A BOTTLE MODEL FOR HEAD-RELATED TRANSFER FUNCTIONS

Bradley S. Ferguson, Robert E. Bogner, Steve Wawryk*

Dept. of Electrical and Electronic Engineering
The University of Adelaide, Adelaide 5005, South Australia
and
* Cooperative Centre for Sensor Signal and Information Processing
bogner@eleceng.adelaide.edu.au

ABSTRACT

We describe a parsimonious model for the direction-dependent transfer function of the pinna. The model describes the transfer function with reference to resonators located in particular physical positions relative to the ear canal. The purpose of the work is to provide a parametric model that permits identification with moderate data-gathering, and filter specification for any direction without the need for interpolation of responses.

1. INTRODUCTION

We set out to produce a parametric model that would be capable of describing the head-related transfer function (hrtf) of the pinna for arbitrary directions, based on observations in a finite number of directions. The initial objective was to produce sound maps of the environment for use by blind persons, but it has emerged that entertainment and ergonomic displays are likely to be the major applications. It was considered desirable to include in the model proper phase or time-domain behavior as there is some belief that such information can be important perceptually for some classes of signals, although this is a question that needs to be settled.

Earlier work, [Kistler and Wightman, 1992], depends on parameterising principal component models of the spectral transformations with minimum phase. The main purpose of such models is apparently that they provide an economical method for defining approximate filters for each of the directions for which measurements have been made. The [Kistler and Wightman, 1992] model is used for approximating the measured transfer function from a few coefficients, but not for interpolating. The interpolation law for the coefficients may not be simple. A significantly different approach which is somewhat better related to the physics of the situtaion is represented in [Chen et al., 1992] where the hrtf is considered to come about from multiple scattering. A more-recent work [Blommer and Wakefield, 1997] uses pole-zero modelling, and discovers that some 40 poles and 40 zeros were needed. That paper uses a logarithmic spectral criterion, which appears sensible from a perceptual viewpoint, and it retains phase details.

We approached the problem by considering that a pole-zero model might be efficient because the system is linear and time-invariant (LTI). We noted that the poles of any LTI system are not influenced by the pattern of excitation, but the zeros may be. In the directional hearing case the direction of arrival would influence the strengths and phases of the pole excitations. We proposed that the migration of the zeros with direction of arrival might be described by moderately-simple nonlinear functions while the poles remained fixed. Several approaches to the pole-zero identification were attempted with poor results until we realized that a key aspect of the directional dependence is the time delay of the excitation of the poles.

Fig.1 illustrates the fact that separate pole systems excited with different delays can produce complicated zero effects that may well escape conventional pole-zero modelling. In general, for a sampled-data representation, the number of zeros corresponds to the order of the numerator of the transfer function, which cannot be less than the differential delay between superimposed time components. For a continuum system the delays would correspond to fractional sampling intervals and the result is, to say the least, not convenient. It is the locations of the zeros of the head-related transfer functions which constitute the all-important directional information that the pinna imposes on the acoustic signal and so we need proper modelling methods.

This work was supported by CSSIP and the University of Adelaide

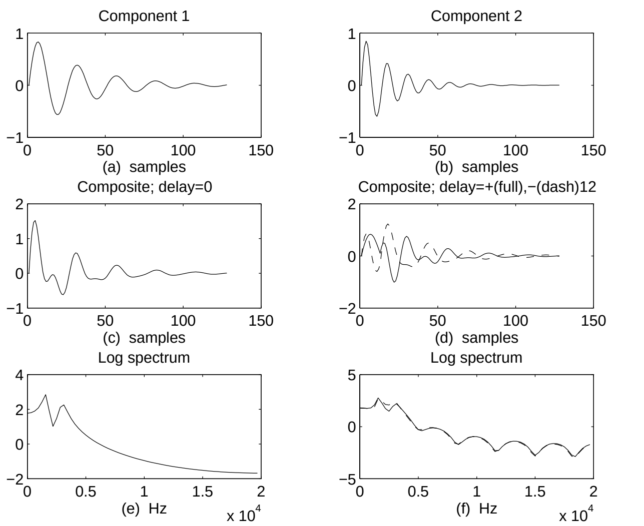

Figure 1: Effects of damped cosines added with delays. (a), (b) pole responses; (c) superimposed, no delay; (d) superimposed with +12 (full line) and -12 (dashed) sample delays; (e) logarithmic magnitude of transfer function with no delay; (f) logarithmic magnitude of transfer function of composite with relative delays. the full and the dashed lines correspond between (d) and (f).

2. RECOGNITION OF A BOTTLE MODEL

In an effort to recognise trends in the data we displayed the unit-pulse time-series of the direction-related impulse responses for sources on the saggital plane at various elevations in several ways. When the data were displayed as an image of intensity (Fig. 2(top image)) we noticed repetitions of arc-shaped patterns, somewhat akin to the ripples on a pond, and that there appeared to be several with different frequencies. The data for this figure are the first 80 samples following detection of a wave of significant amplitude, ie. they are experimental data superficially aligned at the apparent commencement of activity. Each row in each image is a time-series at a particular direction.

Since the data were time-series, we tried separating the patterns by filtering the data into likely bands. The lower four images in Fig. 2 show early results of this process; the bands were chosen simply by counting the ripples and relating them to the repetition frequency, and selecting Fourier components centred on each guessed frequency. The ripple patterns are seen to be separated.

A model that is consistent with these observations is shown in Fig.3 in a simplified 2-dimensional form. In this "bottle" model there are Helmholtz resonators

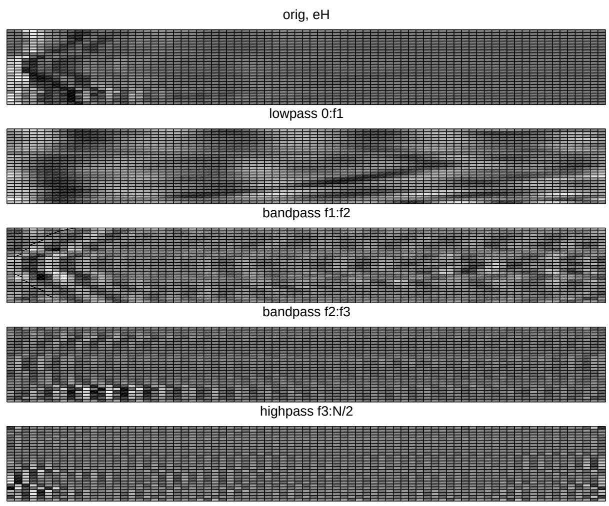

Figure 2: Image presentation of data. Time is from left to right, 80 samples at 44100 Hz sampling rate. Each of the 27 rows corresponds to an angle of elevation, from -40 deg. to 220 deg. at 10 deg. intervals. Top image: original data; lower images are low-pass,and then progressively higher frequency bands.

with their necks located at specific positions in space such that the times at which they are excited depend on the directions of arrival of the field from a distant source. The model neglects diffraction effects as such, but it may approximate them through appropriate poles. The resonances of the ear canal constitute a filter that is convolved with the aggregate of arrivals, and so the poles of the canal appear in all transfer functions. The curvature of the patterns corresponds to the pattern of delay. We see that each delay, for example, T_1, is a sinusoidal function of the direction of arrival, and the amplitude of the sinusoid is equal to the distance of the relevant bottle from the nominal centre of the rotation (which may or may not be the portal of the ear canal).

3. IDENTIFICATION OF BOTTLE MODEL

To fix ideas, consider the simple model shown in Fig.4. This shows a bottle B_k located in the plane of the arrival at position $r_k\angle\theta_k$, and the direction of arrival of the signal from the source is θ_a. The time of arrival at the ear canal of a signal re-transmitted by that bottle is

$$T_{k,a} = \frac{r_a}{c} - \frac{r_k}{c}\cos(\theta_a - \theta_k) + T_k \qquad (1)$$

where r_a is the distance of the source from the centre of rotation, c is the velocity, and T_k is the delay from the kth bottle to the canal portal. Clearly this model

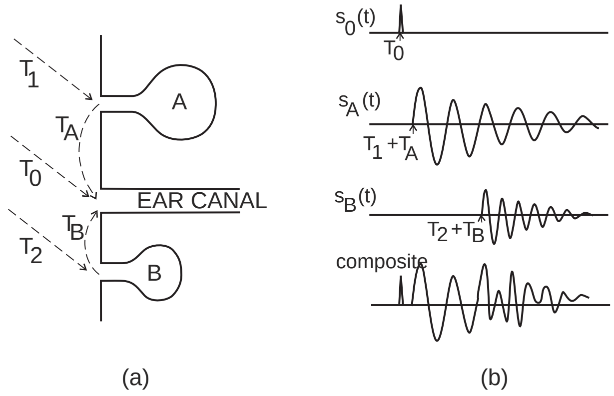

Figure 3: Principle of bottle model. (a) Simple geometric model with two "bottles" (resonators), A and B, showing how signals are received at the ear canal with delays T_0, $T_1 + T_A$ and $T_2 + T_B$, dependent on the direction of arrival. (b) Contributions $s_0(t), s_A(t)$, and $s_B(t)$ and their sum at the ear canal.

displays sinusoidal variation of any resonant response in the ear canal itself; such is accommodated by setting T_k= 0 for the direct path. The important components of this equation for identifying the locations of the bottles are r_k and θ_k.

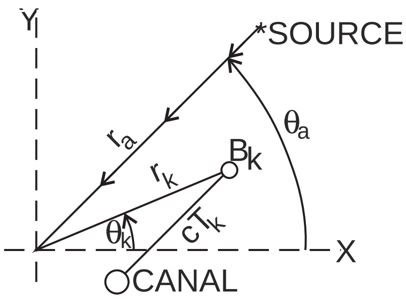

Figure 4: Geometry for kth coplanar bottle B_k and source.

Thus we simply need to identify the r_k and θ_k for each bottle, and since we may be able to constrain the bands to have but one bottle per band this is not a difficult task.

4. THE IDENTIFICATION PROCESS

1. The relevant bands were chosen initially by aggregating the magnitudes of the discrete Fourier transforms (DFTs) of the impulse responses for many directions; some such spectra are shown in Fig.5. Since the poles are unchanged, this process reveals the peaks corresponding to the poles. Actually the sum was smoothed with a rectangular window of 6 samples before peak detection. A later step explains how the bands were refined.

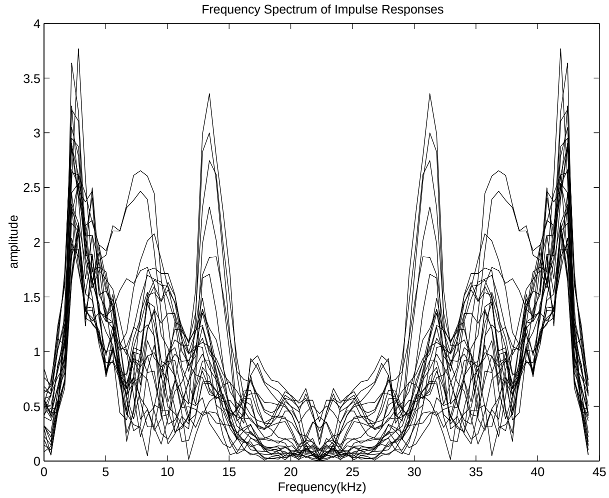

Figure 5: Several spectra as used for initial identification of relevant frequency bands.

2. For each peak a band centred at that frequency was chosen. Then for each direction of arrival the corresponding DFT coefficients were retained and used to synthesize the constrained unit pulse response.

3. In each band a reference waveform was derived by aligning the unit pulse responses using a correlation method. Actually some de-damping was introduced as the band responses tend to be highly damped and the correlations were not well-defined without it. Then from the aggregate of the aligned responses a second-order all-pole system was identified. The correlation coefficient between this second-order response and the optimally-delayed signals from the band was used as an indicator of the merit of the choice of band, and used to refine the choice of band cut-off frequencies. For the data we used the bands were: 1.1 to 6.6 kHz. 6.6 to 13.2 kHz, and 12.1 to 18.7 kHz, but of course these would be dependent on the pinna used. After optimisation of the bands, the resultant second-order response was used as the band reference for identification of the delay for each direction in that band.

4. Then a sinusoid was fitted to the resultant pattern of delays vs angle, as shown in Fig.6. the important parameters of this fitted sinusoid are the amplitude and phase; the frequency is of course one cycle per 2π radians of elevation.

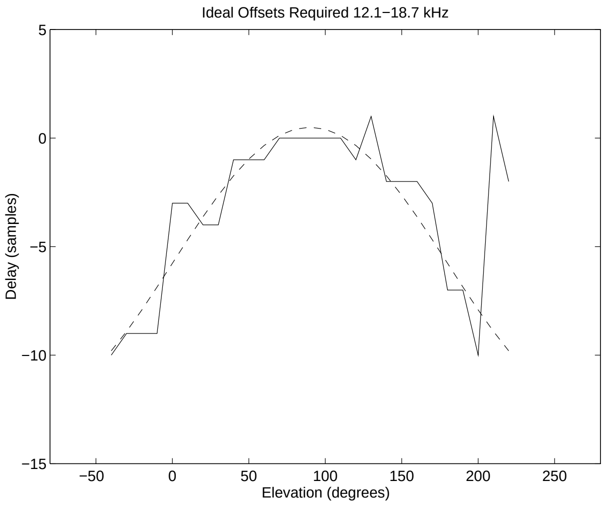

Figure 6: Sinusoid fitted to the pattern of delay vs angle, for the frequency band 12.1-18.7 kHz.

5. From the θ_k and r_k values the locations of all bottles relative to some centre are readily found. An example of the result is shown in Fig.7 which shows positions identified for the dominant four bottles. In fact these positions were found in a 3-dimensional modelling. It is noted that the positions do not coincide with an obvious mechanical bottle formation; it is to be expected that resonances occur not just in bottle structures, but also as standing waves between scatterers. Of course we then expect the coupling to incident fields to be direction-dependent, too.

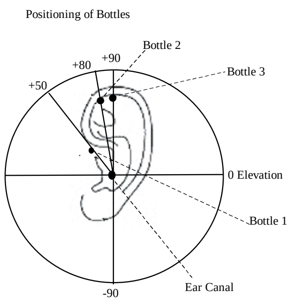

Figure 7: Positions identified for the dominant three bottles.

5. CONCLUSIONS

The novel model describes the directional-related transfer functions in terms of resonators placed at appropriate physical positions. This model is economical in the number of parameters, is amenable to simple calculation of the directional impulse response at any direction. It appears likely that the model can be identified economically from data taken at a few positions; if so, then the model should have great advantages over other models. Further investigation is needed to determine whether incorporation of magnitude factors for the contributions will be advantageous. We do not pretend that the system modes are completely described by poles, as the physical form is really distributed and the effects are probably better attributed to interactions between somewhat distributed scatterers, but this is a matter of degree. Furthermore, the coupling to incident fields to be direction-dependent, too.

6. ACKNOWLEDGEMENTS

It is a pleasure to acknowledge stimulating discussions with David Bies and Sylvan Elhay of The University of Adelaide, Simon Oldfield of DSTO, Brian Connolly and Andrew Reilly of Lake DSP Pty Ltd.

7. REFERENCES

[Chen et al., 1992] Jiashu Chen, Barry D. Van Veen and Kurt E Hecox, "External ear transfer function modelling: A beam-forming approach" J. Acoustical Society of America V.92 N.4 pp.1933-1944, Dec. 1992.

[Kistler and Wightman, 1992] Doris J. Kistler and Frederic L. Wightman "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction" J. Acoustical society of America V.91 N.3 pp.1637-1647, March 1992

[Blommer and Wakefield, 1997] M. A. Blommer and G. H. Wakefield, "Pole-zero approximations for head-related transfer functions using a logarithmic error criterion", *IEEE Trans. On Speech and Audio Processing* V.5 N.3, May 1997, pp.278-287.