HEARING DISTANCE: A LOW-COST MODEL FOR NEAR-FIELD BINAURAL EFFECTS

Simone Spagnol

IUAV - University of Venice

Michele Geronazzo, Federico Avanzini

University of Padova

ABSTRACT

An extremely low-order filter model for source distance rendering in binaural reproduction is proposed in this paper. The main purpose of such model is to cheaply simulate the effect that source-listener distance has on the sound waves arriving at the ears in the near field, a region where the relation between sound pressure and distance is both highly frequency-dependent and nonlinear. The reference for the model is based on an analytical description of a spherical head response, appropriately filtered out so as to include distance-dependent patterns only. To this regard, the model is objectively seen to provide an excellent fit in the whole near field, despite its simplicity.

Index Terms— 3D audio, spatial sound, HRTF

1. INTRODUCTION

Under anechoic conditions, perception of sound source distance is correlated to the signal's intensity, even though it is systematically underestimated for far distances and overestimated for near ones. On the other hand, if the environment is reverberant, a sensation of changing distance occurs if the overall intensity is constant but the proportion of reflected to direct energy, the so-called *R/D ratio*, is altered. Still, a number of other cues contribute to a correct perception of distance, such as distance-dependent spectral effects, familiarity of the involved sounds, and dynamic cues [1].

When the source is in the *far field* (i.e. approximately more than 1.5 m from the center of the listener's head), directional cues are roughly distance-independent. By gradually approaching the *near field* (i.e. when the source is approximately within 1.5 m from the center of the head), it is known that whereas ITD (Interaural Time Difference) remains almost independent from distance, ILD (Interaural Level Difference) dramatically increases across the whole spectrum and in particular at low frequencies. Since distance dependence must then be taken into account in the near field, a prompt characterization of the head's response (Head-Related Transfer Function, HRTF) in such region has to be studied. For the sake of simplicity, the head of the listener can be treated as a rigid sphere [2]; therefore its transfer function will be referred to as *spherical transfer function*, or STF.

Relatively simple STF-like filter structures for sound source rendering in the far field have been proposed to date, e.g. Brown and Duda's first-order filter [3]. These models, although replicating with some degree of approximation the mean magnitude characteristics of the far-field spherical response, do not simulate the rippled behaviour occurring for contralateral sources, and, more importantly, have no parametrization on source distance. Unfortunately, to the authors' knowledge, no real-time model including near-field effects is available in the literature. As a consequence, a proper approximation to distance effects on the spherical head model has to be introduced in order to grant an efficient and effective rendering.

In this paper such an approximation is proposed and analytically evaluated. In Section 2 we briefly introduce the analytical formulation of the distance-dependent STFs, which are approximated following our method outlined in Section 3. A real-time filter model directly derived from the previous analysis is at first objectively evaluated in Section 4.

2. THE SPHERICAL TRANSFER FUNCTION

Within the assumption of an infinitely distant source from the center of the head, the response related to a fixed observation point on the sphere's surface can be described by means of the following transfer function:

$$H(\mu, \theta_{\text{inc}}) = \frac{1}{\mu^2} \sum_{m=0}^{\infty} \frac{(-i)^{m-1}(2m+1)P_m(\cos\theta_{\text{inc}})}{h'_m(\mu)}, \quad (1)$$

where θ_{inc} is the incidence angle (i.e. the angle between the ray from the center of the sphere to the source and the ray to the observation point), P_m and h_m represent the *Legendre polynomial* of degree *m* and the *m*th-order *spherical Hankel function* (where h'_m is its first derivative with respect to its argument), respectively, and μ is the normalized frequency, defined as

$$\mu = f\frac{2\pi a}{c}, \quad (2)$$

where c is the speed of sound and a is the sphere radius.

When the assumption of an infinitely distant source is removed, distance dependence can no longer be ignored. Having defined the normalized distance to the source ρ as the ratio between the absolute distance from the center of the sphere

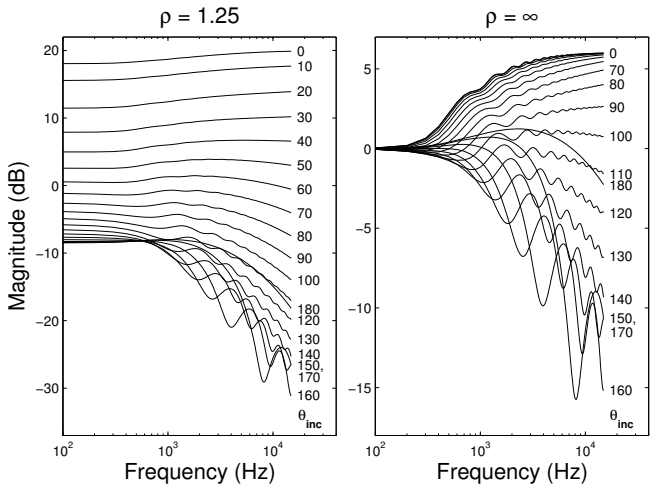

Fig. 1. Near-field and far-field Spherical Transfer Functions: $\rho = 1.25$ (left panel) and $\rho = \infty$ (right panel).

and the sphere radius, the STF can be evaluated by means of the following function [4]:

$$H(\rho,\mu,\theta_{\text{inc}}) = -\frac{\rho}{\mu} e^{-i\mu\rho} \sum_{m=0}^{\infty} (2m+1) P_m(\cos\theta_{\text{inc}}) \frac{h_m(\mu\rho)}{h'_m(\mu)}, \quad (3)$$

for each $\rho > 1$. A description of the evaluation algorithm, based on recursion relations, can be found in [2]. Fig. 1 shows the magnitude of the so calculated transfer function against normalized frequency for 19 different values of the incidence angle and two distances, corresponding to near- and far-field.

In a previous work [5], the authors have used Principal Component Analysis (PCA) in order to study common trends and possible systematic variability in a set of STFs. The results indicated that angular dependence is much more prominent than distance dependence in the transfer function's frequency behaviour. Isolating distance information from the spherical response is thus the first goal towards the design of a cheap and effective model for the STF, as will be performed in the following sections.

3. THE NEAR-FIELD TRANSFER FUNCTION

In order to study the effect of source distance in the near field, a given STF can be normalized to the corresponding far field spherical response yielding a new transfer function, which we refer to as Near-Field Transfer Function (NFTF):

$$H_{NF}(\rho,\mu,\theta_{\text{inc}}) = \frac{H(\rho,\mu,\theta_{\text{inc}})}{H(\infty,\mu,\theta_{\text{inc}})}. \quad (4)$$

Contrarily to STFs, NFTFs are almost non-rippled functions that slightly decay with frequency, in an approximately monotonic fashion. Furthermore, the magnitude boost for small distances is evident in ipsilateral NFTFs whereas it is less prominent in contralateral NFTFs. Such functions are now analyzed in detail in all of their magnitude features with the parallel aim of feeding the modeling process. In the following, we fix for convenience parameter a to the commonly cited 8.75-cm average radius for an adult human head [3]. Final results for a different head radius will just require a uniform rescaling of the frequency axis.

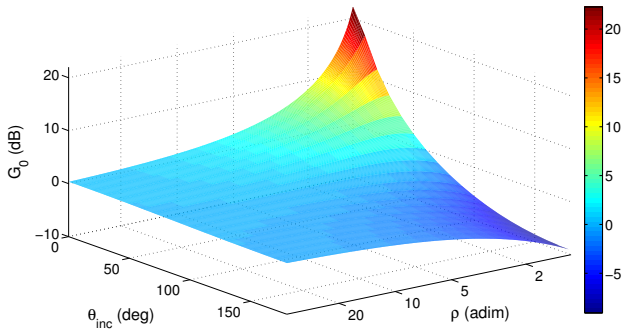

Fig. 2. NFTF gain at DC.

It could be questioned whether analytical NFTFs objectively reflect distance-dependent patterns in measured HRTFs. Unfortunately, most available HRTF recordings are performed in the far field or in its vicinity at one single given distance. Furthermore, a proper NFTF will become more and more sensitive to the geometric features of the complex scatterer (the head) as the sound source approches and, since the sphere can be considered as a simple scatterer, it could become an increasingly worse approximation of the real near field effects. Beside these lawful observations, various studies have attested the suitability of the spherical model for the head in the near field, at least at low frequencies [1].

3.1. DC gain of NFTFs

As a first step towards NFTF analysis, let us look more closely at how the DC gain varies as the source moves away along a given angular direction. For each of 19 incidence angles, $\theta_{\text{inc}} = [0°, 180°]$ at 10-degree steps, Eq. (3) is sampled at DC ($\mu = 0$) for a great number of different, exponentially increasing distances, specifically

$$\rho = 1.15^{1+\frac{k-1}{10}}, \qquad k = 1, \ldots, 250, \quad (5)$$

and its absolute value calculated, yielding dB gain $G_0(\theta_{\text{inc}}, \rho)$. Fig. 2 plots DC gains as functions of distance and incidence angle.

In order to model distance dependence of NFTFs at DC we approximate it as a second-order rational function for all the 19 different incidence angles. This function, that has the form

$$\tilde{G}_0(\theta_{\text{inc}}, \rho) = \frac{p_{11}(\theta_{\text{inc}})\rho + p_{21}(\theta_{\text{inc}})}{\rho^2 + q_{11}(\theta_{\text{inc}})\rho + q_{21}(\theta_{\text{inc}})}, \quad (6)$$

θ_{inc}	p_{11}	p_{21}	q_{11}	q_{21}	RMS [dB]
0°	12.97	−9.691	−1.136	0.219	0.0027
10°	13.19	234.2	18.48	−8.498	0.0223
20°	12.13	−11.17	−1.249	0.346	0.0055
30°	11.19	−9.035	−1.017	0.336	0.0034
40°	9.91	−7.866	−0.83	0.379	0.002
50°	8.328	−7.416	−0.666	0.421	0.0009
60°	6.493	−7.312	−0.503	0.423	0.0002
70°	4.455	−7.278	−0.321	0.382	0.0004
80°	2.274	−7.291	−0.11	0.314	0.0005
90°	0.018	−7.484	−0.13	0.24	0.0005
100°	−2.242	−8.04	0.395	0.177	0.0004
110°	−4.433	−9.231	0.699	0.132	0.0003
120°	−6.488	−11.61	1.084	0.113	0.0002
130°	−8.342	−17.38	1.757	0.142	0.0002
140°	−9.93	−48.42	4.764	0.462	0.0004
150°	−11.29	9.149	−0.644	−0.138	0.0006
160°	−12.22	1.905	0.109	−0.082	0.0003
170°	−12.81	−0.748	0.386	−0.058	0.0002
180°	−13	−1.32	0.45	−0.055	0.0002

Table 1. Coefficients for Eq. (6) and approximation fitness.

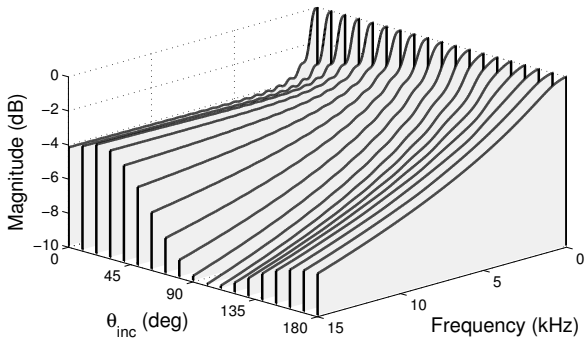

Fig. 3. Frequency trend of normalized NFTFs for $\rho = 1.25$.

where $\theta_{\text{inc}} = 0^\circ, 10^\circ, \ldots, 180^\circ$, is found with the help of the MatLab Curve Fitting Toolbox (`cftool`).

Coefficients $p_{11}(\theta_{\text{inc}})$, $p_{21}(\theta_{\text{inc}})$, $q_{11}(\theta_{\text{inc}})$, and $q_{21}(\theta_{\text{inc}})$ for each of the 19 incidence angles are reported in Table 1, as well as the RMS (root mean square) error measure between real and approximated DC gains for each incidence angle at the 250 evaluated distances. The latter values confirm the overall excellent fitness of the resulting rational functions: in all cases, $RMS(G_0, \tilde{G}_0) < 0.01$ dB. In order to model DC gain for intermediate incidence angles, a simple linear interpolation between adjacent functions can be used. The effective fitness of such an approximation on a dB scale will be objectively evaluated in Section 4, even for incidence angles different from those considered up to now.

3.2. Frequency behaviour of NFTFs

The behaviour of NFTFs at DC having been checked, it has to be studied how much NFTFs depend on frequency and how such dependence can be efficiently modeled. In order to do this, the DC gain G_0 can act as a further normalization factor, thus the following operation is performed for a set of NFTFs computed at the already considered 250 distances and in the frequency range up to 15 kHz, sampled at 10-Hz steps:

$$\hat{H}_{NF}(\rho, \mu, \theta_{\text{inc}}) = \frac{H_{NF}(\rho, \mu, \theta_{\text{inc}})}{G_0(\theta_{\text{inc}}, \rho)}. \tag{7}$$

Fig. 3 shows the frequency behaviour of normalized NFTFs for a fixed small distance, $\rho = 1.25$, and the usual 19 incidence angles. All normalized NFTFs lie in the negative dB space, tending to the 0-dB threshold at most. This means that DC is always the frequency point of the NFTF where the gain is maximum. However, note the different high-frequency trend for ipsilateral and contralateral angles: as an example, at $\theta_{\text{inc}} = 0^\circ$ the magnitude plot looks like that of a high-frequency shelving filter, whereas at $\theta_{\text{inc}} = 180^\circ$ a lowpass behaviour is observed. For intermediate incidence angles, the response for $\rho = 1.25$ gradually morphs from that of a shelving filter to that of a lowpass filter as the angle increases. The same behaviour is observed for all the other near-field distances, the faster switch rate being observed for small distance values.

In light of such result, one could think of approximating the magnitude response of the normalized NFTF through a shelving or lowpass filter, depending on incidence angle and distance. Unfortunately, the switch from a shelving to a lowpass filter at a given incidence angle needs to be smooth in order to avoid listening artifacts; furthermore, a first-order lowpass filter excessively cuts high frequencies with respect to the maximum 10-dB decay observed in the normalized NFTF plots. These shortcomings can be solved, although at the cost of precision loss, by always approximating a normalized NFTF through a first-order high-frequency shelving filter. The implementation chosen for the filter is the one proposed in [6],

$$H_{\text{sh}}(z) = 1 + \frac{H_0}{2}\left(1 - \frac{z^{-1} + a_c}{1 + a_c z^{-1}}\right), \tag{8}$$

$$a_c = \frac{V_0 \tan\left(\pi \frac{f_c}{f_s}\right) - 1}{V_0 \tan\left(\pi \frac{f_c}{f_s}\right) + 1}, \qquad V_0 = 10^{\frac{G_\infty}{20}}, \tag{9}$$

where f_s is the sampling frequency.

Now it has to be highlighted how the two key parameters of the shelving filter, cutoff frequency f_c and asymptotic high-frequency gain G_∞, can be extracted from $\hat{H}_{NF}$ in order to yield a satisfactory approximation. First, the high-frequency gain is calculated as the (negative) dB gain of the NFTF at 15 kHz. The choice of a high frequency point is needed to best model the slope of near contralateral NFTFs in the whole audible range. Second, the cutoff frequency is calculated as the frequency point where $\hat{H}_{NF}$ has a negative dB gain which approximates two thirds of the high-frequency gain. This point is heuristically preferred to the point where the gain is $\frac{G_\infty}{2}$ in order to minimize differences in magnitude between a shelving filter and a lowpass filter for contralateral NFTFs. The quality of the shelving filter approximation is well depicted in Fig. 4 for three different distances at all incidence angles.

The variation of parameters G_∞ and f_c along distance and incidence angle was also studied. Similarly to what was done

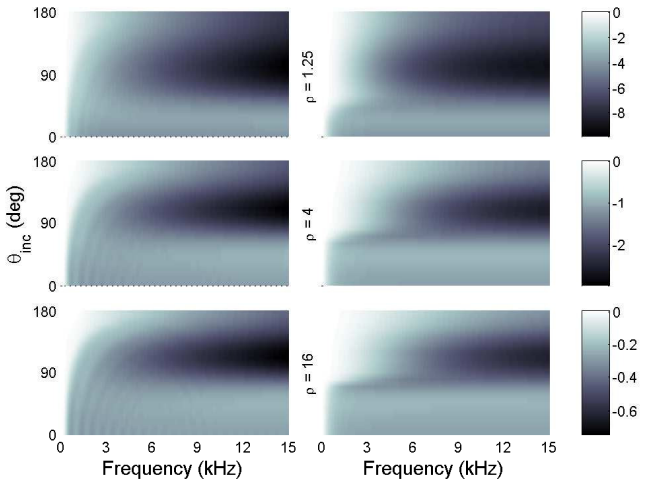

Fig. 4. Normalized NFTFs (left panels) versus shelving-filter approximation (right panels) for $\rho = 1.25$, $\rho = 4$, and $\rho = 16$.

θ_{inc}	p_{12}	p_{22}	q_{12}	q_{22}	RMS [dB]
0°	−4.391	2.123	−0.55	−0.061	0.0007
10°	−4.314	−2.782	0.59	−0.173	0.0016
20°	−4.18	4.224	−1.006	−0.021	0.0057
30°	−4.012	3.039	−0.563	−0.316	0.0116
40°	−3.874	−0.566	0.665	−1.129	0.0199
50°	−4.099	−34.74	11.39	−8.301	0.039
60°	−3.868	3.271	−1.571	0.637	0.0151
70°	−5.021	0.023	−0.875	0.325	0.0097
80°	−6.724	−8.965	0.37	−0.083	0.0112
90°	−8.693	−58.38	5.446	−1.188	0.0179
100°	−11.17	11.47	−1.131	0.103	0.0217
110°	−12.08	8.716	−0.631	−0.12	0.0069
120°	−11.13	21.8	−2.009	0.098	0.0018
130°	−11.1	1.91	0.15	−0.401	0.0008
140°	−9.719	−0.043	0.243	−0.411	0.0014
150°	−8.417	−0.659	0.147	−0.344	0.0012
160°	−7.437	0.395	−0.178	−0.184	0.0006
170°	−6.783	2.662	−0.671	0.05	0.0006
180°	−6.584	3.387	−0.836	0.131	0.0008

Table 2. Coefficients for Eq. (10) and approximation fitness.

for DC gains, a second-order rational function was fitted as follows to the evolution of G_∞ and f_c along distance at given incidence angles:

$$\tilde{G}_\infty(\theta_{\text{inc}}, \rho) = \frac{p_{12}(\theta_{\text{inc}})\rho + p_{22}(\theta_{\text{inc}})}{\rho^2 + q_{12}(\theta_{\text{inc}})\rho + q_{22}(\theta_{\text{inc}})}, \qquad (10)$$

$$\tilde{f}_c(\theta_{\text{inc}}, \rho) = \frac{p_{13}\rho^2 + p_{23}(\theta_{\text{inc}})\rho + p_{33}(\theta_{\text{inc}})}{\rho^2 + q_{13}(\theta_{\text{inc}})\rho + q_{23}(\theta_{\text{inc}})}. \qquad (11)$$

Note the choice of a second-order numerator that allows greater flexibility in the approximation of the central frequency behaviour, which is more complex with respect to that of gains. Table 2 and Table 3 summarize fitness scores and parameters' values for each of the two functional approximations.

The approximation of G_∞ is overall excellent, never exceeding a mean RMS error of 0.04 dB in the considered angular directions. Similarly, the approximation provided by $\tilde{f}_c$ yields a mean RMS error that is below the actual frequency resolution of 10 Hz for almost 70% of the considered incidence angles. Again, an interpolation of adjacent polynomials analogous to the one used for the DC gain is required to model parameters $\tilde{G}_\infty$ and $\tilde{f}_c$ for intermediate angular values.

θ_{inc}	p_{13}	p_{23}	p_{33}	q_{13}	q_{23}	RMS [Hz]
0°	0.457	−0.668	0.174	−1.746	0.699	1.19
10°	0.455	0.142	−0.115	−0.01	−0.348	0.92
20°	−0.87	3404	−1699	7354	−5350	3.36
30°	0.465	−0.913	0.437	−2.181	1.188	7.01
40°	0.494	−0.669	0.658	−1.196	0.256	19.14
50°	0.549	−1.208	2.02	−1.59	0.816	30.67
60°	0.663	−1.756	6.815	−1.296	1.166	21.65
70°	0.691	4.655	0.614	−0.889	0.76	60.32
80°	3.507	55.09	589.3	29.23	59.51	29.59
90°	−27.41	10336	16818	1945	1707	36.16
100°	6.371	1.735	−9.389	−0.058	−1.118	4.54
110°	7.032	40.88	−44.09	5.635	−6.18	2.53
120°	7.092	23.86	−23.61	3.308	−3.392	2.72
130°	7.463	102.8	−92.27	13.88	−12.67	2.33
140°	7.453	−6.145	−1.809	−0.877	−0.19	2.9
150°	8.101	−18.1	10.54	−2.23	1.295	5.28
160°	8.702	−9.05	0.532	−0.96	−0.023	2.15
170°	8.925	−9.03	0.285	−0.905	−0.079	3.71
180°	9.317	−6.888	−2.082	−0.566	−0.398	3.87

Table 3. Coefficients for Eq. (11) and approximation fitness.

4. A MODEL FOR NFTF RENDERING

The analysis performed in the previous section allows the straightforward construction of a filter model for distance rendering, that can be easily integrated with an infinite-distance spherical model of the head following one of the implementations available in the literature [3]. In fact, if the latter is modeled through a filter $H^{\infty}_{\text{sphere}}$ that takes the incidence angle θ_{inc} as input, the information given by the NFTF can be provided by a cascade of a multiplicative gain G_0 and a shelving filter H_{sh}.

The general filter structure is sketched in Fig. 5. Based on distance and incidence angle information, the "Parameter Extraction" computation block linearly interpolates functions $\tilde{G}_0$, $\tilde{G}_\infty$ and $\tilde{f}_c$ using Eq. (6), Eq. (10), and Eq. (11) respectively; afterwards, $\tilde{G}_0(\theta_{\text{inc}}, \rho)$ is used as multiplicative factor whereas $\tilde{G}_\infty(\theta_{\text{inc}}, \rho)$ and $\tilde{f}_c(\theta_{\text{inc}}, \rho)$ are feeded as parameters to the shelving filter.

A crucial question is the overall goodness of model H_{dist}, that is, whether all the introduced approximations objectively unsettle the magnitude response of original NFTFs as computed through Eq. (3) and Eq. (4). The quality of the approximation offered by the model is attested through a measure of spectral distortion widely used in recent literature [7]:

$$SD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(20\log_{10}\frac{|H(f_i)|}{|\tilde{H}(f_i)|}\right)^2} \quad \text{[dB]}, \qquad (12)$$

where H is the original response (here H_{NF}), $\tilde{H}$ is the reconstructed response (here H_{dist}), and N is the number of available frequencies in the considered range, that in this case is limited between 100 Hz and 15 kHz. The error measure was calculated either for spatial locations that were considered during the analysis process and new ones, thanks to the functional representation and interpolation of the key parameters. Specifically, the magnitude of H_{dist} was computed via

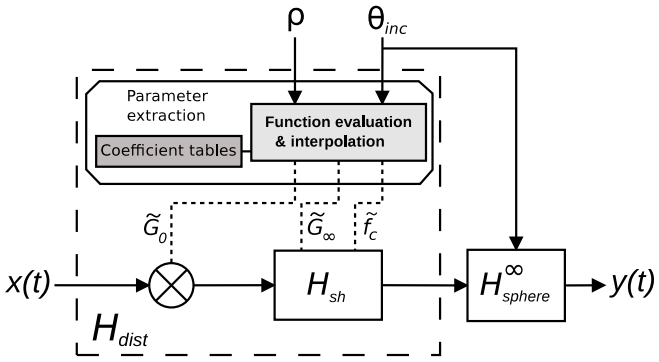

Fig. 5. A model for a spherical head including distance dependence.

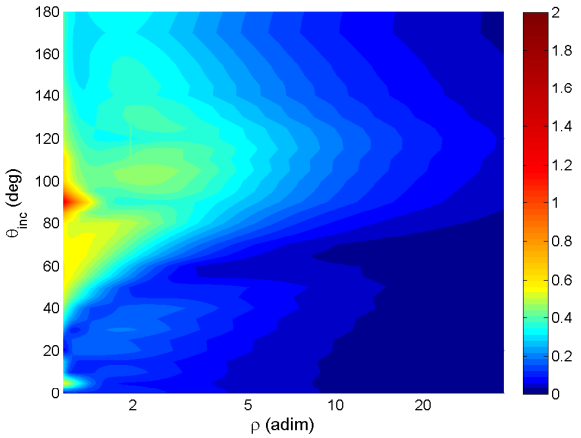

Fig. 6. Spectral distortion introduced by model H_{dist}.

the model for the usual 250 distances, this time at 5-degree angular steps ($\theta_{\text{inc}} = 0°, 5°, 10°, \ldots, 180°$), and compared to the magnitude response of the corresponding original NFTFs up to 15 kHz. The distance-dependent spectral distortion plot for the 37 considered incidence angles is shown in Fig. 6.

Notice that the overall fitness of the approximation is excellent in almost the whole near field, being the SD lower than 1 dB in all of the considered source locations except for the very nearest ones around $\theta_{\text{inc}} = 90°$. The evidenced discrepancy is well explained by the frequency behaviour of the normalized NFTF at these positions which is halfway between that of a shelving filter and of a lowpass filter. Also note how there is no evident relative SD increase between reconstructed NFTFs for angles that were considered in the analysis process and for interpolated angular directions, except for a small dump at $\theta_{\text{inc}} = 5°$ and $\rho < 1.5$. As a consequence, linear interpolations of the key coefficients are already effective as they are, not needing to be improved through higher-order interpolations and/or a denser sampling along the angular axis during the analysis process.

Finally, the almost null SD for high ρ values indicates that near-field effects gradually dissolve with distance just like in the analytical NFTF, i.e. the contribution of H_{dist} tends to 1 as $\rho \to \infty$. This result attests the validity of the above model for the whole spatial range, including the far field.

5. CONCLUSIONS AND OPEN ISSUES

A first-order filter model for near-field effects rendering, thought for real-time binaural listening applications, was studied in this paper. The fit to analytical responses provided by the model was objectively proved to be overall appropriate; experimental evaluations on its subjective effectiveness are a subsequent necessary step.

Further work should take into consideration alternative filter structures to the single, first-order shelving filter, such as a higher-order shelving filter or a lowpass filter realization allowing slope control for contralateral positions. Also, if one remains within the assumption that ITD does not change with distance, the design of an all-pass section counterbalancing the effect that the shelving/lowpass filter's phase response has on ITD needs to be carried on. Last but not least, the choice of the far-field head filter to be coupled with the distance rendering model will turn out to be pivotal for a good STF approximation.

6. REFERENCES

[1] P. Zahorik, D. S. Brungart, and A. W. Bronkhorst, "Auditory distance perception in humans: a summary of past and present research," *Acta Acustica united with Acustica*, vol. 91, pp. 409–420, 2005.

[2] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," *J. Acoust. Soc. Am.*, vol. 104, no. 5, pp. 3048–3058, November 1998.

[3] C. P. Brown and R. O. Duda, "A structural model for binaural sound synthesis," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 5, pp. 476–488, 1998.

[4] W. M. Rabinowitz, J. Maxwell, Y. Shao, and M. Wei, "Sound localization cues for a magnified head: Implications from sound diffraction about a rigid sphere," *Presence*, vol. 2, pp. 125–129, 1993.

[5] S. Spagnol and F. Avanzini, "Real-time binaural audio rendering in the near field," in *Proc. 6th Int. Conf. Sound and Music Computing (SMC09)*, 2009, pp. 201–206.

[6] U. Zölzer, Ed., *Digital Audio Effects*, J. Wiley & Sons, New York, NY, USA, 2002.

[7] M. Otani, T. Hirahara, and S. Ise, "Numerical study on source-distance dependency of head-related transfer functions," *J. Acoust. Soc. Am.*, vol. 125, no. 5, pp. 3253–3261, May 2009.