3 shows examples of simulated HRTFs to get a better impression of the errors induced by the mesh grading. For our hybrid ap-

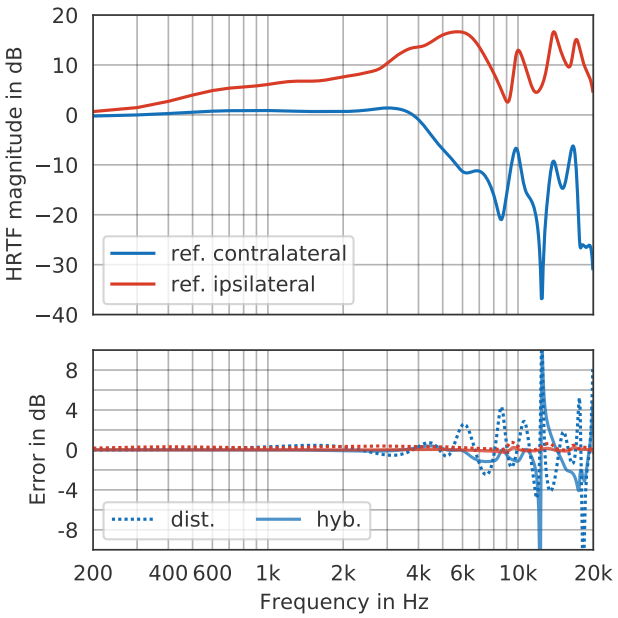

Figure 3: Reference HRTFs for an ipsilateral and contralateral source on the horizontal plane (top) and errors induced by the tested mesh grading approaches (bottom) (dist: min: 0.7 mm max: 7.5 mm, hyb: min: 0.5 mm max: 6.0 mm).

proach, errors above 1 dB occur only at positions where the HRTF magnitude is below -10 dB. While, the average error is smaller for the hybrid approach, errors for both approaches can be considered to be perceptually negligible if considering the resolution of the auditory system of about a third octave.

The computation times, needed for the HRTF, show an improvement of about 10 to 30% for the hybrid approach compared to the distance-based mesh grading. However, both are slower than the uniform remeshing. As expected, the computation time increases with the number of elements with one exception for meshes with about 30,000 elements. This is a side effect of the ML-FMM acceleration, which is optimized for uniform cluster sizes. The computation took 3.29 hours on average for calculating the reference using 8 parallel processes on an 8 core Intel Xeon E5-2637 with 64 GB RAM. The fastest simulations of graded meshes took 1.6 hours (distance-based) and 1.23 hours (hybrid) on average.

Conclusion

We propose a hybrid mesh grading algorithm based on local mesh curvature and distance to the ear. We compare our implementation to uniform remeshing and distance-based mesh grading and publish it as MIT lincesed open source software. HRTFs were simulated using twenty human head meshes and the evaluation of the HRTFs showed that our hybrid mesh grading could further decrease the computation time and increase the accuracy in comparison to the well established distance-based mesh grading. In our study about 40 000 elements (l_{min} = 0.5 mm, l_{max} = 10 mm) were the best parameters to obtain accurate simulations and a low computation time.

Future work could investigate parameters for automatically generating meshes that are valid for different frequency ranges to further speed up the computation. Moreover, it would be interesting to test in how far the proposed algorithm can be applied to head and torso meshes. First informal tests suggested that this might be the case.

References

[1] H. Ziegelwanger, W. Kreuzer, and P. Majdak. “A priori mesh grading for the numerical calculation of the head-related transfer functions”. *Applied Acoustics*, 114 (Dec. 2016), pp. 99-–110. doi: 10.1016/j.apacoust.2016.07.005.

[2] S. Marburg. “Six boundary elements per wavelength: Is that enough?”. *Journal of Computational Acoustics* 10.01 (Mar. 2002), pp. 25—51. doi: 10.1142/S0218396X02001401.

[3] M. Dunyach, D. Vanderhaeghe, L. Barte, and M.Botsch. “Adaptive Remeshing for Real-Time Mesh Deformation”. *Eurographics Short Papers*, (2013), pp. 29–32. doi: 10.2312/conf/EG2013/short/029-032.

[4] D. Sieger and M. Botsch. “The Polygon Mesh Processing Library”, (2020), *http://www.pmp-library.org*

[5] F. Brinkmann, A. Lindau, S. Weinzierl, S. van de Par, M. Müller-Trapet, R. Opdam, and M. Vorländer. “A High Resolution and Full-Spherical Head-Related Transfer Function Database for Different Head-Above-Torso Orientations”. en. *J. Audio Eng. Soc.* 65.10 (Oct. 2017), S. 841-–848. doi: 10.17743/jaes.2017.0033.

[6] F. Brinkmann, M. Dinakaran, R. Pelzer, J. J. Wohlgemut, F. Seipel, D. Voss, P. Grosche, and S. Weinzierl. *The HUTUBS head-related transfer function (HRTF) database*. 2019. doi: 10.14279/depositonce-8487.

[7] H. Ziegelwanger, W. Kreuzer, and P. Majdak. “Mesh2HRTF: An open-source software package for the numerical calculation of head-related transfer functions”. *22nd Int. Congress Sound and Vibration*, (2015), Florence, Italy.

[8] M. Botsch and L. Kobbelt. “A remeshing approach to multiresolution modeling”. en. *Proceedings of the 2004 Eurographics/ACM SIGGRAPHsymposium on Geometry processing - SGP ’04.*, France: ACM Press, 2004, pp. 185-192. doi: 10.1145/1057432.1057457.