Head-related transfer function interpolation in azimuth, elevation, and distance

Hannes Gamper
Department of Media Technology, Aalto University School of Science, FI-00076 Aalto, Finland
hannes.gamper@aalto.fi

Abstract: Although distance-dependent head-related transfer function (HRTF) databases provide interesting possibilities, e.g., for rendering virtual sounds in the near-field, there is a lack of algorithms and tools to make use of them. Here, a framework is proposed for interpolating HRTF measurements in 3-D (i.e., azimuth, elevation, and distance) using tetrahedral interpolation with barycentric weights. For interpolation, a tetrahedral mesh is generated via Delaunay triangulation and searched via an adjacency walk, making the framework robust with respect to irregularly positioned HRTF measurements and computationally efficient. An objective evaluation of the proposed framework indicates good accordance between measured and interpolated near-field HRTFs.

PACS numbers: 43.66.Pn [QJF]
Date Received: August 12, 2013 **Date Accepted:** October 22, 2013

1. Introduction

The head-related transfer function (HRTF) describes the filtering that sound undergoes when traveling from a sound source to the ears of the listener, due to shadowing and reflections caused by the listener's torso, head, and pinnae. HRTFs are a function of the sound source's azimuth and elevation with respect to the listener's head. For distances greater than 1 m, HRTFs can be considered distance-independent.[1,2] Therefore, HRTFs are typically measured at a fixed distance of 1–2 m from the center of the test subject's head, over a range of azimuth and elevation angles.[2] The region further than 1 m away from the center of the head is referred to as the *far-field*, whereas the *near-field* denotes the region within 1 m.[2]

A virtual source can be spatialized by filtering its sound signal with a measured HRTF. If no HRTF measurement is available for the desired source position, interpolation can be used to obtain an HRTF estimate from HRTFs measured at nearby positions. A number of approaches have been proposed previously to interpolate HRTF measurements in 2-D (i.e., azimuth and elevation), obtained at a fixed distance.[3–5]

For virtual sources in the near-field, experiments have shown the distance cues encoded into near-field HRTFs to improve the localization performance of listeners.[1] Near-field virtual sources may be used, for example, to signify urgency or increase immersion.[1] To account for the distance-dependence of HRTFs in the near-field, approaches have been proposed to estimate near-field HRTFs from measurements taken at a single distance.[2,6–9] Recently, HRTF databases have been published containing measurements taken at various distances in the near-field.[10–12] However, there is a lack of algorithms and tools to make use of such databases.[1] Specifically, to the best of the author's knowledge, no HRTF interpolation framework has been proposed in the literature that allows the direct interpolation of distance-dependent HRTF measurements without relying on a particular layout of the HRTF measurement positions.

Lentz *et al.*[13] proposed to interpolate two HRTFs to arrive at an HRTF estimate between two measured distances. Based on the PKU&IOA HRTF database,[10] Villegas and Cohen[14] proposed a virtual panning tool that interpolates measurements obtained at various distances. The interpolation is performed by determining the eight measurement points forming a volume enclosing the desired source position. Both 3-D interpolation methods mentioned above are *ad hoc* methods tailored for a specific HRTF database with a regular measurement grid structure. However, the assumption of a regular grid structure may be violated in practice. When measuring HRTFs on a spherical grid, the measurement points are typically distributed more sparsely toward the poles than at the equator, in accordance with the decreasing localization accuracy of humans toward extreme elevations. Other potential causes for irregularities in the HRTF measurement grid are movements of human subjects during HRTF measurements[12] and positioning errors of the mechanical measurement setup.

Here, a 3-D HRTF interpolation framework is proposed that (i) does not rely on a regular HRTF measurement grid, (ii) allows the direct interpolation of HRTF measurements taken at various distances, and (iii) is computationally efficient, without resorting to *ad hoc* methods. The contribution of this article lies in the use of a standard triangulation method to efficiently group HRTF measurements into sets for interpolation, the use of a fast search algorithm to select a suitable HRTF set, and the use of tetrahedral interpolation with barycentric weights to interpolate the selected HRTF set. Furthermore, an objective evaluation of the performance of the proposed 3-D interpolation framework is presented.

2. Proposed method

Given a set of HRTFs measured at a fixed distance, an interpolated HRTF can be obtained from three measurement points forming a triangle enclosing the desired source *direction*.[15,16] The present work extends this approach to include the desired source distance through direct interpolation of HRTF measurements obtained at various distances. The proposed approach is based on finding four measurement points forming a tetrahedron that encloses the desired source *position* [see Fig. 1(b)]. Similarly to previously proposed methods for interpolating HRTF measurement points lying on the surface of a sphere, the tetrahedral interpolation proposed here is based on the assumption that an HRTF estimate for a desired source position can be obtained by interpolating nearby HRTF measurements. Next, the process of grouping HRTF measurement points into non-overlapping tetrahedra is discussed.

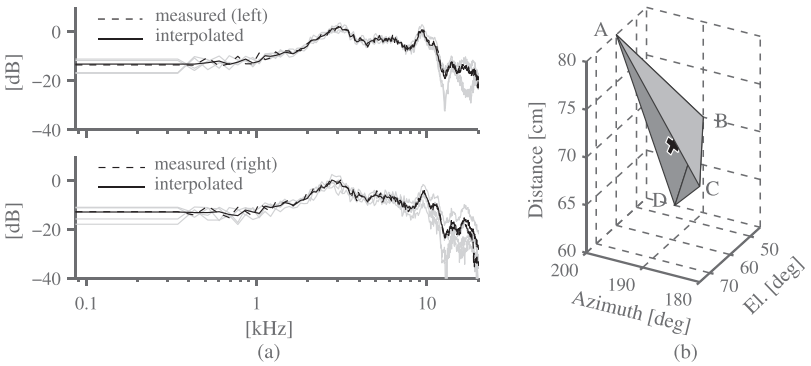

Fig. 1. (a) Magnitude spectra of measured (dashed line) and estimated (solid line) HRTF via tetrahedral interpolation of measured HRTFs (light lines); (b) desired source position × and HRTF measurement positions forming tetrahedron used for interpolation. The measurements are taken from the database by Yu *et al.* (Ref. 11).

2.1 Triangulation of measurement points

A set of points in 2-D can be grouped into non-overlapping triangles via triangulation. When using triangles for interpolation, it is desirable that they be nearly equiangular. The Delaunay triangulation is optimal in this sense, and it maximizes the minimum angle of the generated triangles.[17] For points lying on a plane, the Delaunay triangulation generates triangles such that the circumcircle of each triangle contains no other points.[17] This concept is applicable to irregularly spaced points as well as to higher dimensions. In 3-D, the Delaunay triangulation yields tetrahedra such that the circumsphere of each tetrahedron contains no other points.

Efficient algorithms exist to perform Delaunay triangulation in 2-D and 3-D.[17] Figure 2 illustrates the tetrahedral mesh generated via Delaunay triangulation of a set of measurement points for (i) an HRTF database with a highly regular measurement grid, and (ii) a database with grid irregularities. The mesh consists of non-overlapping tetrahedra that fill the space occupied by the measurement grid. Any point inside that space is enclosed by exactly one tetrahedron, except if the point lies on a vertex, edge, or facet shared by multiple tetrahedra.

2.2 Calculation of interpolation weights

Once a tetrahedral mesh of HRTF measurement points has been generated via triangulation, an HRTF estimate for any point $\boldsymbol{X}$ lying inside the mesh can be obtained by interpolating the vertices of the tetrahedron enclosing $\boldsymbol{X}$. Consider a tetrahedron formed by the vertices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ as depicted in Fig. 1(b). Any point $\boldsymbol{X}$ inside the tetrahedron can be represented as a linear combination of the vertices

$$\boldsymbol{X} = g_1\boldsymbol{A} + g_2\boldsymbol{B} + g_3\boldsymbol{C} + g_4\boldsymbol{D}, \tag{1}$$

where g_i are scalar weights. With the additional constraint

$$\sum_{i=1}^{4} g_i = 1, \tag{2}$$

the weights g_i are the *barycentric coordinates* of point $\boldsymbol{X}$.[18] The barycentric coordinates can directly be used as interpolation weights for estimating the HRTF $\hat{\boldsymbol{H}}_{\mathrm{x}}$ at point $\boldsymbol{X}$ as the weighted sum of the HRTFs, $\boldsymbol{H}_i$, measured at $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$, respectively,

$$\hat{\boldsymbol{H}}_{\mathrm{x}} = \sum_{i=1}^{4} g_i \boldsymbol{H}_i. \tag{3}$$

Subtracting $\boldsymbol{D}$ from both sides of Eq. (1) yields

$$\boldsymbol{X} - \boldsymbol{D} = [\,g_1 \quad g_2 \quad g_3\,]\boldsymbol{T}, \tag{4}$$

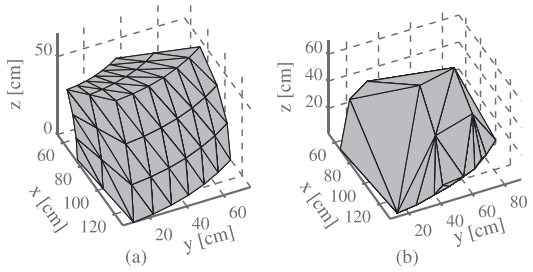

Fig. 2. Delaunay triangulation of measurement points in (a) the PKU&IOA HRTF database (Ref. 10) and (b) the HRTF database by Bolaños and Pulkki (Ref. 12).

where

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{A} - \boldsymbol{D} \\ \boldsymbol{B} - \boldsymbol{D} \\ \boldsymbol{C} - \boldsymbol{D} \end{bmatrix}. \tag{5}$$

Given a desired source position, $\boldsymbol{X}$, the barycentric interpolation weights are found by evaluating

$$[\, g_1 \quad g_2 \quad g_3 \,] = (\boldsymbol{X} - \boldsymbol{D})\boldsymbol{T}^{-1}, \tag{6}$$

and with Eq. (2),

$$g_4 = 1 - g_1 - g_2 - g_3. \tag{7}$$

Note that $\boldsymbol{T}$ depends solely on the geometry of the tetrahedron and is independent of the desired source position, $\boldsymbol{X}$. Therefore, $\boldsymbol{T}^{-1}$ can be pre-calculated for all tetrahedra during initialization and stored in memory. This reduces the operational count for finding the interpolation weights via Eqs. (6) and (7) to 12 additions and 9 multiplications per tetrahedron.

Barycentric weights are well-suited for interpolation:

(i) For a point lying inside a tetrahedron, the barycentric weights, g_i, are positive: $0 < g_i < 1$.
(ii) For a point moving inside a tetrahedron, the weights change smoothly as a function of the vertex-distance.[18]
(iii) For a point lying on a vertex $\boldsymbol{A}$, the barycentric weights are 1 at $\boldsymbol{A}$ and 0 otherwise; hence, the interpolation at $\boldsymbol{A}$ is exact.
(iv) For a point lying on an edge or facet of the tetrahedron, only the vertices forming that edge or facet have nonzero barycentric weights. Furthermore, the vertex weights for all tetrahedra sharing that edge or facet are identical.

The above properties are particularly advantageous for the display of moving virtual sources, as the interpolation via barycentric weights does not cause discontinuities in the interpolated HRTFs. For a source moving smoothly from one tetrahedron to another across a shared vertex, edge, or facet, the HRTF estimate changes smoothly, including at the crossing point. Another interesting observation is that for points lying on the boundary of the tetrahedral mesh, the tetrahedral interpolation reduces to a triangular interpolation [the gains in this case are actually equal to the gains used in vector-base amplitude panning[19]]. Therefore, virtual sources in the far-field can be interpolated via projection onto the mesh boundary and appropriate gain adjustment. However, a detailed analysis of this approach is outside the scope of this paper. Figure 1(a) illustrates the HRTF interpolation at a source position, ×, using barycentric weighting of the HRTF measurements at the vertices of the enclosing tetrahedron [see Fig. 1(b)].

2.3 Selecting a tetrahedron for interpolation

Given a tetrahedral mesh of HRTF measurements obtained via triangulation and a desired source position, $\boldsymbol{X}$, a tetrahedron suitable for interpolation can be found by evaluating the barycentric coordinates of the tetrahedra; $\boldsymbol{X}$ lies inside a tetrahedron if and only if all barycentric coordinates are positive. Therefore, a straightforward way to find a suitable tetrahedron for interpolation is to iterate through all tetrahedra in the mesh until one is found that satisfies this condition. The running time of this "brute-force" approach increases linearly with the number of iterations needed to find a suitable tetrahedron. An example run of this brute-force search is shown in Fig. 3(a).

Given the large number of tetrahedra generated for dense HRTF measurement grids, and the tight processing time constraints of real-time audio applications, it is desirable to speed up the process of selecting a tetrahedron for interpolation. A more efficient way to locate a point in a triangulation is via an *adjacency walk*.[18] Starting from a

random tetrahedron, evaluate the barycentric coordinates and walk to the adjacent tetrahedron across the triangle formed by the vertices with the three largest barycentric coordinates; terminate when all barycentric coordinates are positive [see Fig. 3(b), light gray tetrahedra]. The theoretical complexity of the adjacency walk for non-homogenous meshes is $O((n)^{1/3})$.[18] This constitutes a substantial improvement in terms of scalability over the $O(n)$ brute-force approach. As shown in Fig. 3(c), the worst-case performance of the adjacency walk is well below 0.1 ms even for the largest tested database (31 752 tetrahedra). To reduce the number of steps needed for the adjacency walk to terminate, a tetrahedron close to the desired source position, $\boldsymbol{X}$, can be chosen as the starting point for the walk. A simple yet efficient way to find the closest neighbors to a point in 3-D is by querying an *octree* representation of the HRTF measurement points.[20] A cuboid containing all points forms the root of the octree. Starting from the root cuboid, the octree is generated by recursively dividing every cuboid into eight equal-sized cuboids. The subdivision of a cuboid stops when it contains at most N points, making it a leaf of the octree. N is chosen to yield the desired spatial resolution of the octree. To find a tetrahedron close to a desired source position, $\boldsymbol{X}$, the octree is searched for the leaf cuboid enclosing $\boldsymbol{X}$. A tetrahedron with a vertex contained in that leaf cuboid lies close to $\boldsymbol{X}$, and can be used as a starting point for the adjacency walk, thus reducing the iterations needed for the walk to terminate [see Fig. 3(b), dark gray] as well as the running time of the selection algorithm [see Fig. 3(c)].

2.4 Pre-processing and initialization

To minimize computational load at run-time, some pre-processing is done during initialization. First, the Delaunay triangulation is performed and the resulting tetrahedral mesh is stored. For each tetrahedron, the inverse $\boldsymbol{T}^{-1}$ in Eq. (6) is pre-calculated to speed up the calculation of barycentric weights, and an adjacency map is created that lists the adjacent tetrahedra in the mesh to enable fast tetrahedron selection via the adjacency walk algorithm. Finally, an octree representation of the measurement points is generated.

3. Evaluation

To evaluate the performance of the proposed HRTF interpolation framework, experiments are carried out on distance-dependent HRTF measurement databases of the KEMAR mannequin: the PKU&IOA database[10] and the database by Yu *et al.*[11] For the database by Yu *et al.*,[11] the triangulation of the measurement points is shown in Fig. 4(a), along with plots of the magnitude spectra of the measured HRTFs. To test the robustness of the proposed tetrahedral interpolation against irregularities in the measurement grid, a reduced HRTF database with an irregular measurement grid is constructed from both databases by randomly removing measurement points. Using the reduced

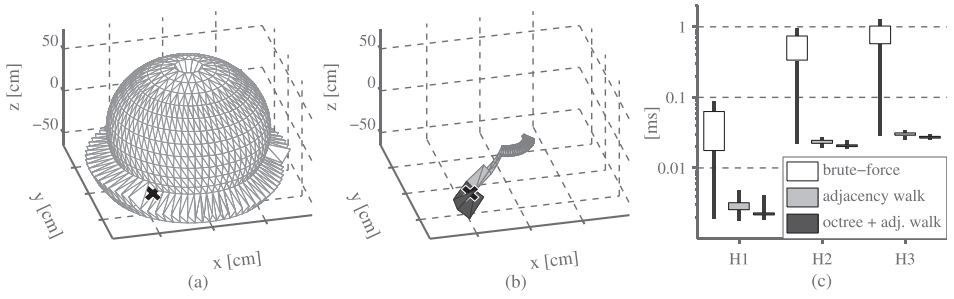

Fig. 3. Tetrahedron selection for source position × via (a) brute-force search (18 772 iterations) and (b) adjacency walk (Ref. 18) with a random starting tetrahedron (light gray, 105 iterations) and with a tetrahedron selected via an octree query (Ref. 20; dark gray, five iterations); (c) running-times for 1000 random source positions averaged over 100 repetitions on a computer with a 2 GHz quad-core processor for three HRTF databases [H1 (Ref. 12), H2 (Ref. 11), H3 (Ref. 10)]; vertical lines extend from minimum to maximum, boxes from lower to upper quartile.

databases, HRTF estimates at the removed points are obtained via the proposed tetrahedral interpolation and compared to the actual HRTF measurements in the original database. This allows to objectively evaluate the performance of the interpolation algorithm.

The HRTF interpolation is performed on the magnitude spectra of the measured HRTFs, as proposed by Zotkin *et al.*[21] The phase of the interpolated HRTF can be derived from a spherical head model[21] independent of source distance,[1] and is not considered here. Figure 1 illustrates the interpolation at a removed measurement point ×. The triangulation of the database by Yu *et al.*[11] with 50% and 90% of measurement points removed is depicted in Fig. 4(a) (top). As can be seen, the triangulation copes well with the irregularities in the reduced measurement grids. Figure 4(a) (bottom) shows the interpolated HRTF magnitude spectra with 50% and 90% of measurement points removed. With 50% of HRTF measurements removed, the tetrahedral interpolation is able to reproduce the main spectral features of the measured HRTFs. Unsurprisingly, some of those features are lost with 90% of measurement points removed. To objectively evaluate the performance of the interpolation framework, the root-mean square error (RMSE) between interpolated and measured HRTF magnitudes is calculated over third-octave bands with center frequencies from 500 Hz to 16 kHz. Figure 4(b) depicts the RMSE as a function of the percentage of points removed. The RMSE is calculated over ten repetitions of the random removal for each percentage. The 0% condition is obtained by removing a single random point from the database. This is repeated 100 times to obtain a representative RMSE estimate. The 0% condition serves as a control; with a single point removed, in most cases, the tetrahedral interpolation reduces to a simple triangular or linear interpolation of neighboring measurement points. As expected, the RMSE increases with the percentage of points removed.

4. Summary and conclusion

A framework for the 3-D interpolation of near-field HRTF measurements obtained at various distances was proposed. The main steps involved are (i) Delaunay triangulation of measurement points to obtain a tetrahedral mesh, (ii) selection of a tetrahedron enclosing the desired position, and (iii) interpolation of the HRTF measurements at the vertices of the enclosing tetrahedron using barycentric weights. An objective evaluation shows

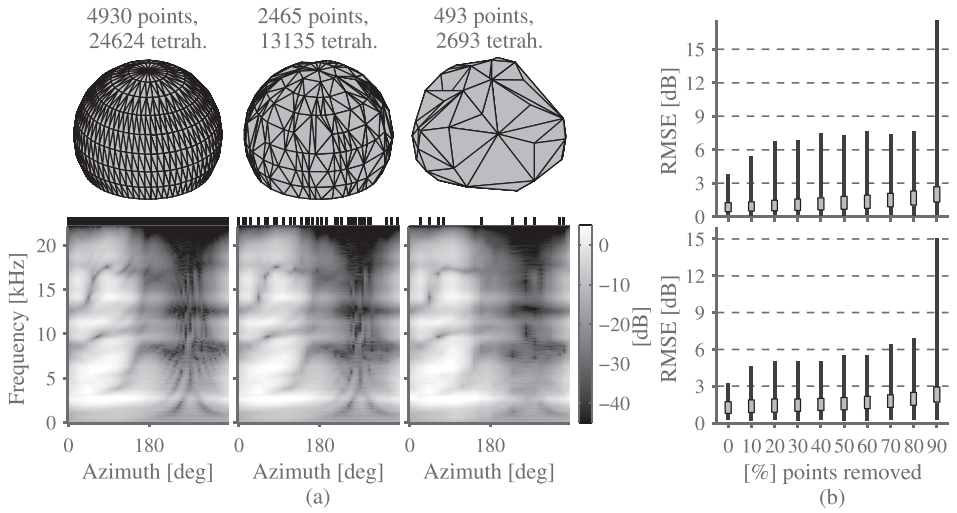

Fig. 4. (a) Triangulation (top) and HRTF magnitude spectra (bottom) of the HRTF database by Yu *et al.* (Ref. 11) at 0 degrees elevation, 50 cm distance, using all measurement points and with 50% and 90% of points randomly removed. Ticks mark magnitude spectra obtained directly from measured HRTFs (i.e., without interpolation). (b) RMSE for Qu *et al.* (Ref. 10; top) and Yu *et al.* (Ref. 11; bottom), as a function of the percentage of points removed; vertical lines extend from minimum to maximum, boxes from lower to upper quartile.

good accordance between interpolated and measured HRTFs in the near-field. The proposed framework proved well-suited for the (real-time) interpolation of HRTF databases containing measurements at several distances and exhibiting an irregular measurement grid, thus enabling the spatialization of virtual sources in the near-field. A MATLAB® demonstration of the proposed interpolation framework is available online.[22]

Acknowledgments

This work was supported by the Helsinki Graduate School in Computer Science and Engineering (HeCSE), the MIDE program of Aalto University, the Nokia Research Foundation, and Tekniikan edistämissäätiö (TES). The author wishes to thank Jonathan Botts for helpful comments.

References and links

[1]D. S. Brungart, "Near-field virtual audio displays," Presence: Teleoper. Virtual Environ. **11**, 93–106 (2002).
[2]A. Kan, C. Jin, and A. van Schaik, "A psychophysical evaluation of near-field head-related transfer functions synthesized using a distance variation function," J. Acoust. Soc. Am. **125**, 2233–2242 (2009).
[3]E. Wenzel and S. Foster, "Perceptual consequences of interpolating head-related transfer functions during spatial synthesis," in *Proc. IEEEWASPAA*, New Paltz, NY, 1993, pp. 102–105.
[4]K. Hartung, J. Braasch, and S. J. Sterbing, "Comparison of different methods for the interpolation of head-related transfer functions," in *Proc. 16th Conf. Audio Eng. Soc.*, Helsinki, Finland, 1999, pp. 319–329.
[5]G. Ramos and M. Cobos, "Parametric head-related transfer function modeling and interpolation for cost-efficient binaural sound applications," J. Acoust. Soc. Am. **134**, 1735–1738 (2013).
[6]R. Duraiswami, D. N. Zotkin, and N. A. Gumerov, "Interpolation and range extrapolation of HRTFs," in *Proc. IEEE ICASSP*, Montreal, Canada, 2004, pp. 45–48.
[7]D. Menzies and M. Al-Akaidi, "Nearfield binaural synthesis and ambisonics," J. Acoust. Soc. Am. **121**, 1559–1563 (2007).
[8]D. Romblom and B. Cook, "Near-field compensation for HRTF processing," in *Proc. Conv. Audio Eng. Soc.*, San Francisco, CA, 2008.
[9]S. Spors and J. Ahrens, "Efficient range extrapolation of head-related impulse responses by wave field synthesis techniques," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 49–52.
[10]T. Qu, Z. Xiao, M. Gong, Y. Huang, X. Li, and X. Wu, "Distance-dependent head-related transfer functions measured with high spatial resolution using a spark gap," IEEE Trans. Audio, Speech, Lang. Process. **17**, 1124–1132 (2009).
[11]G.-Z. Yu, B.-S. Xie, and D. Rao, "Characteristics of near-field head-related transfer function for KEMAR," in *Proc. Conf. Audio Eng. Soc.*, Tokyo, Japan, 2010.
[12]J. G. Bolaños and V. Pulkki, "HRIR database with measured actual source direction data," in *Proc. Conv. Audio Eng. Soc.*, San Francisco, CA, 2012.
[13]T. Lentz, I. Assenmacher, M. Vorländer, and T. Kuhlen, "Precise near-to-head acoustics with binaural synthesis," J. Virtual Reality Broadcast. **3**, 1–12 (2006).
[14]J. Villegas and M. Cohen, "HRIR∼: Modulating range in headphone-reproduced spatial audio," in *Proc. ACMSIGGRAPH Conf. Virtual-Reality Continuum and Its Applications in Industry* (ACM, New York, 2010), pp. 89–94.
[15]F. P. Freeland, L. W. P. Biscainho, and P. S. R. Diniz, "Interpositional transfer function of 3D-sound generation," J. Audio Eng. Soc. **52**, 915–930 (2004).
[16]M. Queiroz and G. H. M. A. de Sousa, "Efficient binaural rendering of moving sound sources using HRTF interpolation," J. New Music Res. **40**, 239–252 (2011).
[17]F. Aurenhammer, "Voronoi diagrams—A survey of a fundamental geometric data structure," ACM Comput. Surv. **23**, 345–405 (1991).
[18]R. Sundareswara and P. Schrater, "Extensible point location algorithm," in *Proc. Int. Conf. Geometric Modeling and Graphics*, London, UK, 2003, pp. 84–89.
[19]V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," J. Audio Eng. Soc. **45**, 456–466 (1997).
[20]H. Samet, "Implementing ray tracing with octrees and neighbor finding," Comput. Graphics **13**, 445–460 (1989).
[21]D. Zotkin, R. Duraiswami, and L. Davis, "Rendering localized spatial audio in a virtual auditory space," IEEE Trans. Multimedia **6**, 553–564 (2004).
[22]H. Gamper, "3-D HRTF interpolation," http://www.mathworks.com/matlabcentral/fileexchange/43809 (Last viewed October 31, 2013).