SELECTION AND INTERPOLATION OF HEAD-RELATED TRANSFER FUNCTIONS FOR RENDERING MOVING VIRTUAL SOUND SOURCES

*Hannes Gamper**

Dept. of Media Technology
Aalto University
Finland
Hannes.Gamper@aalto.fi

ABSTRACT

A variety of approaches have been proposed previously to interpolate head-related transfer functions (HRTFs). However, relatively little attention has been given to the way a suitable set of HRTFs is chosen for interpolation and to the calculation of the interpolation weights. This paper presents an efficient and robust way to select a minimal set of HRTFs and to calculate appropriate weights for interpolation. The proposed method is based on grouping HRTF measurement points into non-overlapping triangles on the surface of a sphere by calculating the convex hull. The resulting Delaunay triangulation maximises minimum angles. For interpolation, the HRTF triangle that is intersected by the desired sound source vector is selected. The selection is based on a point-in-triangle test than can be performed using just 9 multiplications and 6 additions per triangle. A further improvement of the selection process is achieved by sorting the HRTF triangles according to their distance from the sound source vector prior to performing the point-in-triangle tests. The HRTFs of the selected triangle are interpolated using weights derived from vector-base amplitude panning, with appropriate normalisation. The proposed method is compared to state-of-the-art methods. It is shown to be robust with respect to irregularities in the HRTF measurement grid and to be well-suited for rendering moving virtual sources.

1. INTRODUCTION

The head-related transfer function (HRTF) describes the filtering that sound travelling from a sound source to the ears of a listener undergoes due to shadowing and reflections from the listener's torso, head, and pinnae. The HRTF changes as a function of azimuth and elevation of the sound source and thus provides cues for the human auditory system to determine the source direction. To render a virtual source emanating sound from a certain direction, the directional cues corresponding to the desired azimuth and elevation need to be encoded in the sound. The encoding can be done by filtering the sound with an HRTF corresponding to the desired direction. Some databases of measured HRTFs available online include the MIT KEMAR database [1], the CIPIC database [2], and the LISTEN database [3]. These databases contain a large number of measurement points covering a range of azimuth and elevation angles. To display a virtual sound source at a direction for which no measured HRTF is available, an appropriate HRTF needs to be estimated from available HRTF measurements. A common approach to obtain an HRTF estimate for an arbitrary azimuth and elevation angle from an HRTF measurement database is to interpolate a subset of the measured HRTFs. Although various ways to interpolate HRTF subsets have been proposed previously, relatively little attention has been given to the way in which the HRTF subset is selected.

Here, a robust and efficient way to select a minimal subset of an HRTF measurement database for interpolation is proposed. The subset is minimal in the sense that it contains the minimum number of measured HRTFs that allow rendering a moving virtual source without abrupt spectral changes. The interpolation of a minimal subset of HRTFs can be performed as a linear combination of the selected HRTF subset with appropriate weights, in the time or frequency domain. Here, a simple method to calculate interpolation weights is proposed and compared to methods proposed previously.

2. PRIOR WORK

2.1. Subset selection

Approaches to HRTF interpolation have been proposed that take into account all or a large number of measurement points, for example using spherical splines [4] or via rational state-space interpolation [5]. However, while these approaches potentially yield more accurate HRTF estimates than methods that use only a small subset of HRTF measurements for interpolation, they come at an increased computational cost. The computational complexity of the HRTF interpolation algorithm becomes an issue when rendering moving virtual sources, especially if multiple sources and/or room reflections are to be rendered simultaneously, or if the rendering is done on a device with reduced computational power, for example a mobile phone.

To minimise computational complexity, approaches have been proposed to select a minimal subset of HRTF measurements for interpolation. The reasoning behind such approaches is that an HRTF estimate for a desired, non-measured direction can be obtained by interpolating HRTFs measured at directions close to the desired direction. Figure 1 illustrates various HRTF subset selection approaches for a given source direction. One subset selection approach is to select the measurement points nearest to the desired source direction. Proposed methods include finding the 3 nearest measurement points [6, 7] and 4 nearest measurement points [8, 4] (see fig. 1, N3 and N4). Distance measures proposed to determine the nearest neighbouring measurement points include the great-circle distance [4] and the Euclidean distance [7]. Another subset

* 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). Acknowledgements to Jonathan Botts for fruitful discussions.

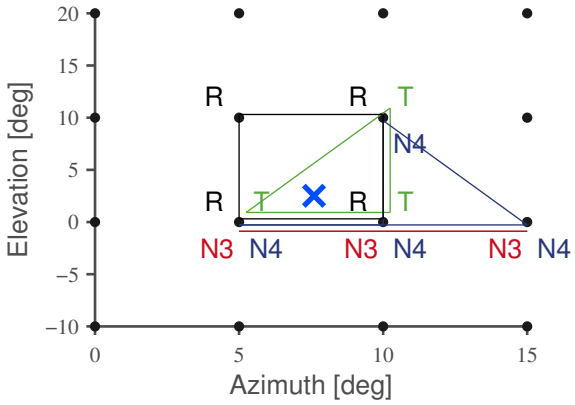

Figure 1: *Subset selection approaches: nearest-3 (N3), nearest-4 (N4), enclosing triangle (T) and enclosing rectangle (R). The dots denote measurement points in the MIT KEMAR HRTF database [1], the cross denotes the desired source direction.*

selection approach is to select the nearest measurement points that enclose the desired source direction. Proposed approaches include finding a rectangle [9, 10] or triangle [11, 12] enclosing the desired direction (see fig. 1, R and T).

2.2. Calculation of interpolation weights

A straightforward way to interpolate a set of HRTF measurements is to calculate a weighted average of the selected HRTFs. Bilinear interpolation can be used to calculate weights to interpolate measured HRTFs linearly with respect to azimuth and elevation. Bilinear interpolation has been proposed for interpolating 3 measurement points [11], and 4 measurement points arranged in a regular grid [9]. Geometric approaches calculate weights based on the distance of the desired direction from each measurement point. For the interpolation of 3 or more measurement points, weights can be calculated from the inverse of the Euclidean distance [7] or the great-circle distance (i.e., the distance along the sphere) [4, 13]. The interpolation of measured HRTFs can be interpreted as a superposition of the signals of virtual loudspeakers positioned at the measurement points [12]. With this interpretation, the interpolation weights are equivalent to the panning gains of these virtual loudspeakers. Panning gains for arbitrary loudspeaker setups can be calculated using vector-base amplitude panning (VBAP) [14].

3. PROPOSED METHOD

3.1. Vector formulation

An HRTF database contains HRTFs measured at various directions with respect to the measured test subject. The goal of HRTF interpolation is to obtain an HRTF estimate corresponding to a desired source position. HRTFs can be considered distance-independent for distances greater than 1 m [15]. Therefore, assuming the desired source is further than 1 m away, the measured directions and the desired source position can be represented as unit vectors, denoted as measurement vectors $\mathbf{h}_i$ and source vector $\mathbf{s}$, respectively. Given this vector formulation, the proposed methods for subset selection and HRTF interpolation assume that an HRTF estimate for the desired source direction $\mathbf{s}$ can be obtained as a linear combination of 3 HRTF measurements at $\mathbf{h}_1$, $\mathbf{h}_2$, and $\mathbf{h}_3$ that form a curved triangle on the unit sphere and enclose $\mathbf{s}$. This formulation is analogous to the vector-base formulation in VBAP [14].

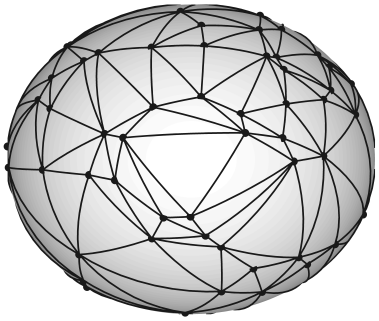

Figure 2: *100 random HRTF measurement points on the unit sphere and the triangulation of their convex hull.*

3.2. Algorithm initialisation

Processing time requirements can be tight when rendering moving virtual sound sources. Therefore, it is a good idea to offload as much processing as possible to the initialisation of the algorithm. Goal of the subset selection algorithm is to find a measurement triplet $\{\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3\}$ that encloses the desired source direction $\mathbf{s}$. To obtain a good HRTF estimate, $\mathbf{h}_1$, $\mathbf{h}_2$, and $\mathbf{h}_3$ should be close to $\mathbf{s}$ in terms of their distance along the sphere. Furthermore, it makes sense to have a consistent mapping of a source direction $\mathbf{s}$ and the measurement points used for interpolation, i.e., there should be a unique representation of an HRTF estimate at $\mathbf{s}$ as the linear combination of measurement points $\mathbf{h}_i$. To meet these requirements, the measurement points are grouped into triplets that form non-overlapping triangles on the unit sphere. Efficient algorithms exist to perform this triangulation. With all measurement points lying on the surface of a sphere, the triangulation can be done by calculating the convex hull of the measurement points. This yields a Delaunay triangulation [16]. The Delaunay triangulation maximises the minimum angle of the generated triangles, which is a desirable property for interpolation [16, 14]. Figure 2 illustrates the Delaunay triangulation for 100 random measurement points. As can be seen, the triangulation deals well with grid irregularities.

The result of the triangulation is stored in memory as a list of measurement point triplets. For a desired source direction $\mathbf{s}$, the linear combination of a measurement point triplet $\{\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3\}$ is given as:

$$\mathbf{s} = g_1\mathbf{h}_1 + g_2\mathbf{h}_2 + g_3\mathbf{h}_3, \tag{1}$$

where g_1, g_2, and g_3 are scalar weights. With $\mathbf{g} = [g_1 \;\; g_2 \;\; g_3]^{\mathrm{T}}$ and $\mathbf{H} = [\mathbf{h}_1 \;\; \mathbf{h}_2 \;\; \mathbf{h}_3]$, the weights are obtained as:

$$\mathbf{g} = \mathbf{H}^{-1}\mathbf{s}. \tag{2}$$

To minimise the computational effort of evaluating eq. (2) at runtime, the inverse $\mathbf{H}^{-1}$ can be calculated for every measurement triplet during initialisation and stored in memory. This reduces

the operation count required for evaluating eq. (2) at runtime to 9 multiplications and 6 additions per measurement triplet.

3.3. HRTF subset selection

The criterion for selecting a measurement triplet $\{\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3\}$ for interpolation is that the weights calculated via eq. (2) are positive:

$$g_1, g_2, g_3 \geq 0. \tag{3}$$

This is equivalent to $\mathbf{s}$ lying inside the curved triangle formed by $\{\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3\}$ on the unit sphere. At runtime, the subset selection algorithm iterates through all measurement triplets until a suitable triplet is found. The performance of this "brute-force" implementation for various publicly available HRTF databases is illustrated in fig. 3.

Next, an improvement to the subset selection algorithm is proposed. Instead of iterating through the triplets in an arbitrary order, the triplets are sorted according to their distance from the desired source direction $\mathbf{s}$, so that the selection algorithm first iterates through triplets close to $\mathbf{s}$. The triplet distance is calculated as the distance between $\mathbf{s}$ and the centre of the triangle formed by the measurement triplet. A straightforward way to express the centre $\mathbf{c}$ of a (curved) triangle is to take the arithmetic mean of the triangle vertices $\{\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3\}$:

$$\mathbf{c} = (\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3)/3. \tag{4}$$

The distance d along the sphere between $\mathbf{s}$ and the triangle is then given as:

$$d = \arccos(\mathbf{s} \bullet \mathbf{c}), \tag{5}$$

where $\bullet$ is the inner product of $\mathbf{s}$ and $\mathbf{c}$. However, for the purpose of ordering the measurement triangles with respect to their distance to $\mathbf{s}$, it is sufficient to calculate the inner product $\hat{d}_{\mathrm{ip}}$:

$$\hat{d}_{\mathrm{ip}} = \mathbf{s} \bullet \mathbf{c}. \tag{6}$$

Sorting all triangles by descending $\hat{d}_{\mathrm{ip}}$ is equivalent to sorting them by ascending distance d. As $\mathbf{c}$ can be calculated via eq. (4) during initialisation and stored in memory, evaluating eq. (6) at runtime reduces to 3 multiplications and 2 additions per measurement triplet (whereas the brute-force approach evaluates eq. (2) which requires 9 multiplications and 6 additions per triplet). Efficient algorithms exist for sorting the triangles according to distance, given the list $L_{1,\mathrm{N}} = \{\hat{d}_{\mathrm{ip},1}, \ldots, \hat{d}_{\mathrm{ip},\mathrm{N}}\}$, where N is the number of measurement triplets. Assuming that a suitable triangle for interpolation is among the K triangles closest to $\mathbf{s}$, but not necessarily *the* closest, it is not necessary to sort the whole list. Instead, quick selection can be used to partition $L_{1,\mathrm{N}}$ into 2 unsorted lists: $A_{1,\mathrm{K}}$, containing the K closest triangles, and $B_{\mathrm{K}+1,\mathrm{N}}$, containing the remaining triangles [17]. With K set to 0.02 N, $A_{1,\mathrm{K}}$ consists of the closest 2 percent of all triangles. To speed up the quick selection, K is set to an admissible range rather than a fixed value [17]. Here, K is set to 0.02 N$\pm$ 0.004 N.

To find a suitable triangle for interpolation, the subset selection algorithm starts by iterating over $A_{1,\mathrm{K}}$. If no suitable triangle is found among the closest triangles, the selection algorithm continues to iterate over the remaining triangles $B_{\mathrm{K}+1,\mathrm{N}}$. Figure 3 illustrates the effect of first iterating over the closest triangles on the total processing time of the subset selection algorithm (fig. 3, dark grey dots), for 1000 random source directions $\mathbf{s}$ and three different HRTF databases. The brute-force implementation (fig. 3, light grey dots) iterates over all triangles regardless of their distance to $\mathbf{s}$, until a suitable one is found. As expected, the processing time of the brute-force implementation increases linearly with the iteration index j of the selected triangle. Therefore, the processing time of the brute-force subset selection algorithm is dependent on the desired source direction $\mathbf{s}$.

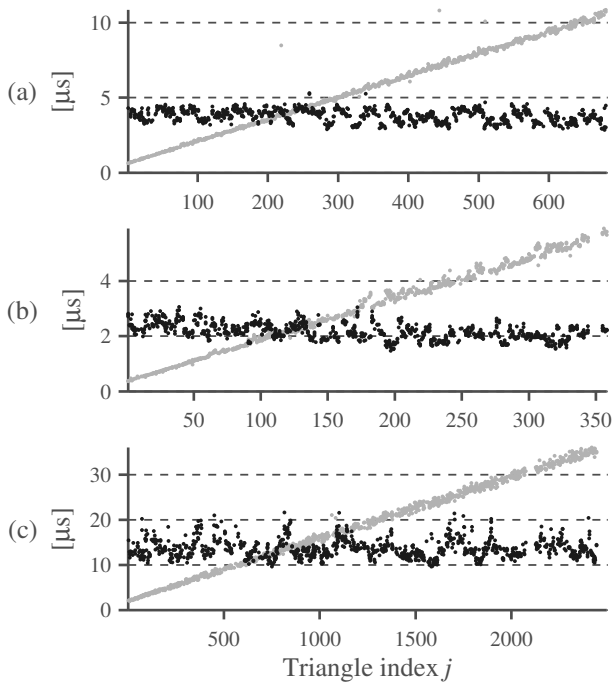

Figure 3: *Processing times of brute-force iteration (light grey dots) and iteration over the* K *closest triangles after quick selection (dark grey dots), for 1000 random source directions, using the HRTF databases: (a) MIT KEMAR [1]; (b) LISTEN [3]; (c) CIPIC [2]. Each dot denotes the time required to find an HRTF measurement triplet j suitable for interpolation for a random source direction, averaged over 500 repetitions, on a computer with a 2GHz quad-core processor.*

By using quick selection to find the K closest triangles and iterating over those first, the processing time of the subset selection algorithm is approximately constant regardless of the triangle index (see fig. 3, dark grey dots), that is, a suitable triangle is found within a certain time span for any source direction $\mathbf{s}$. Constant processing time is a desirable property for spatial sound rendering algorithms as it allows for a constant rendering update rate when rendering for example moving sources.

3.4. Calculation of interpolation weights

The weights $\mathbf{g}$, calculated via eq. (2) to check whether a measurement triplet $\{\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3\}$ is suitable for interpolation, can directly be used as interpolation weights, after normalisation:

$$g_{i,\mathrm{int}} = \frac{g_i}{\sum_{j=1}^{3} g_j} \qquad \Rightarrow \qquad \sum_{i=1}^{3} g_{i,\mathrm{int}} = 1 \tag{7}$$

where $g_{i,\mathrm{int}}$ is the interpolation weight of measurement point $\mathbf{h}_i$.

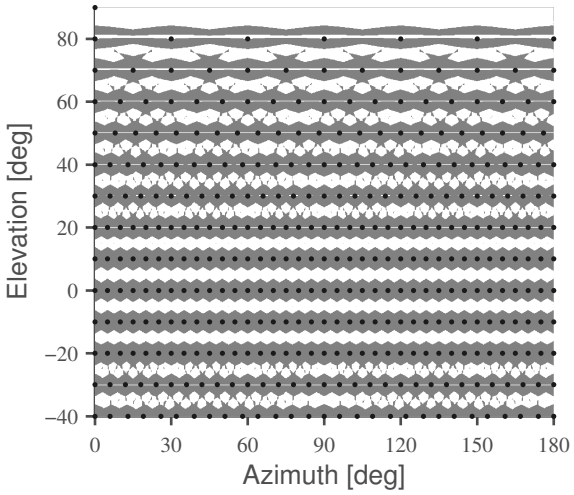

Figure 4: *Dots denote measurement points in the MIT KEMAR HRTF database [1]. Source directions in grey areas are not enclosed by their respective 3 nearest measurement points.*

If the desired source $\mathbf{s}$ lies on an edge of a measurement triangle, the vertex opposite to that edge has a weight of zero, thus effectively only the two measurement points forming the edge are selected for interpolation. Similarly, if $\mathbf{s}$ lies on a vertex, the other two vertices of the same triangle have zero weights, and only the HRTF measured at $\mathbf{s}$ is selected.

4. EXPERIMENTS

To evaluate the proposed approaches for HRTF subset selection and calculation of interpolation weights, experiments are carried out using both modelled and measured HRTFs.

4.1. HRTF subset selection

Even though the proposed subset selection approach is very efficient computationally (see fig. 3), an even faster method is to simply select the 3 nearest measurement points for interpolation [6, 7]. However, a drawback of this approach is that the 3 points nearest to the desired source direction $\mathbf{s}$ do not necessarily enclose $\mathbf{s}$, and may for example form a line rather than a triangle (see N3 in fig. 1). In the case of the nearest measurement points forming a line, it is questionable whether using all 3 points for interpolation would yield a better HRTF estimate than using for example only the 2 nearest measurement points. Figure 4 shows large areas for which source directions are not enclosed by their respective 3 nearest measurement points (with respect to their distance along the sphere) in the MIT KEMAR HRTF database [1]. This is mostly due to the nearest measurement points lying on a line with constant elevation, and the desired source direction $\mathbf{s}$ lying above or below that line (see section 4.2 for a discussion on the consequences).

Given the drawbacks of the nearest-3 selection algorithm, the additional computational cost of the subset selection algorithm proposed here, incurred for testing whether a selected measurement triplet encloses the desired source direction, seems justified.

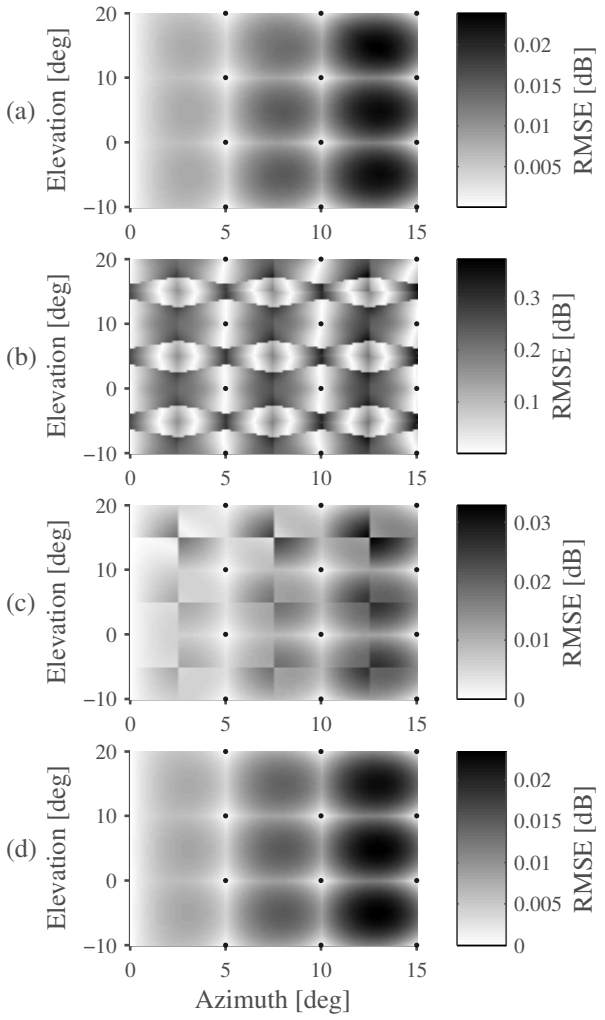

Figure 5: *RMSE for interpolating a simple head-shadowing model sampled at regular azimuth and elevation intervals (dark dots) using (a) normalised VBAP weighting (proposed here); (b) inverse distance weighting; (c) bilinear interpolation of 3 measurement points; (d) bilinear interpolation of 4 measurement points.*

4.2. Interpolation of modelled HRTF data

Using a simple model for the angle-dependent effect of head shadowing on sound reaching the ears [18], the performance of the proposed method for calculating interpolation weights is compared to methods proposed in the literature. The model is implemented as an IIR filter:

$$H_{\mathrm{hs}}(z,\varphi) = \frac{\left(\frac{c}{a} + \alpha(\varphi) f_{\mathrm{s}}\right) + \left(\frac{c}{a} - \alpha(\varphi) f_{\mathrm{s}}\right) z^{-1}}{\left(\frac{c}{a} + f_{\mathrm{s}}\right) + \left(\frac{c}{a} - f_{\mathrm{s}}\right) z^{-1}}, \qquad (8)$$

with

$$\alpha(\varphi) = 1.05 + 0.95 \cos\left(\frac{180}{150}\left(\varphi + \frac{\pi}{2}\right)\right), \qquad (9)$$

where c denotes the speed of sound, a the head radius, φ the azimuth angle, and f_{s} the sampling rate. Note that φ is given in *interaural-polar coordinates* [19]. In interaural-polar coordinates,

H_{hs} is independent of elevation. Given azimuth ϕ and elevation δ in standard vertical-polar coordinates, φ is obtained as

$$\varphi = \arcsin\left(\sin(\phi)\cos(\delta)\right). \tag{10}$$

Using the head-shadowing model given by eq. (8), approximate HRTFs are obtained at regular intervals of azimuth and elevation. Four different methods are used to calculate interpolation weights and to interpolate those approximate HRTFs:

a) the normalised VBAP weights proposed here;

b) inverse distance weighting [7];

c) bilinear interpolation of 3 measurement points [11];

d) bilinear interpolation of 4 measurement points [9].

The methods are compared in terms of the root-mean-square error (RMSE) between the interpolated and the modelled HRTF magnitude response. The purpose of the comparison is to illustrate qualitative, rather than quantitative, differences between the compared interpolation approaches.

Figure 5 illustrates the RMSE of the four methods for a range of azimuth and elevation angles. As expected, all algorithms are virtually error-free at measurement points. Moving away from measurement points, the interpolation error increases. The interpolation error of the bilinear interpolation of 3 measurement points exhibits slight discontinuities (see fig. 5c), caused by interpolation discontinuities. The discontinuities coincide with shared edges of measurement triangles: As the source position crosses a shared edge between two measurement triangles, the measurement points used for interpolation switch abruptly from one triangle to the other, causing a discontinuity in the interpolated HRTF. Discontinuities are undesirable when rendering moving sound sources, as they could potentially cause audible discontinuities in the sound. The inverse distance weighting shows more severe discontinuities (see fig. 5b). These are partially caused by the fact that the nearest points used for interpolation may form a line, rather than a triangle enclosing the desired source direction (see fig. 4). The other reason why inverse distance weighting produces discontinuities in the interpolation is the same as for the bilinear interpolation of 3 measurement points: The transition between one selected subset of measurement points to another is discontinuous. The interpolation of the normalised VBAP weighting proposed here (see fig. 5a) and of bilinear interpolation of 4 measurement points (see fig. 5d) is nearly identical for the tested range of azimuth and elevation angles. Both methods interpolate smoothly in azimuth and elevation, without discontinuities, which makes them good candidates for rendering moving sound sources. The reason these two methods do not produce discontinuities is that for a source moving close to an edge of a selected measurement subset, the weights of vertices not lying on that edge gradually diminish, until they become zero for a source lying on the edge. This allows for a smooth transition from one selected measurement subset, across an edge, to another subset. The advantage of the proposed VBAP weighting over the bilinear interpolation of 4 measurement points is that it selects at most 3 measurement points for interpolation and does not depend on a regular measurement grid (see fig. 2). As the performance of the two methods is very similar otherwise and the bilinear interpolation is not well defined in the presence of grid irregularities, bilinear interpolation is dropped from the following experiments.

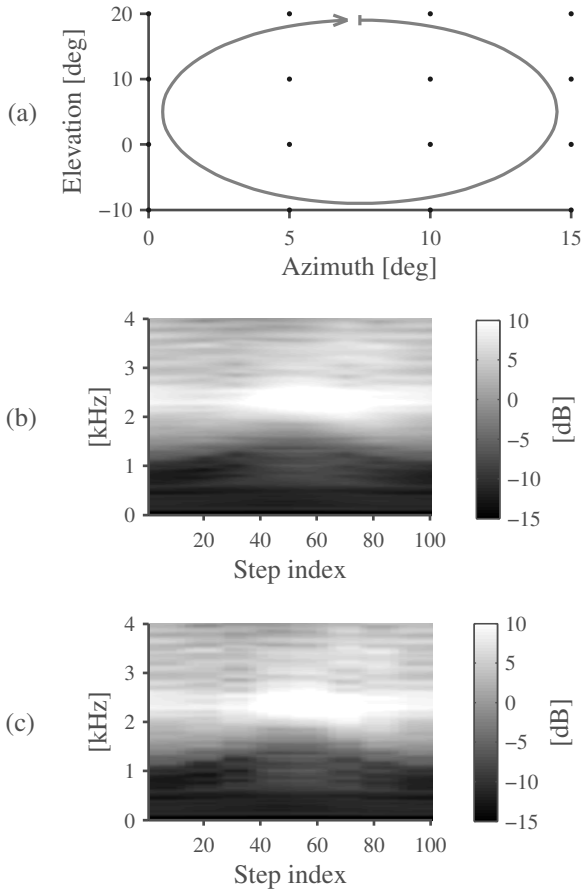

Figure 6: *HRTF interpolation for a moving source; (a) the source path, discretised at 100 steps; (b) spectrogram of HRTF interpolation using the proposed methods for subset selection and calculation of interpolation weights; (c) spectrogram of HRTF interpolation by inverse distance weighting of the nearest 3 HRTFs.*

4.3. Rendering a moving virtual source

To illustrate the effect of interpolation discontinuities, the rendering of a moving sound source is simulated using the MIT KEMAR HRTF database [1]. After subset selection and interpolation weight calculation, the interpolation is performed on the magnitude of the measured HRTFs in the frequency domain, as proposed by Zotkin et al. [7]. The phase of the interpolated HRTF can be derived from a spherical head model [7] and is not considered here.

Figure 6 shows spectrograms of the interpolated HRTFs for a source moving along a circular path (fig. 6a). The proposed methods for subset selection and calculation of interpolation weights produce smooth HRTF estimates (fig. 6b). Interpolation using the nearest 3 measurement points and inverse distance weighting produces discontinuities visible in the spectrogram (fig. 6c). Informal listening tests indicate that these discontinuities may be audible.

4.4. Interpolation of measured HRTF data

To compare the overall performance of the proposed method to other approaches, interpolation is performed on the MIT KEMAR

Table 1: *Average RMSE for interpolating the measured left- and right-ear HRTFs in (a) the MIT KEMAR HRTF database [1], (b) the LISTEN HRTF database [3], and (c) the CIPIC HRTF database [2]. The minimum RMSE for each row is in bold.*

		proposed [dB]	bilinear 3 [dB]	nearest-3 [dB]	nearest-2 [dB]
(a)	left	**1.809**	1.811	1.916	1.811
	right	**2.015**	2.017	2.142	2.017
(b)	left	3.480	3.517	**3.432**	3.557
	right	**3.195**	3.228	3.248	3.291
(c)	left	2.195	2.201	**2.185**	2.201
	right	**2.217**	2.222	2.226	2.222

HRTF database [1], by removing one measurement point at a time and comparing the measured HRTF to the HRTF estimated via interpolation for the same point. The methods compared are:

a) the method for subset selection and interpolation weight calculation proposed here;

b) bilinear interpolation of 3 measurement points;

c) inverse distance weighting of the 3 nearest measurement points;

d) inverse distance weighting of the 2 nearest measurement points (for comparison).

The interpolation is performed on the magnitude spectra, as proposed by Zotkin et al. [7]. Table 1 summarises the results in terms of the RMSE averaged over all measurement points. For the MIT KEMAR database (table 1a), all methods have nearly identical performance, except for the nearest-3 interpolation which performs slightly worse. This is due to the measurement grid structure of the MIT KEMAR database, which causes the nearest-3 subset selection to perform badly (see figs. 4 and 5). The algorithm is actually outperformed by interpolation using just the 2 nearest measurement points. As hypothesised, there is no advantage in using 3 measurement points for interpolation, if all 3 points lie on a line (see section 4.1). For the LISTEN database (table 1b), the RMSE of all tested algorithms is similar. The RMSE is higher than for the MIT KEMAR database, as the measurement grid of the LISTEN database is less dense. The results for the CIPIC database are analogous (table 1c). The improved performance of the nearest-3 interpolation for the LISTEN and CIPIC database indicates that for these databases the 3 nearest measurement points are more likely to form a triangle enclosing the point to be estimated. The method proposed here performs well for all 3 tested HRTF databases.

Overall, for the HRTF databases used in this comparison, the differences between the tested methods are small. This is partly due to the way the RMSE is calculated (by removing one measurement point at a time) and the fact that the tested databases exhibit highly regular measurement grids (with the exception of some irregularities in the MIT KEMAR and LISTEN databases at extreme elevation values, where the measurement grid is sparser). This is also the reason for the good performance of the inverse-distance interpolation of the nearest 2 measurement points: After removing one measurement point from the HRTF database, the nearest measurement points usually lie to the left and right of the removed point, at equal distance and elevation. For this special case, all algorithms except the nearest-3 interpolation potentially select those two neighbouring points and calculate the same interpolation weights, thus resulting in an identical interpolation. Obviously, this is not true for any point that does not lie precisely half-way between 2 measurement points at equal elevation. Interpolating at positions that do not coincide exactly with measurement points, or using an HRTF database with a more irregular grid structure, would produce larger differences with respect to the measurement points selected and the weights calculated for interpolation, and thus the resulting HRTF interpolation.

5. SUMMARY AND CONCLUSION

A method for selecting a subset of an HRTF measurement database for HRTF interpolation and calculating interpolation weights was presented. The main steps of the algorithm are:

1. *Initialisation*: group HRTF measurement points into non-overlapping triangles by calculating their convex hull and pre-calculate the inverse and centres of the triangles;
2. *source update*: iterate through the measurement triplets (after optional sorting according to distance) until a suitable triplet for interpolation is found;
3. *interpolation*: calculate interpolation weights via vector-base amplitude panning (VBAP) and normalisation.

After selecting an HRTF subset and calculating appropriate interpolation weights, the actual interpolation can be performed as a linear combination of the time- or frequency-domain responses. The proposed algorithm proves to be computationally efficient and robust with respect to irregular measurement grids. The measurement triangles obtained by calculating the convex hull of the measurement points represent a Delaunay triangulation, which maximises the minimum angle of all triangles, a property advantageous for interpolation [16, 14].

The interpolation weights calculated via VBAP and normalisation change smoothly with azimuth and elevation, allowing for smooth interpolation of moving sources without interpolation discontinuities. The accuracy of the proposed method for interpolating an HRTF database is comparable to or better than other state-of-the-art methods, while not suffering from interpolation discontinuities for moving sources.

6. REFERENCES

[1] W. G. Gardner and K. D. Martin, "HRTF measurements of a KEMAR," *J. Acoust. Soc. Am.*, vol. 97, no. 6, pp. 3907–3908, 1995.

[2] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *Proc. IEEE WASPAA*, New Paltz, NY, Oct. 2001, pp. 99–102.

[3] IRCAM, "LISTEN HRTF database," http://recherche.ircam.fr/equipes/salles/listen/, last accessed March 12, 2013.

[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, Apr. 1999, pp. 319–329.

[5] F. Keyrouz and K. Diepold, "A new HRTF interpolation approach for fast synthesis of dynamic environmental interaction," *J. Audio Eng. Soc.*, vol. 56, no. 1/2, pp. 28–35, 2008.

[6] J.-M. Jot, V. Larcher, and O. Warusfel, "Digital signal processing issues in the context of binaural and transaural stereophony," in *Proc. 98th Conv. Audio Eng. Soc.*, Paris, France, Feb. 1995.

[7] D. Zotkin, R. Duraiswami, and L. Davis, "Rendering localized spatial audio in a virtual auditory space," *IEEE Trans. Multimedia*, vol. 6, no. 4, pp. 553–564, 2004.

[8] E. Wenzel and S. Foster, "Perceptual consequences of interpolating head-related transfer functions during spatial synthesis," in *Proc. IEEE WASPAA*, New Paltz, NY, Oct. 1993, pp. 102–105.

[9] L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen, "Creating interactive virtual acoustic environments," *J. Audio Eng. Soc.*, vol. 47, no. 9, pp. 675–705, 1999.

[10] E. H. A. Langendijk and A. W. Bronkhorst, "Fidelity of three-dimensional-sound reproduction using a virtual auditory display," *J. Acoust. Soc. Am.*, vol. 107, no. 1, pp. 528–537, 2000.

[11] F. P. Freeland, L. W. P. Biscainho, and P. S. R. Diniz, "Interpositional Transfer Function of 3D-Sound Generation," *J. Audio Eng. Soc.*, vol. 52, no. 9, pp. 915–930, 2004.

[12] M. Queiroz and G. H. M. a. de Sousa, "Efficient binaural rendering of moving sound sources using HRTF interpolation," *J. New Music Research*, vol. 40, no. 3, pp. 239–252, 2011.

[13] S. Carlile, C. T. Jin, and V. Van Raad, "Continuous virtual auditory space using HRTF interpolation: acoustic and psychophysical errors," in *Proc. IEEE Pacific Rim Conf. Multimedia*, Sydney, Australia, Dec. 2000, pp. 220–223.

[14] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456–466, 1997.

[15] D. S. Brungart and W. M. Rabinowitz, "Auditory localization of nearby sources. head-related transfer functions," *J. Acoust. Soc. Am.*, vol. 106, no. 3, pp. 1465–1479, 1999.

[16] F. Aurenhammer, "Voronoi diagrams – a survey of a fundamental geometric data structure," *ACM Comput. Surv.*, vol. 23, no. 3, pp. 345–405, 1991.

[17] C. Martínez, A. Panholzer, and H. Prodinger, "The analysis of range quickselect and related problems," *Theor. Comput. Sci.*, vol. 412, no. 46, pp. 6537–6555, 2011.

[18] V. Pulkki, T. Lokki, and D. Rocchesso, "Spatial effects," in *DAFX: Digital Audio Effects*, 2nd ed., U. Zölzer, Ed. Chichester, UK: Wiley, March 2011, pp. 139–184.

[19] V. R. Algazi, R. O. Duda, R. Duraiswami, N. A. Gumerov, and Z. Tang, "Approximating the head-related transfer function using simple geometric models of the head and torso," *J. Acoust. Soc. Am.*, vol. 112, no. 5, pp. 2053–2064, 2002.