MEASURING THE DIRECTION OF EARLY REFLECTIONS.

T. W. Roberts Acoustics Research Centre, University of Salford, Salford M5 4WT, UK.
M. R. Avis Acoustics Research Centre, University of Salford, Salford M5 4WT, UK.

1 INTRODUCTION

Many devices to measure the direction of room reflections have been described, but matching the performance of the human auditory system remains a challenge. Techniques based on beamforming or sound field decomposition do provide unambiguous results but require an impractical number of microphone positions. Other techniques based on intensity or time of flight (TOF) do much better in this regard but at the cost of ambiguities in certain situations such as the simultaneous arrival of reflections. TOF techniques are of practical interest as only four microphones are required in order to extract the azimuth and elevation of a given reflection. The performance of a tetrahedral microphone array with TOF analysis is studied in detail and the conditions where unambiguous results can be obtained are discussed.

2 MINIMUM AUDIBLE ANGLE

The *Minimum Audible Angle* (MAA) is the angle at which two identical sounds can just be resolved by a listener as being separate. The MAA is a function of both the frequency (or spectral content) of the sounds and the direction from which they arrive. In humans, the best case MAA is of the order of 1° for pure tones between 500 Hz and 750 Hz presented from directly in front of the listener[1]. While the worst case MAA can be large and even indeterminate, any device to measure the direction of reflections for the study of human perception should arguably strive to do at least as well as if not better than the best case performance of the human auditory system at a given frequency.

3 AVAILABLE TECHNIQUES

The most obvious method for measuring the direction of an isolated steady state sound in anechoic conditions is to use a high directivity receiver scanned in azimuth and elevation until a maximum signal level is found. The direction of the device at the maximum indicates the direction of the source. For a source in a non-anechoic environment, sound energy will also arrive at a receiver position from directions other than the source. A directional receiver will be able to observe such sources (according to its directivity) but analysis in time of the impulse response can also be used as a result of the different paths taken by the sound (an hence their different times of arrival).

As an extension to the directivity approach described above, a number of methods make use of multiple transducers of known directivity pointing in different directions. Knowledge of the directivity of each microphone and the signals obtained are used to derive bearings to the source. D'Antonio et al used an arrangement of six cardiod microphones mounted orthogonally in directions labelled front, back, up, down, left and right. Impulse responses were obtained from each microphone with reflections separated using time of flight. A $\left(1+\cos\theta\right)$ polar pattern was then assumed for each microphone with bearings of the sources derived from the relative signal levels obtained in pairs of microphones[2]. Analogous methods have been used with tetrahedral arrays and SoundField® microphones for the measurement of loudspeaker positions[3, 4] and for more general room acoustics studies[5, 6].

The directional characteristics of sound fields can also be studied by measuring the vector intensity $\vec{I}(t) = p(t)\vec{u}(t)$ [7]. Many practical devices are available, a common arrangement employing pairs of phase-matched pressure transducers mounted close together. Euler's equation is used to derive the velocity component from a pair and an array of three pairs can be used to obtain the full vector intensity[8].

A four microphone time of flight technique which does not depend on directivity, but uses a geometrical approach to obtain source positions has also been demonstrated[9]. The original study used omni-directional microphones, with one at the origin of a right handed set of axes, and three mounted the same distance along each of the axes. A similar study mounted the microphones at the vertices of a tetrahedron reporting resolution within 3° for well separated reflections[10].

While high directivity, single element microphones are available[11], devices having higher order directivity can be constructed from 3D arrays of simple elements[12, 13] with associated signal processing based on beamforming[14, 15]. One advantage of such arrays is that the look direction can generally be altered in post-processing rather than having to physically scan the device. Many variations on the basic delay and sum beamforming algorithm can be employed but all have an upper frequency limit imposed by spatial aliasing, and a low frequency limit arising from the fact that the array becomes more omni-directional as the frequency is reduced[14-16]. A number of commercial systems based on beamforming arrays are available[17, 18].

Gover and co-workers describe two 32 element arrays optimised for different frequency ranges with the elements located at the vertices of a pentakis dodecahedron. Excitations are played into the space to be measured and recorded using a multi-channel sound card. Sixty different look directions (orientations of the main lobe of the array) are derived in post-processing allowing a map of sound energy versus direction to be obtained for a single application of the excitation[19-22]. Rafaely an co-workers report measurements using a single microphone sequentially translated to 882 positions on the surface of nested spheres of different radii. At each point an impulse response is obtained with the results processed to provide a map of average power with angle[23].

4 LIMITATIONS

The angular resolution of devices based on directivity is governed by their beam width. If the beam width is large compared to the angular spacing of the sources then the sources will not be resolved. The analogous concept in optics is the well known Rayleigh Criterion which gives the minimum angular separation between two sources that can still be resolved as being distinct. In their operating frequency ranges, the directivity of the 32 element arrays of Gover and co-workers had an FWHM of $\approx 28°$ with the angle between look directions between 22° and 24°[20]. The wide beam width coupled with the widely spaced look directions mean that any fine structure at the scale of the best case MAA in humans is not resolvable. With 882 microphone positions, the measurements of Rafaely and co-workers achieve an angular resolution of $\approx 9°$, still an order of magnitude higher than the best case MAA in humans[23]. This highlights the fundamental problem with beamforming techniques in that a very large number of elements are required in order to approach the angular resolution of the human auditory system. At the time of writing, the largest beamforming microphone array known to the authors is the (planar) LOUD array having 1020 elements[24].

The complexity of beamforming must be weighed up against the fact that all of the other methods outlined suffer from a number of inherent flaws. Gover gives a discussion of these with the fundamental issue for direction of arrival determination being an inability to disambiguate reflections arriving from different directions at the same time[19]. For an environment containing a single source, this situation is likely to arise when the source and receiver are located at points of high symmetry and can usually be mitigated through careful positioning based on prior knowledge of the space. Such prior knowledge is not required by beamforming techniques.

Sekiguchi and co-workers report being able to measure the direction of reflections to within 3° using the four microphone TOF technique[10]. Given that this performance approaches the best case MAA in humans and the low complexity of the approach, it is interesting to study further the factors allowing such accuracy to be obtained.

5 TOF ERROR ANALYSIS

In order to study the potential accuracy of the four microphone TOF technique, an analysis of the measurement errors is required. The equations to obtain the Cartesian components of the position of a source from the times of flight registered at the microphone positions of a tetrahedral array can be derived through simple application of Pythagoras' theorem and knowledge of the properties of the tetrahedron and are given in equations (1.1), (1.2) and (1.3)[10].

$$X = \frac{c^2}{2a}\left(t_3^2 - t_4^2\right) \tag{1.1}$$

$$Y = \frac{c^2}{2a\sqrt{3}}\left(t_3^2 + t_4^2 - 2t_2^2\right) \tag{1.2}$$

$$Z = \frac{c^2}{2a\sqrt{6}}\left(t_2^2 - 3t_1^2 + t_3^2 + t_4^2\right) \tag{1.3}$$

The procedure for combining measurement uncertainties is well known[25, 26]. If a derived result is a function of a number of individual measurements (that is, $Z = f\left(x_0, x_1, \ldots, x_{N-1}\right)$) and the errors on the individual measurements are independent of each other (that is, they have zero covariance) then the variance of the derived measurement σ_Z^2 is given by equation (1.4) where $\sigma_{x_i}^2$ is the variance of the measurement of the variable x_i.

$$\sigma_Z^2 = \sum_{i=0}^{N-1} \sigma_{x_i}^2 \left(\frac{\partial f}{\partial x_i}\right)^2 \tag{1.4}$$

Applying equation (1.4) to equations (1.1), (1.2) and (1.3) results in equations (1.5), (1.6) and (1.7) for the variances of the measurements of the Cartesian components of the source respectively.

$$\sigma_X^2 = \frac{c^4}{a^2}\sigma_t^2\left(t_3^2 + t_4^2\right) + X^2\left(4\frac{\sigma_c^2}{c^2} + \frac{\sigma_a^2}{a^2}\right) \tag{1.5}$$

$$\sigma_Y^2 = \frac{c^4}{3a^2}\sigma_t^2\left(4t_2^2 + t_3^2 + t_4^2\right) + Y^2\left(4\frac{\sigma_c^2}{c^2} + \frac{\sigma_a^2}{a^2}\right) \tag{1.6}$$

$$\sigma_Z^2 = \frac{c^4}{6a^2}\sigma_t^2\left(9t_1^2 + t_2^2 + t_3^2 + t_4^2\right) + Z^2\left(4\frac{\sigma_c^2}{c^2} + \frac{\sigma_a^2}{a^2}\right) \tag{1.7}$$

For sources close to the array, the X, Y and Z components along with the flight times are all small so the overall error will be small compared to sources that are far from the array. For sources far from the array, the individual flight times will increasingly approximate the flight time from the source to the centre of the array, t (that is, $t_1 \to t, t_2 \to t, t_3 \to t, t_4 \to t$). Defining R as the distance from

the centre of the array to the source (such that $R = ct$) leads to equation (1.8) (where $\varepsilon_i = \{X, Y, Z\}$).

$$\left(\frac{\sigma_{\varepsilon_i}}{\varepsilon_i}\right)^2 = 2\left(\frac{cR}{a\varepsilon_i}\right)^2 \sigma_t^2 + 4\left(\frac{\sigma_c}{c}\right)^2 + \left(\frac{\sigma_a}{a}\right)^2 \tag{1.8}$$

Assuming that any jitter in the measuring system is small compared to the sample rate, one approximation of the timing error is to take one half the time between samples. With this assumption, equation (1.8) becomes equation (1.9).

$$\left(\frac{\sigma_{\varepsilon_i}}{\varepsilon_i}\right)^2 = \sqrt{2}\left(\frac{cR}{af_s\varepsilon_i}\right)^2 + 4\left(\frac{\sigma_c}{c}\right)^2 + \left(\frac{\sigma_a}{a}\right)^2 \tag{1.9}$$

The implications of equation (1.9) on the array performance are:

1. The last two terms represent a constant contribution from fluctuations in the speed of sound (primarily arising from changes in temperature and humidity in the environment) and the positioning errors of the transducers. For short measurements indoors the second term is likely to be negligible.
2. The effect of the last two terms can be removed completely for the measurement of a given component by aligning the corresponding axis of the array perpendicular to the direction to the source such that $\varepsilon_i = 0$.
3. The first term represents a variable contribution from timing errors that varies from source to source. The error is greatest for sources further from the array.
4. The larger the array, or the higher the sample rate, the smaller the effect of the timing error on a given source.

6 SIMULATION

The effect of timing errors can easily be observed using a simple image source model of an array in an environment having one source and one reflection. The environment chosen was a closed room of dimensions 3x3x3 m having a perfectly reflecting floor and perfectly absorbing walls and ceiling. The source was located at position (-0.793,0.031,-0.381) m and the centre of the array at (0.443,0.031,-0.381) m with the origin of coordinates at the centre of the room. The array size was chosen to be the same as that of Sekiguchi and co-workers having *a*=0.17 m[10].

The image source model was used to generate theoretical times of flight to each microphone in the array from the direct sound and first reflection. These theoretical flight times were then individually perturbed by increasing amounts with the results fed back through equations (1.1), (1.2) and (1.3) to compute the positions of the sources. The difference between the source positions obtained on the reflection and the theoretical results are plotted for each component in Figure 1. It can be seen that the error increases linearly as a function of the timing error, except for the X and Y components which do not depend on all of the measured flight times. In these cases no change in the component as a result of the perturbation from that component is observed (as expected from equations (1.5), (1.6) and (1.7)).

It is obvious that the timing errors can lead to significant errors in locating a particular source. In the example shown in Figure 1, a sample rate of 22,050 Hz (corresponding to $\sigma_t = 22.7\mu s$) leads to worst case errors of ~0.1, ~0.15 and ~0.18 m for the X, Y and Z components respectively. This translates into errors of ~0.2° elevation and ~7° azimuth. To reduce the error in azimuth to within 1°

the error on the Y component must be reduced by an order of magnitude, requiring a sample rate of the order of 100kHz ($\sigma_t = 5\mu s$). This is well within the capabilities of currently available sound cards where sample rates of 96 kHz and higher are common. Alternatively, the array size could be increased (a factor of ~4.5 would be required to have the same effect) or a combination of both.

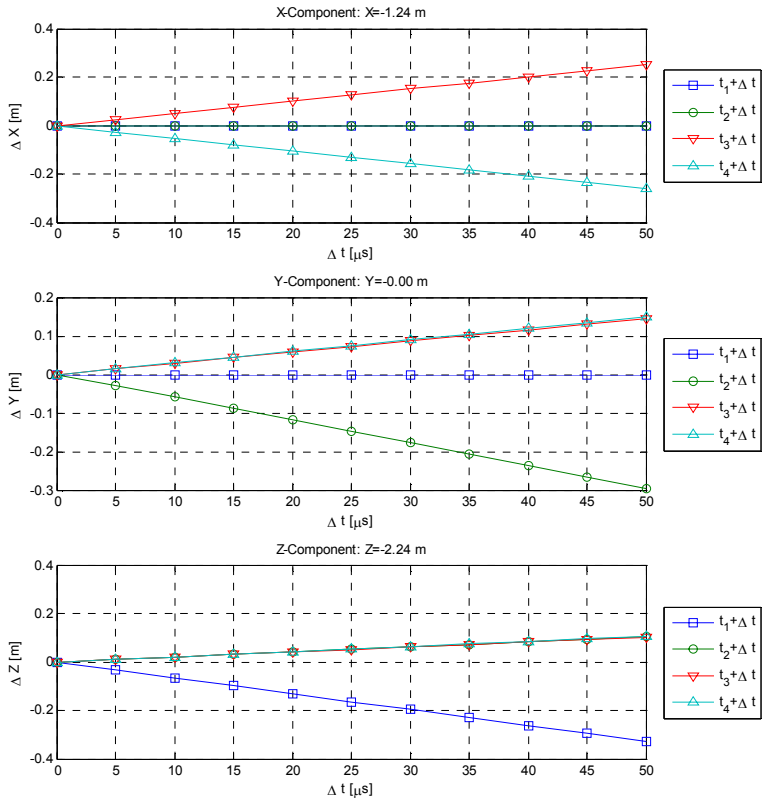

Figure 1: Simulated errors resulting from individually perturbing the flight times at each transducer by an amount Δt. The results are for the first reflection in the environment described in the text.

7 MEASUREMENTS

A tetrahedral array having transducer spacing of 17cm was constructed using a four ½" free field measurement microphones (type MCE212 with GRAS 26CA ICP pre-amps) and measured in the semi-anechoic chamber at the University of Salford. The positioning error of the acoustic centre of each transducer in the array on its mountings is estimated to be 0.01m.

A 32 channel NetdB-PRO-132 DAQ system and laptop was used to acquire the signals at a sample rate of 51.2 kHz (later down-sampled to 22.05 kHz in post-processing). The chamber was excited with a continuous white noise source played through a 6" loudspeaker in a sealed box as depicted in Figure 2. The output of the noise source was also recorded by the DAQ and the loudspeaker was found to produce output up to ~6 kHz. Impulse responses were obtained from estimates of the CPSD between a given microphone signal and the recorded white noise excitation, and the PSD of the white noise excitation itself. Impulse responses were obtained for the array located different distances from the source, on a line through the middle of the chamber.

The four microphones were calibrated for level at 1kHz using a B&K type 4230 portable calibrator. An approximate measurement of the phase mismatch between the four microphones was performed by picking one microphone as a common reference and comparing each microphone to that reference. The maximum discrepancy to the common reference was found to be ~6° for

frequencies up to 6 kHz. The phase and level calibrations were applied to the reduction of the data to impulse responses for each microphone.

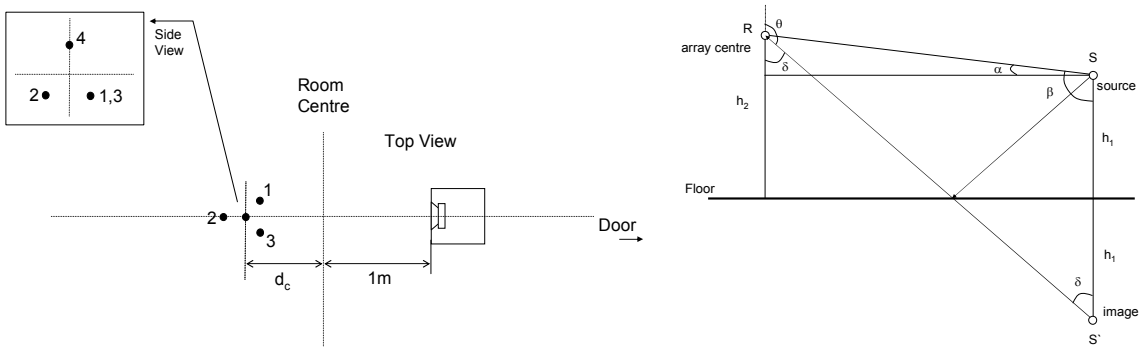

Figure 2: Left: top view of the tetrahedral array and source within the semi-anechoic chamber. Right: side view of the source, array and virtual image of the first reflection.

Flight times were obtained by correlating the impulse response for each microphone with an estimate of that of just the source taken from the direct sound. A peak finding algorithm[27] was then used to find the positions of the peak as shown in the examples in Figure 4. The resulting flight times were then fed into equations (1.1), (1.2) and (1.3) to obtain the positions of the virtual sources.

Equation (1.10) gives the expected spherical polar angle θ of the first reflection derived from the image source model in Figure 2.

$$\theta = \pi - \sin^{-1}\left[\frac{\sqrt{\left(SA^2 + \left(h_2 - h_1\right)^2\right)}}{S'R}\sin\left(\frac{\pi}{2} + \tan^{-1}\left(\frac{h_2 - h_1}{SA}\right)\right)\right] \tag{1.10}$$

This equation is plotted in Figure 3 against the source to receiver distance along with the TOF results and their error bars. It can be seen that the results are an excellent match within the error bars which are of the order of $\pm 10^{\circ}$.

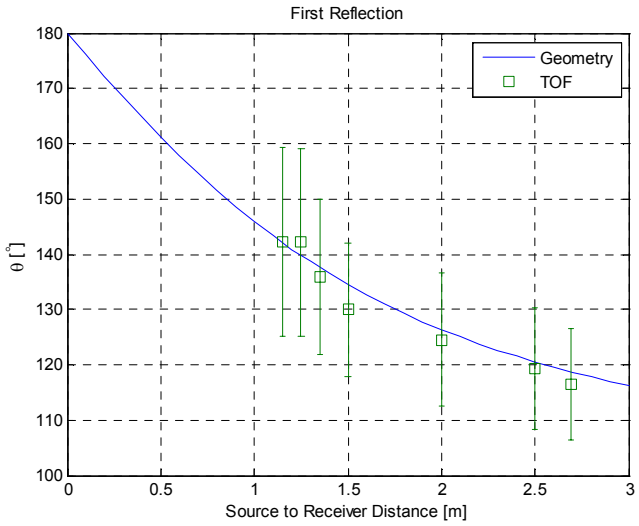

Figure 3: The spherical polar angle θ for the first reflection obtained from the TOF analysis compared to the expected angle from the image source model in Figure 2.

A further configuration where large sheets of plasterboard were placed either side of the source-receiver axis was also measured, a scenario where multiple simultaneous reflections are expected. An example result for one microphone is shown on the right in Figure 4 where even before the issue of multiple simultaneous reflections is encountered, the problem of accurate peak identification is plain to see. While the peak finding algorithm has identified some peaks, others have not be been found and without prior knowledge it is not clear what is a peak and what is an artefact. This is compounded when consistent results are required from all transducers in order to be able to apply the TOF equations.

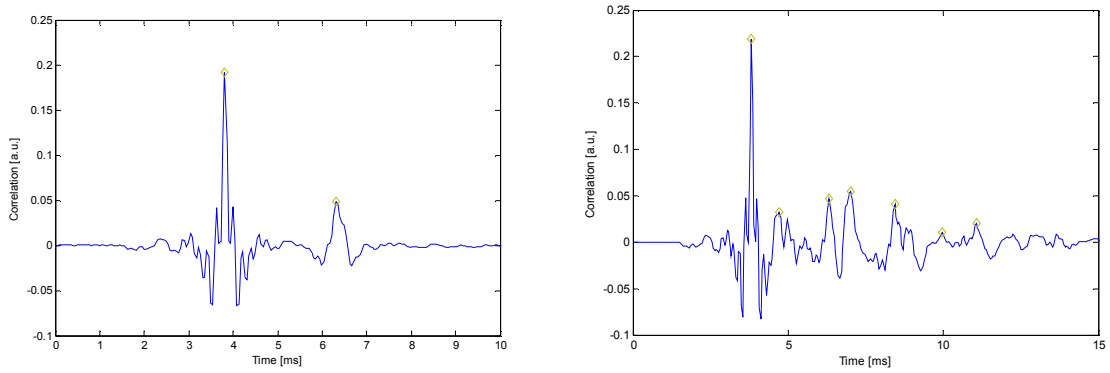

Figure 4: Cross-correlations between an estimate of the source impulse response and the full impulse response from one microphone in the array. Left: direct sound plus floor reflection. Right: direct sound plus floor reflection plus reflections from two parallel sheets of plasterboard either side of and parallel to the source-receiver axis. The results of automatic peak finding are shown as diamonds.

8 CONCLUSIONS

The four microphone TOF technique has been shown to be capable in theory of matching the best case MAA of the human auditory system if a suitable combination of sample rate and array size are chosen. When taking measurements, source and receiver positions of high symmetry must be avoided in order to prevent multiple simultaneous reflections. Care must also be taken with regard to identifying the flight times in the measured impulse responses.

9 ACKNOWLEDGEMENTS

We would like to thank Rhonda Wilson, Michael Capp and Alan Wood at Meridian Audio for suggesting this topic and for useful discussions. We would also like to thank David Waddington and Liam Kelly at the University of Salford for the loan of the equipment used in this study and for support in using it.

10 REFERENCES

1. Mills, A.W., On the Minimum Audible Angle, Journal of the Acoustical Society of America **30**(4) 237-246. (1958).
2. D'Antonio, P., F. Becker, and C. Bilello, Sound Intensity and Interaural Cross-Correlation Measurements using Time-Delay Spectrometry (Paper No. 2543), in AES 83rd Convention. 1987.
3. Laborie, A., R. Bruno, and S. Montoya, Reproducing multichannel sound on any speaker layout, in AES 118th Convention Paper Number 6375. 2005: Barcelona, Spain.
4. Humphrey, R.A., Automatic Loudspeaker Location Detection for use in Ambisonic Systems, MEng Thesis, University of York, (2006).

5. Essert, R., Measurement of Spatial Impulse Responses with a Soundfield Microphone, Paper 5pAA3, in Third joint meeting of the Acoustical Society of America and the Acoustical Society of Japan. 1996: Honolulu, Hawaii.
6. Fazenda, B. and J. Romero, 3-Dimensional Room Impulse Response Measurements in Critical Listening Spaces, Proceedings of the Institute of Acoustics **30**(6) 232-239. (2009).
7. Fahy, F., Sound Intensity: Elsevier Applied Science.(1989).
8. Guy, R.W. and A. Abdou, A measurement system and method to investigate the directional characteristics of sound fields in enclosures, Journal of Noise Control Engineering **41**(1) 8-18. (1994).
9. Yamasaki, Y. and T. Itow, Measurement of spatial information in sound fields by closely located four point microphone method, Journal of the Acoustical Society of Japan (E) **10**(2) 101-110. (1989).
10. Sekiguchi, K., S. Kimura, and T. Hanyu, Analysis of Sound Field on Spatial Information Using a Four-Channel Microphone System Based on Regular Tetrahedron Peak Point Method, Applied Acoustics **37** 305-323. (1992).
11. Robertson, A.E., Microphones, 2nd ed: Iliffe.(1963).
12. Broadhurst, A.D., An Acoustic Telescope for Architectural Acoustic Measurements, Acustica **46** 299-310. (1980).
13. Broadhurst, A.D., Sparse Volume Array for Architectural Acoustic Measurements, Acustica **50** 33-38. (1982).
14. Cox, H., R.M. Zeskind, and M.M. Owen, Robust Adaptive Beamforming, IEEE Transactions on Acoustics, Speech and Signal Processing **ASSP-35**(10) 1365-1376. (1987).
15. Brandstein, M. and D. Ward, Microphone arrays : signal processing techniques and applications. Digital signal processing: Springer.(2001).
16. Rafaely, B., Analysis and design of spherical microphone arrays, IEEE Transactions on Speech and Audio Processing **13**(1) 135-143. (2005).
17. Acoustic Camera. [cited 23rd September 2009]; Available from: http://www.acoustic-camera.com.
18. Brüel and Kjær, Pulse array-based noise source identification solutions. [cited; Available from: http://www.bksv.com/doc/bp2144.pdf.
19. Gover, B.N., Development and Use of a Spherical Microphone Array for Measurement of Spatial Properties of Reverberant Sound Fields, PhD Thesis, University of Waterloo, (2001).
20. Gover, B.N., J.G. Ryan, and M.R. Stinson, Microphone array measurement system for analysis of directional and spatial variations of sound fields, Journal of the Acoustical Society of America **112**(5) 1980-1991. (2002).
21. Gover, B.N., J.G. Ryan, and M.R. Stinson, Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array, Journal of the Acoustical Society of America **116**(4) 2138-2148. (2004).
22. Gover, B.N., Directional measurement of airborne sound transmission paths using a spherical microphone array, Journal of the Audio Engineering Society **53**(9) 787-795. (2005).
23. Rafaely, B., I. Balmages, and L. Eger, High-resolution plane-wave decomposition in an auditorium using a dual-radius scanning spherical microphone array, Journal of the Acoustical Society of America **122**(5) 2661-2668. (2007).
24. Weinstein, E., K. Steele, A. Agarwal, and J. Glass. LOUD: A 1020-Node Microphone Array and Acoustic Beamformer. in International Congress on Sound and Vibration (ICSV). 2007. Cairns, Australia.
25. Squires, G.L., Practical Physics, 4th ed: Cambridge University Press.(2001).
26. Rabinovich, S.G., Measurement Errors and Uncertainties - Theory and Practice, 2nd ed: American Institute of Physics.(2000).
27. Jun, G. Find peak value (fpeak.m). [cited October 2008]; From MATLAB Central - A MathWorks Web Resource]. Available from: http://www.mathworks.com/matlabcentral/fileexchange/4242.