A Filter Based Approach to Sound Source Simulation Through an Outward Facing Spherical Array of Loudspeakers

Calum Armstrong[1] and Jude Brereton[2]

[1]AudioLab, University of York, York, UK
e-mail: calum.armstrong@alumni.york.ac.uk
[2]AudioLab, University of York, York, UK
e-mail: jude.brereton@york.ac.uk

September 23rd 2016.

Abstract

Presented is the design and implementation of an independent 17-point spherical loudspeaker array functional as a sound source simulator. A novel filter based approach is used to dynamically modify the spectral balance of the audio signals feeding the array in an attempt to simulate the acoustic radiation patterns of a human head. Standardised test methods are further developed to measure sound source directivity with the use of automatic analysis software. It is shown that this frequency domain simulation method, coupled with an effective loudspeaker array, can result in a better approximation of human-like directivity compared to current sound source simulators used within virtual acoustics simulations.

1 Introduction

The directivity of a sound source is defined as the frequency dependent 3-dimensional distribution of acoustical energy radiated from a source during active emittance. However, despite multiple papers concerning the variable directivity patterns of the human voice [1, 2, 3, 4, 5] (both with regards to frequency output and phoneme production) little work has been undertaken in the production of a sound source capable of simulating such patterns both accurately and dynamically in an attempt to replicate a human source.

The results presented here demonstrate a method of dynamic human sound source simulation with a view to improving upon the validity of Room Impulse Response (RIR) capture processes and further more general sound source applications within the up and coming Virtual Reality industry. Improving upon the accuracy and adaptability of current static human sound source simulators (Head and Torso Simulators) the proposed method takes into account recent publications of phoneme dependent directivity and therefore the need for real-time manipulation of directivity patterns for human-like simulation.

Human directivity patterns were measured for 3 phonemes: /u:/, /i:/, and /a:/ (as defined by the International Phonetic Alphabet). A single result was simulated under the proposed methodology. Simulated directivity pattern measurements were then taken and compared to the original findings.

2 Related Work

A pioneering study by Farnsworth and Dunn (FD) [1] examined the directivity of the human voice over 15s of extended speech, finding greater degrees of directional variation with an increase in frequency. Similar experiments were later repeated with similar results by Flanagan [6], making use of a dummy mannequin and oral transducer. Kuhn [7] discussed how the shape of the human head, torso, and pinna all contribute to an individual's directivity whilst differences in directivity due to phoneme were later addressed by Meyer [2], concluding that noticeable changes in vocal timbre will ensue with just a 40° rotation of a human sound source. Phoneme dependent directivity has since been studied further and in much greater detail by Katz [3, 8] supporting absolutely the dependency of directivity on phoneme as well as an individual's characteristics.

Chu and Warnock (CW) [9] examined in more detail the directivity characteristics of extended human speech relating its importance to the design of open plan offices. Results were analysed over 1/3 octave frequency bands and appear consistent with those gathered by Flanagan and FD, although no direct comparison was made. Most recently, in 2012, Monson and Story [4] conducted a study comparing spoken and sung performances of various styles from various performers. They concluded that no differences in directivity were present when comparing modes of production (i.e. spoken or sung) and only limited differences were found when comparing genders and production levels (i.e. loud or quiet). Once again large differences in directivity were, however, found dependent of phoneme, supporting the work of Katz. It should be noted, however, that a paper by Cabrera, Davis and Connolly [5] contradicted these findings to some extent, showing that at least for opera singers, directivity can vary significantly between soloists, although they agreed that performance style made little difference.

Included in the study conducted by CW [9] and further introduced by Bozzoli and Farina [10, 11] and Halkosaari, Vaalgamaa and Karjalainen (HVK) [12] are comparisons of real human directivity data and data gathered from HATS (Head and Torso Simulators). Specifically, Brüel & Kjær (B&K) mouth simulators of types: unknown (CW), 4128 and 4130 (BF) and 4128 and 4227 (HVK) were studied. Conclusions show in general similar results with regards to directivity; however, differences were shown, at times, to exceed 5dB (BF) and differences between transfer functions measured between point receivers by HVK were seen as high as 14dB. Similarities are best noted at lower frequencies ($<\approx$5,000Hz) in the forward direction whilst peak differences are seen as frequency increases and measurements are taken from within the rear hemisphere of radiation. In general, HATSs have been measured as being more directional that a real human, meaning a greater concentration of radiated energy is seen output in the forward direction.

It is important to note, however, that when comparing humans and HATSs only the average directivity of extended speech has ever been considered. Results obtained by Meyer [2] and Katz [3] both show significant changes in directivity with phoneme. Therefore, although on average the directivity of a HATS was found to be similar to that of a human being, the directivity of any individual phoneme may still be revealed to be vastly unrepresentative. It is in fact likely that due to the fixed dimensions of a HATS their directivity will remain far more constant than that of a real human, given the alterations in mouth aperture and jaw position during speech.

In more recent years (late 1990s onwards) there have been several papers published [13, 14, 15, 16, 17] that present the use of spherical arrays of transducers as sound source simulators. Whilst promise is shown in the potential of these types of simulators, the papers do not discuss in depth the simulation / attempted replication of directivity patterns. Although noted as an application area [17, 14], focus is instead drawn to the mathematical modelling and simulation of predetermined patterns. The current research discusses far more the control of directivity rather than the practical simulation of it. For example, without exception each paper found on the topic discussed directivity modelling with regards to spherical harmonics despite current directivity pattern measurement techniques not including the analysis of such.

3 Directivity Measurements

3.1 Definitions

Two terms were defined to aid with the classification of data. Angle Dependent Frequency Spectra (ADFS), Frequency Dependent Source Directivity (FDSD).

ADFS were defined as a set of frequency response measurements made around a source. They describe the variance of a sound source's observed frequency magnitude output as a result of a change in angle of observation. Alternately, FDSD was defined as the directivity of a sound source, given an excitation signal of specific frequency. Both concepts stem from the knowledge that any real source will radiate sound unevenly, and complete knowledge of one implies complete knowledge of the other - they are simply alternative mappings of the same data. Both a sound source's ADFS and FDSD are individual and depend on the layout / mechanical qualities of the source as well as any acoustical absorption / blocking provided by the source's physical structure.

To conform with standard protocols and to ease in the plotting of data, a third term - Frequency-banded Angular Directivity (FAD) - was defined as the average directivity of a sound source within a particular frequency band. The bulk of data analysis was handled in this format and allowed for comparison with previous studies.

3.2 Method

FDSD patterns were measured and averaged over 10 octave frequency bands with centre frequencies ranging from 31.5Hz to 16,000Hz. Analysis was limited to relative directivity only i.e. only the differences in ADFS relative to the frequency response of the source observed from an on-axis observation point were considered in FAD plots. These were refereed to as relative ADFS.

The relative ADFS of 69 angles of observation over 3 orthogonal planes (as described in Fig. 1) were computed. Measurements were made at 10° increments about the horizontal plane and 20° increments about each of the two vertical planes. For each angle tested a pair of simultaneous recordings were taken during source excitation: an on-axis “reference” recording taken from an angle of $\phi = 0$; $\theta = 0$ and a “test” recording taken from the angle under test, see Fig 2. Recordings were captured with identical Earthworks M30 measurement microphones and were taken inside an anechoic chamber. An example of the experimental set-up can be seen in Fig. 3.

The Long Term Average Spectrum over the entire active region of each recording was calculated and the reference recording spectrum subtracted from the test recording spectrum for each simultaneously recorded pair. The resulting differences in the ADFS were then considered the relative ADFS of the source. This method alleviated the hardware demanding requirement to make 69 simultaneous recordings whilst tolerating the inevitable unreliability of repeated excitation signals in the cases of a human subject. By compensating each test ADFS measured with respect to an on-axis recording of the source's output during the corresponding isolated recording a reliable set of data was gathered. The directivity of the source could then be obtained from the differences in magnitudes of different frequency components seen across the multiple angles of observation.

Throughout each measurement the sound source under test was required to produce a frequency-component rich sound in order to accurately measure the directivity of the source over as wide a range of frequencies as possible. A digital source output computer generated white noise but this was not possible to obtain via a human subject as sound source. Instead, the subject was asked to produce a sine-sweep style vocal glissando for each of the phonemes /u:/, /i:/, and /a:/. The fundamental frequency of the glissando started at 220Hz and extended upwards for two octaves.

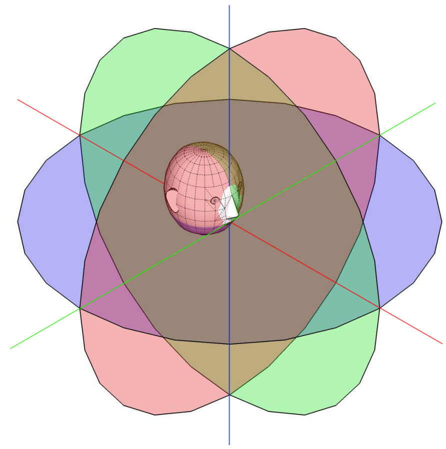


Figure 1: The three orthagonal planes around which ADFS measurements were taken about a source.

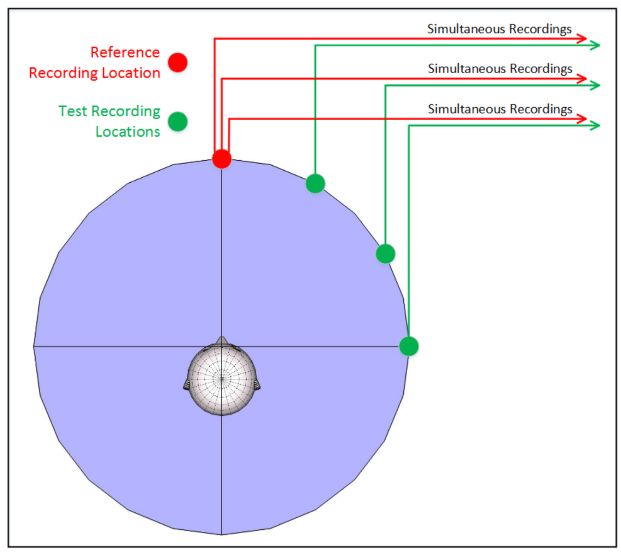


Figure 2: Graphical representation of the simultanious reference recording technique used to obtain sound source directivity measurements.

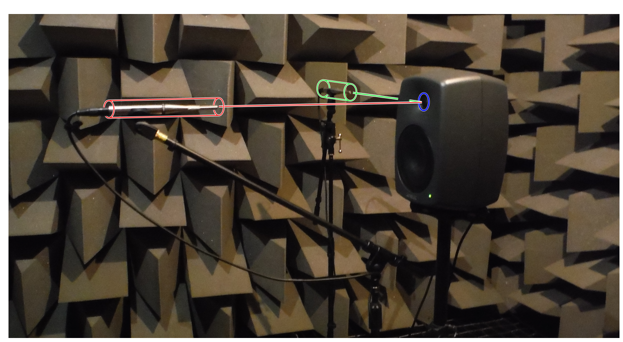


Figure 3: Example of a simultanious reference recording measurement being taken with the reference miciphone highlighted in red.

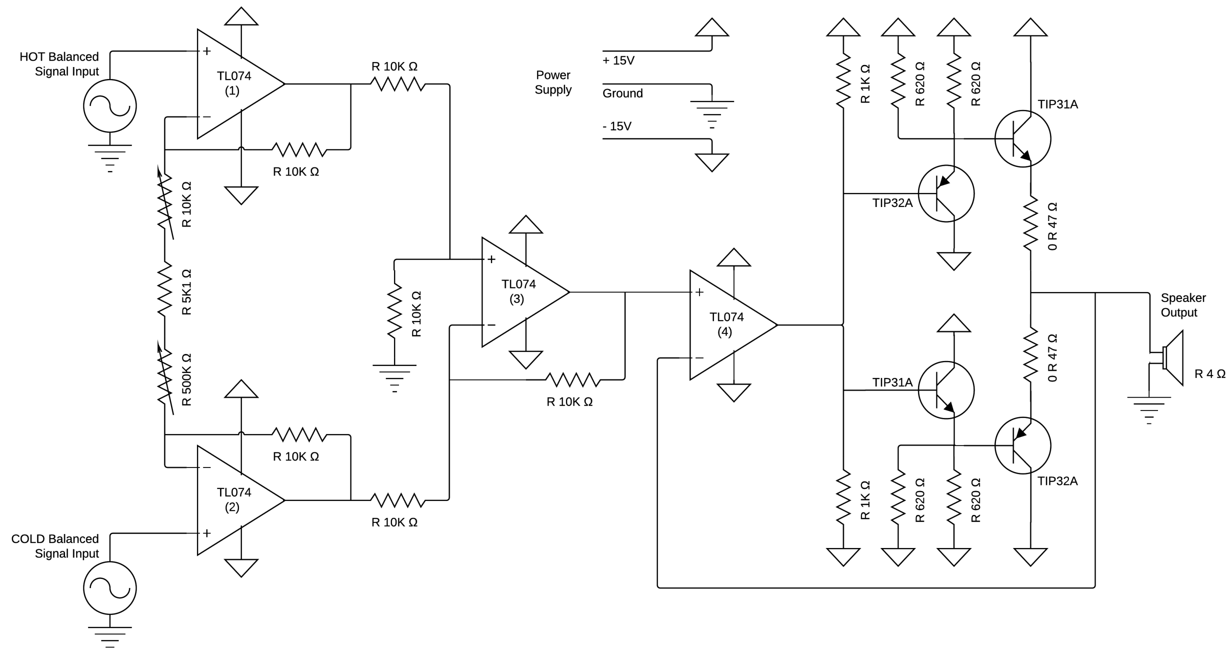


Figure 4: Circuit Diagram of the Class AB Amplifiers installed in the Globe

4 The Globe

4.1 Design

The Globe, as pictured in Fig. 5 and Fig. 6, is an integrated hardware-software solution to sound source simulation. Simply put, it is a spherical independent 17-Point loudspeaker array with built in amplifier and balanced XLR input for each speaker.

During testing, speaker inputs were driven by a multi-output audio interface [18] running alongside professional recording software [19]. Whilst the recording software used was not ideally tailored to the application it provided a quick and easy solution to trialling the filter based methodology.

Figure 5: The Globe - A 17-point sperical loudspeaker array.

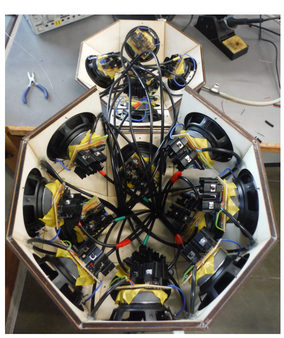


Figure 6: Installation of amplifiers inside the Globe.

Speakers were chosen as low-cost, flat-response transducers of minimal size and weight minimising the overall dimensions of the device. The housing was constructed from low-cost laser-cut plywood, although it is advisable that more thought should be given to the acoustic impact of the housing in future. Amplifier circuits were custom designed, built for purpose and comprised a differential input with voltage gain, a feedback loop, and a high power transistor output. A circuit diagram is given in Fig. 4.

4.2 Filter Based Simulation

Directivity simulation was achieved via inbuilt channel plug-in effects available within the recording software. A master audio file was routed to 17 bus channels, with each bus channel output to a separate physical output on the audio interface. A Graphical Equalizer was then inserted into each bus. By altering the relative frequency response of each output channel feeding the Globe's array overall directivity could be manipulated.

Graphical Equalizer settings were adjusted according to the ADFS data gathered from the human directivity pattern measurements. The angle of each speaker within the Globe's array was considered and a corresponding measurement angle selected from the tests. In cases where no measurement angle approximately matched that of a speaker a simple linear interpolation was used to obtain approximated data. The relative frequency response (ADFS) measured at that angle was then translated to the frequency response of the Graphical Equalizer inserted into the bus channel feeding that speaker.

Using this method, the Globe's relative directivity pattern was governed solely by the frequency responses of the individual Graphical Equalizers inserted into the bus channels. Absolute directivity was, however, a result of both the relative directivity and further the frequency response of the audio file selected for playback on the master channel and output directly (without any filtering) to the front - central speaker on the Globe. Imperfections due to the frequency responses of the individual speakers were at this point ignored as they were not relevant to proof of concept.

By making use of the real-time automation capabilities of modern DAWs, it is possible to adjust bus channel filtering settings in real-time. Exploiting this results in the interactive manipulation of the directivity pattern of the Globe. This can be achieved both prior to, and during a playback / simulation sequence. Further, these alternations can be programmed to occur automatically during playback.

Although time constraints prevented this from being extensively tested, the ability to switch on / off the filtering of audio channels was available and allowed for limited real-time directivity pattern manipulation i.e. switching between an omnidirectional (no filtering) and human-like (filtering) sound source.

5 Results

Fig. 7 shows the results obtained for the upper 7 frequency bands analysed from a variety of FAD measurements. (a-c, f-h, k-m) show the results of individual phoneme tests for a particular human subject. (d, i, n) show the average FAD of the subject over each phoneme measured. (e, j, o) show the FAD measurements made of the Globe during simulation of the results shown in (d, i, n) as discussed in section 4.2.

Results show surprising consistency over the three phonemes measured. Despite previous studies [2, 3, 8, 4] finding large differences it appears as though restricting analysis to relative directivity only diminishes these almost entirely. It may be the case that the differences in directivity measured in previous studies are far more related to the frequency response of the excitation signal than first thought and therefore appear more prominent in absolute directivity measurements.

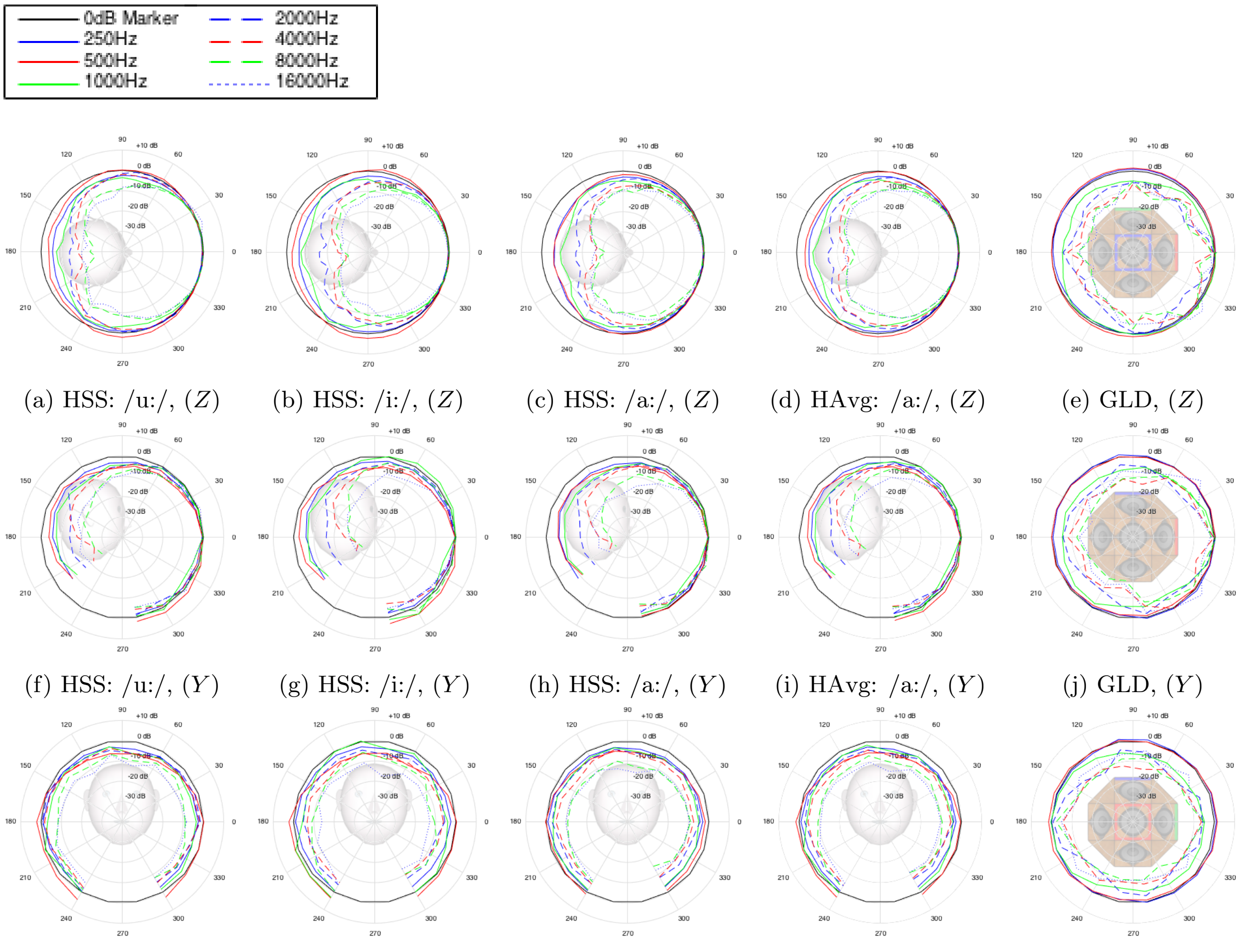
Measurements shown in Fig. 7 (e, j, o) show similarity to the simulated results, however, appear somewhat jerky and undefined. The reason for this can however be explained by the imperfect omnidirectional response of the Globe. The filter based method discussed here assumes that with no filtering applied a perfect omnidirectional response can be simulated. This perfectly flat response is then modified and adjusted to match the directivity pattern being simulated. An imperfect omnidirectional response, however, skews the process. To alleviate this issue and examine only the method of frequency based simulation the omnidirectional response of the Globe was measured and subtracted from the results shown in (e, j, o). The results of this calibration are shown in Fig. 8.

Fig. 8 (a, b, c) shows much higher degrees of similarity to Fig. 7 (d, i, n). Although discrepancies and inaccuracies are present the general patterns across almost all frequency bands are consistent. What is noticeable, however, is an overall lacking in the ability to attenuate frequency bands to the levels required by the simulation. For example, in the rear hemisphere of the Z-Plane angular attenuations in excess of 25dB are seen in the human response whilst simulated attenuations were unable to exceed even 20dB. It is expected that this may be due to crossover interference from neighbouring speakers in the array. It is likely that a more complex algorithm requires developing that considers simulated attenuation levels taking into account the acoustic contribution of all speakers in the array rather than just the speaker facing the direction in question directly.

6 Conclusion

Results obtained from these experiments prompt two separate conclusions to be drawn. The first concerns the Globe as a self-standing sound source simulator whilst the second concerns the greater success of the filter based approach to sound source simulation. Whilst the simulated directivity pattern measurements of the Globe do not currently offer any significant improvement over alternate sound source simulators the method developed does offer significant advantages that may yet prove to be beneficial.

Results shown here indicate that directivity pattern simulation is possible via the proposed method although further work, calibration, and product engineering is required to improve this. Whilst this alone does not offer any improvement over a well calibrated HATS the functionality and possibility of real time directivity pattern manipulation does present a

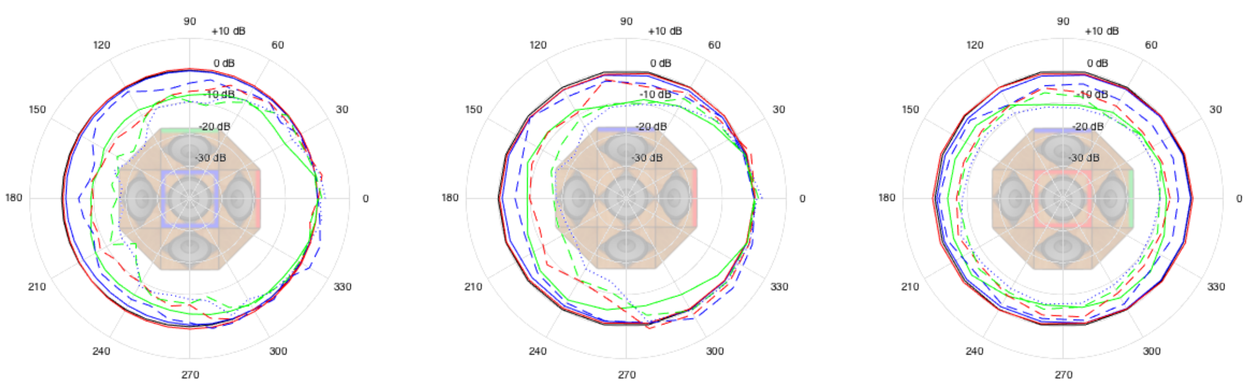


(a) HSS: /u:/, (Z) (b) HSS: /i:/, (Z) (c) HSS: /a:/, (Z) (d) HAvg: /a:/, (Z) (e) GLD, (Z)

(f) HSS: /u:/, (Y) (g) HSS: /i:/, (Y) (h) HSS: /a:/, (Y) (i) HAvg: /a:/, (Y) (j) GLD, (Y)

(k) HSS: /u:/, (X) (l) HSS: /i:/, (X) (m) HSS: /a:/, (X) (n) HAvg: /a:/, (X) (o) GLD, (X)

Figure 7: FAD plots covering the Z- Y- and X-Planes for: three Human Sound Source (HSS) phonemes: /u:/, /i:/ and /a:/; an average HSS result (HAvg); the Globe during simulation of the average HSS results (GLD). Spectral measurements are shown with respect to angle in (dB).



(a) GLODD, Z-Plane (b) GLODD, Y-Plane (c) GLODD, X-Plane

Figure 8: FAD plots covering the Z- Y- and X-Planes for the Globe during simulation of average HSS results after omnidirectional calibration (GLODD). Spectral measurements are shown with respect to angle in (dB).

significant advantage over other techniques and may warrant future consideration of this method.

Further, the placement of simulation techniques within the digital domain (digital filters) rather than the physical domain (head and torso shaped moulds) opens up the ability to achieve the likes of sound source rotation / adaptation far more easily than current techniques. Device maintenance and component improvement also become far easier when the physical properties of the simulator become of less importance.

References

[1] H. Dunn and D. Farnsworth, "Exploration of pressure field around the human head during speech," *The Journal of the Acoustical Society of America*, vol. 10, no. 3, pp. 184–199, 1939.

[2] A. Marshall and J. Meyer, "The directivity and auditory impressions of singers," *Acta Acustica united with Acustica*, vol. 58, no. 3, pp. 130–140, 1985.

[3] B. F. Katz and C. dAlessandro, "Directivity measurements of the singing voice," in *19th International Congress On Acoustics Madrid*. Sociedad Espanola de Acustica, Sep. 2-7 2007.

[4] B. B. Monson, E. J. Hunter, and B. H. Story, "Horizontal directivity of low-and high-frequency energy in speech and singing," *The Journal of the Acoustical Society of America*, vol. 132, no. 1, pp. 433–441, 2012.

[5] D. Cabrera, P. J. Davis, and A. Connolly, "Long-term horizontal vocal directivity of opera singers: Effects of singing projection and acoustic environment," *Journal of Voice*, vol. 25, no. 6, pp. e291–e303, 2011.

[6] J. L. Flanagan, "Analog measurements of sound radiation from the mouth," *The Journal of the Acoustical Society of America*, vol. 32, no. 12, pp. 1613–1620, 1960.

[7] G. F. Kuhn, "The effect of the human torso, head, and pinna on the azimuthal directivity and on the median plane vertical directivity," *The Journal of the Acoustical Society of America*, vol. 65, no. S1, pp. S8–S9, 1979.

[8] B. F. Katz, F. Prezat, and C. dAlessandro, "Human voice phoneme directivity pattern measurements," *The Journal of the Acoustical Society of America*, vol. 120, no. 5, pp. 3359–3359, 2006.

[9] W. T. Chu and A. Warnock, "Detailed directivity of sound fields around human talkers," Tech. Rep. RR-104, National Research Council Canada, Tech. Rep., 2002.

[10] F. Bozzoli, A. Farina, and M. Viktorovitch, "Balloons of directivity of real and artificial mouth used in determining speech transmission index," in *118th Audio Engineering Society Convention*. Audio Engineering Society, May 28-31 2005.

[11] F. Bozzoli and A. Farina, "Directivity balloons of real and artificial mouth simulators for measurement of the speech transmission index," in *115th Audio Engineering Society Convention*. Audio Engineering Society, Oct. 10-13 2003.

[12] T. Halkosaari, M. Vaalgamaa, and M. Karjalainen, "Directivity of artificial and human speech," *Journal of the Audio Engineering Society*, vol. 53, no. 7/8, pp. 620–631, 2005.

[13] R. Avizienis, A. Freed, P. Kassakian, and D. Wessel, "A compact 120 independent element spherical loudspeaker array with programable radiation patterns," in *120th Audio Engineering Society Convention*. Audio Engineering Society, May 20-23 2006.

[14] F. Zotter, H. Pomberger, and A. Schmeder, "Efficient directivity pattern control for spherical loudspeaker arrays," *Journal of the Acoustical Society of America*, vol. 123, no. 5, p. 3643, 2008.

[15] A. M. Pasqual, P. Herzog, and J. R. de França Arruda, "Theoretical and experimental analysis of the electromechanical behavior of a compact spherical loudspeaker array for directivity control," *The Journal of the Acoustical Society of America*, vol. 128, no. 6, pp. 3478–3488, 2010.

[16] A. M. Pasqual, A. de França, J. Roberto, and P. Herzog, "Application of acoustic radiation modes in the directivity control by a spherical loudspeaker array," *Acta acustica united with Acustica*, vol. 96, no. 1, pp. 32–42, 2010.

[17] H. Pomberger, "Angular and radial directivity control for spherical loudspeaker arrays," Diploma Thesis, University of Music and Dramatic Arts, 2008.

[18] Saffire Pro 40. Product Website. Focusrite. [Online]. Available: https://us.focusrite.com/firewire-audio-interfaces/saffire-pro-40 [Accessed: May 06 2016]

[19] Logic Pro X. Product Website. apple. [Online]. Available: http://www.apple.com/uk/logic-pro/ [Accessed: May 06 2016]