



A quasi-binaural approach to head-movement-aware evaluation of spatial acoustics

Chungeun Kim, Russell Mason and Tim Brookes

Institute of Sound Recording, University of Surrey, Guildford, GU2 7XH United Kingdom

PACS: 43.55.MC, 43.60.JN, 43.66.PN, 43.66.QP

ABSTRACT

This research incorporates the nature of head movement made in listening activities, into the development of a quasi-binaural acoustical measurement technique for the evaluation of spatial impression. A listening test was conducted where head movements were tracked whilst the subjects rated the perceived source width, envelopment, source direction and timbre of a number of stimuli. It was found that the extent of head movements was larger when evaluating source width and envelopment than when evaluating source direction and timbre. It was also found that the locus of ear positions corresponding to these head movements formed a bounded sloped path, higher towards the rear and lower towards the front. This led to the concept of a signal capture device comprising a torso-mounted sphere with multiple microphones. A prototype was constructed and used to measure three binaural parameters related to perceived spatial impression - interaural time and level differences (ITD and ILD) and interaural cross-correlation coefficient (IACC). Comparison of the prototype measurements to those made with a rotating Head and Torso Simulator (HATS) showed that the prototype could be perceptually accurate for the prediction of source direction using ITD and ILD, and for the prediction of perceived spatial impression using IACC. Further investigation into parameter derivation and interpolation methods indicated that 21 pairs of discretely spaced microphones were sufficient to measure the three binaural parameters across the sloped range of ear positions identified in the listening test.

INTRODUCTION

Attempts to overcome the difficulties involved in acoustical evaluations of listening spaces or reproduction systems with subjective listening tests have resulted in the development of physical measures that can represent the human perception of acoustical attributes. Also, a number of signal capture devices to replace human listeners have been developed which enable the measurement of relevant parameters. This research extends these attempts, focusing on objective evaluation of spatial attributes of sound such as source width and envelopment, as well as source direction. Uniquely, this research takes into account the nature of head movement in the development of a signal capture system, towards more accurately representing human listening behaviour. The ultimate goal of this research is to design an objective evaluation system of spatial impression, including source direction, that incorporates the nature of head movement.

In the following sections, the background information is summarised from previous studies, as are related previous findings by the authors about the nature of head movement. Then the experiments which led to the development of the signal capture system for head-movement-aware binaural signal measurements are described, and details of the proposed system are given.

BACKGROUND

This section briefly introduces the key concepts related to this research – spatial impression and related acoustical

parameters, and findings from previous studies regarding head movement in listening activities. Then the authors' previous research investigating the nature of head movement in the specific case of spatial impression evaluation activity is summarised.

Concepts of spatial impression and related parameters

The concept of spatial impression was suggested in the 1960s as the feeling of the source broadening or the music "beginning to gain body and fullness", or a feeling related to "envelopment" [1, 2]. It was widely accepted by later researchers as an important acoustical quality of a listening space [3], and gradually categorised into two aspects, now generally known as source width and envelopment [4-6]. Investigations into binaurally measurable acoustical parameters that can effectively indicate the perceived spatial impression have found that the interaural cross-correlation coefficient (IACC) can be used for this purpose [7-11]. Fluctuations of Interaural time and level differences (ITD and ILD) were also found to be related to the perception of spatial impression, but to be less practical than IACC as spatial impression indicators [11].

Previous studies related to the nature of head movements in source localisation activities

Previous studies of head movement in listening activities showed in general that making head movement helps listeners locate the sound source, by resolving front-back

confusions which may occur in localisation [12-15]. The investigations into the pattern of head movement tended to focus on rotational movement (or orientation) and categorise it into three directions: in azimuth, elevation, and roll. The rotational movements in azimuth are known to have the most significant effect on the accuracy of localisation [16, 17]. However, little has yet been found about the nature of head movement or its effects in other subjective audio evaluation activities where source localisation is not of primary interest. This led the authors to a detailed investigation into the nature of head movement in more general listening activities evaluating a range of attributes including source location.

Nature of head movement found in the evaluation of spatial impression

An experiment was conducted by the authors in the form of subjective tests, where the listeners were allowed to move their heads freely whilst listening to various types of sound with the purpose of evaluating source location, source width, envelopment, and timbre [18]. The head movements were recorded with a head tracker which allowed collection of the head position and orientation coordinates.

Firstly, statistical analysis of the recorded data in conjunction with the various conditions of the test settings revealed that, across all listeners, the range of movement was relatively small when judging timbre, larger when localising sources, and greatest when evaluating source width and envelopment. This indicated that head movement is meaningful in the listeners' judgment of spatial impression, as well as of source location, and therefore should be taken into consideration in the implementation of an objective evaluation system in order to be representative of human listeners.

Secondly, modal analysis of the collected head orientation data showed that the corresponding ear positions during the tests formed a locus that is confined and sloped, higher towards the rear and lower towards the front [19]. This suggested that a head-movement-aware signal capture system only needs to cover a limited range of positions, and that this range is not along the horizontal plane as previously assumed. Figure 1 describes this finding in the form of a spherical histogram of the estimated right ear position calculated from the recorded head orientation data. The plot shows where the right ear was positioned the most during the listening tests, viewed from the right hand side of the head.

INTRODUCTION AND INITIAL VALIDATION OF PROTOTYPE SIGNAL CAPTURE DEVICE

In this section, an experiment is described which contributed to the development of a prototype signal capture device for head-movement-aware binaural signal measurements considering the ear locus described above.

Multiple-microphone sphere as a prototype signal capture device

An intuitive way to implement head movement in a binaural signal capture system is to rotate a dummy head that is representative of human head and torso in terms of shape and dimensions, sometimes called head and torso simulator (HATS). However, rotating the capturing device during the measurements is impractical for time-varying signals, and the need for multiple measurements taken at a range of positions increases the measurement time. These problems can be solved by the use of a sphere with multiple microphones at the corresponding ear positions. This method enables simultaneous capture of the binaural signals at multiple head orientations. However, there is a possibility that the

measurement accuracy may be reduced compared to the head model, due to the simplification of the head shape. This experiment was designed to examine this accuracy reduction in the use of the sphere as the signal capture model and to test its validity.

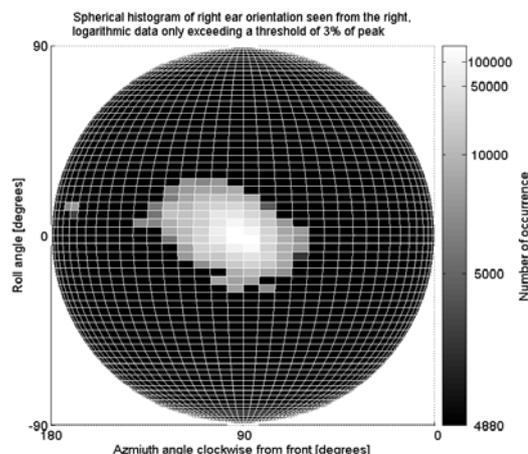


Figure 1. Spherical histogram of right ear position corresponding to the head orientation data collected during various acoustical evaluations. Each grid corresponds to 5° angular distance horizontally, and 2.5° vertically. View from the right hand side of the head (the nose is effectively half way up the right hand side of the sphere). Logarithmic data were used. Only data above a threshold of 3% of the peak are displayed.

Initial investigation of binaural parameter measurement performance

The sphere model was a plastic sphere of 17.2cm diameter, the dimensions of which were determined by averaging the sizes of the various head models in previous studies [20, 21]. Two omni-directional microphones (Countryman B3) were placed on the surface of the sphere through small holes, 180 degrees apart from each other. Three binaural parameters considered to be related to perceived spatial impression, as discussed earlier – ITD, ILD and IACC – were derived from a number of simulated binaural responses. The binaural responses were created by convolving white Gaussian noise with one or more binaural impulse responses, which were measured in a quasi-anechoic manner [22], with the head model rotating from 0 to 360 degrees azimuth at 2.5-degree intervals, on a rotating table (Outline ET2-ST2). Two configurations of the sound sources were simulated such that a wide range of ITD, ILD and IACC values could be derived. Firstly, a point source with varying lateral angle (from 0° to 360° azimuth) with respect to head orientation was simulated. Secondly, a varying number of decorrelated noise sources around the frontal direction were simulated with the intention of generating different levels of source width or envelopment – from a single source at 0° azimuth to 18 decorrelated sources spaced 20° apart covering the whole azimuth [19]. Figure 2 describes some of the source specifications introduced for this experiment as an example. The measurements were repeated with a HATS (a Cortex Manikin MK2 whose dimensions conform to international standard IEC TR 60959 [23]) as the reference, and the measurements made using the sphere were compared to this. Also, some variations were made to the sphere model, such as adding a torso taken from a KEMAR dummy head [21], a pair of pinnae also from the KEMAR dummy head, and a nose. This was to investigate the effects of these body parts on the differences of the measured parameters.

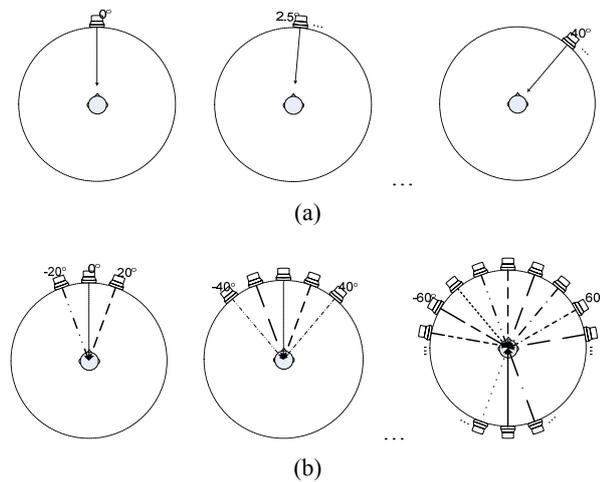


Figure 2. Examples of devised sound sources for the experiment: (a) point source at various directions, and (b) spanned decorrelated sources of various angular widths.

The ITD, ILD and IACC were calculated in a number of different frequency bands after the captured signals had been passed through a gammatone filterbank [24]. This was to take account of the dependency of the relationship between IACC and perceived width on frequency [25], and to allow for a more detailed observation of the results. The IACC was calculated from the following cross-correlation coefficient function using the binaural signals:

$$C(\tau) = \frac{\int_{t_1}^{t_2} f_l(t) f_r(t + \tau) dt}{\sqrt{\int_{t_1}^{t_2} f_l^2(t) dt \int_{t_1}^{t_2} f_r^2(t) dt}}, \quad (1)$$

where $f_l(t)$ and $f_r(t)$ are the signals at the left and the right ears, t_1 and t_2 define the period of measurement, and τ is an offset between $f_l(t)$ and $f_r(t)$. IACC is taken as the maximum value of $C(\tau)$ over the range of $|\tau| \leq 1$ ms. This range of τ is specified such that the maximum possible ITD (when the two ears are in line with the path of sound propagation) can be included. From this procedure the ITD can also be found: it is the value of τ at which $C(\tau)$ is a maximum [26]. The ILD was calculated by subtracting the mean sound pressure level (SPL) of the signal at the left ear from that at the right ear.

The three parameters measured using the HATS and each variation of the sphere were compared in terms of the perceptibility of their differences. As the criteria, the just noticeable differences (JNDs) of ITD, ILD and IACC were derived through a review of a number of previous related studies [27]. Firstly, the ITD JND was found to be affected by the reference ITD (the initial ITD value at which the JND measurements are made) and by the IACC value. Considering these relationships, the ITD JND for any given source direction and frequency was specified as a function of the reference ITD and IACC values obtained using the HATS in the same condition. Secondly, a fixed ILD JND value was determined, by averaging those found from previous works which varied within a small range for similar experimental conditions (noise source, 0dB reference). Thirdly, the IACC JND was found to decrease as the reference value increased, approximately following Weber's law which was suggested in [28], although the results seem to have been affected by various factors such as the source type, sound level and duration. From the collected previous findings, a relationship between the IACC JND and the reference was derived by adopting and slightly modifying Weber's law.

Results and discussion

Comparisons of ITD and ILD between the HATS and the various versions of the sphere generally showed that the differences are not perceptually negligible at large ranges of source directions and frequencies, regardless of the variations of the sphere. By way of an example, Figure 3 shows the ITD and ILD differences between the HATS and the sphere mounted on the torso (which was found slightly more accurate than other variations of the sphere). In the plot, the radius corresponds to a frequency band (the centre frequency of the gammatone filter). The centre of the disc corresponds to the lowest frequency band, centred at 100Hz, and the outmost points along the boundary of the disc correspond to the highest, centred at approximately 10.1kHz. The angles correspond to the source direction with the head assumed to be at the centre of the disc. The bright area inside the disc indicates conditions where the difference exceeds the JND and would therefore be perceivable, and the dark area indicates conditions where the difference is within the JND and would be perceptually negligible.

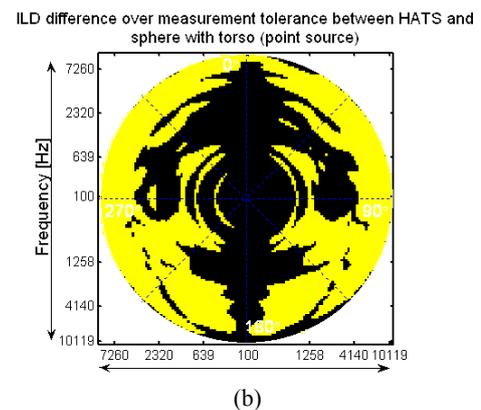
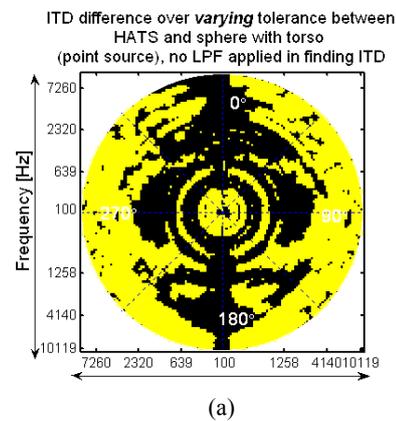


Figure 3. Plots of ITD and ILD differences between the HATS and the sphere on a torso, compared to their JNDs as the measurement tolerances. The bright area inside the disc indicates that the difference exceeds the JND for the corresponding source direction and frequency band, and the dark area indicates that the difference is within the JND.

On the other hand, the IACC differences were found to be perceptually negligible for the majority of source directions, widths and frequencies. Figure 4 shows the IACC results, with the sphere mounted on the torso. In the case of point sources, it can be said that the HATS can be replaced with the sphere with torso for IACC measurements at frequencies below approximately 4kHz, with perceptually negligible errors. In the case of spanned sources, the IACC difference was found to be larger than the measurement tolerance at certain frequency bands and for certain angular widths of the

sources, except when fewer point sources were simulated around the front. However, the difference is perceptually negligible at lower frequencies and is always negligible below approximately 1kHz.

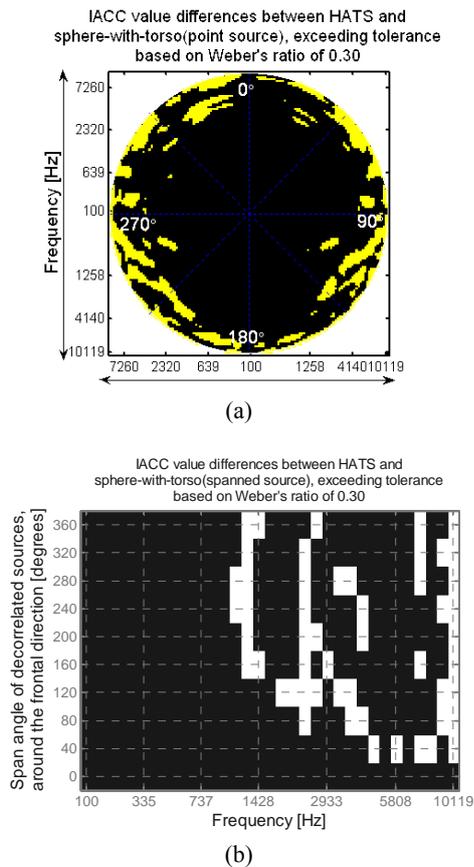


Figure 4. Plots of IACC differences between the HATS and the sphere with torso, compared to the JNDs: (a) for the simulated point source at various directions, and (b) for the simulated spanned sources with various angular widths. The bright area indicates that the difference exceeds the measurement tolerance, and the dark area indicates that the difference is within the tolerance.

These findings indicate that some versions of IACC, specified only for low frequencies, can be measured with the sphere model with the torso instead of the HATS. It has previously been found that IACC values calculated over low frequency ranges are closely related to the perceived source width (below the 2kHz octave band) or envelopment (below the 4kHz octave band) [4, 29, 30]. This means that, when the angular source width is below approximately 80° , the sphere with torso will be capable of predicting the source width or envelopment using IACC in the manner described in those previous works.

Evaluation of source direction prediction performance

The results of the ITD and ILD comparisons for the point source have shown that the sphere model is not an accurate substitute for the HATS for these measurements, for the configurations tested. However, the measurements of ITD and ILD are intended for use in predicting the perceived sound source location, and it is still possible that the sphere model is useable for this purpose. A set of tests was designed in order to clarify whether the sphere model could predict the source direction accurately, despite the differences in the measured ITD and ILD values between the sphere and the HATS.

Firstly, from the measured ITD and ILD data using the HATS and the sphere with torso, ITD and ILD databases were created in the form of look-up tables [31] from which a source direction corresponding to a given set of ITD and ILD could be calculated. Secondly, each of these tables was tested using different sets of ITD and ILD values for a range of intended source directions. These test sets were generated by means of averaging two adjacent values for a single intended source direction. For example, the ITD values for the intended source direction of 2.5° azimuth (for all the introduced frequency bands from 100 to 10119Hz) were created by averaging those measured at 2.5° and at 5° . This way, a test set was created for the intended source directions of -177.5° to 172.5° in intervals of 12.5° , per each of the two parameters (ITD and ILD), and per each of the models (HATS and sphere with torso). Then for each of the test ITD or ILD data, the look-up tables were used to convert the measured values to a probability score for each candidate source direction, indicating how close the look-up values were to the given test values at each angle. From this a prediction was made of the source direction, as the angle for which the sum of probabilities for ITD and ILD, averaged over all the measured frequency bands, was the largest. The predicted source angles were finally compared to the intended directions.

Results and discussion

It was found from the comparison that when correct comparisons are made, i.e. the HATS look-up table is used for the HATS test data and the sphere look-up table is used for the sphere test data, the prediction is almost always accurate. Figure 5 shows an example for the case of predicting source direction using the lookup table created using the data collected with the sphere with torso.

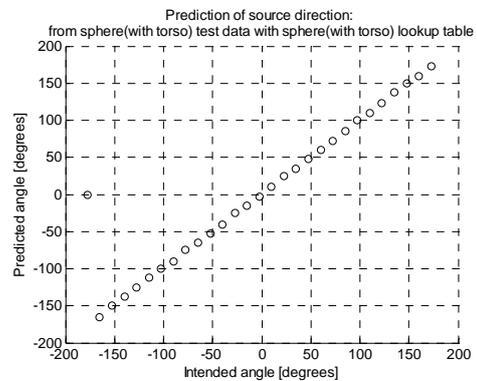


Figure 5. Result of horizontal source direction prediction from test ITD and ILD sets using a lookup table created with actual sphere-with-torso measurements. Only one suspected case of front-back confusion is found.

This indicates that although the ITD and ILD measured with the sphere with torso were not perceptually close enough to those measured with the HATS, the sphere model can still predict horizontal source direction reliably. As a test of this approach, predictions were also made using the incorrect look-up tables (i.e. using the HATS look-up table to predict the source direction of the sphere measurements, and vice versa), and this showed a much greater inaccuracy in the results, including significant numbers of errors of translations around the cones of confusion.

As the IACC differences between the two models were mostly perceptually negligible as seen from Figure 4, and the source location can be accurately predicted from the ITD and ILD data using a look-up table approach, it seems that the sphere model (with torso) is valid for the potential head-

movement-aware signal capture system for the prediction of overall spatial impression and source direction.

Summary

A sphere-based signal capture device which enables head-movement-aware measurements of ITD, ILD and IACC in a practical way by means of multiple microphones was introduced and tested in this section. The effects of the head shape simplification were initially investigated in the experiment, where the three parameters obtained from a variety of simulated sound sources were compared to those measured with a HATS. It was found that whilst the ITD and ILD differences between the HATS and the sphere measurements are not perceptually negligible in general, the IACC differences are perceptually negligible, making the sphere valid as a potential objective prediction device of source width and envelopment. Secondly, the source direction prediction accuracy was tested using the ITD and ILD data obtained with the sphere with torso. The results showed that although the perceptual accuracy of the ITD and ILD measurements made using the sphere model may be perceptually distinguishable from those of the HATS, the horizontal source direction can still be predicted accurately with the sphere with torso.

INVESTIGATION INTO PARAMETER DERIVATION TECHNIQUES FOR PERCEPTUALLY ACCURATE PROTOTYPE DESIGN

In the previous section, the use of the sphere for binaural parameter measurements was validated, in comparison to the HATS, for the prediction of source direction, source width and envelopment. In order to develop the sphere model to allow quasi-binaural head-movement-aware measurements using multiple microphones, the pattern and distribution of the microphones need to be determined. This section addresses this issue using investigations into various ITD, ILD and IACC measurement techniques that would enable perceptually accurate interpolation of these parameters from discrete microphone measurements.

Experimental design

The experiment was designed such that the ITD, ILD and IACC at a single pair of potential ear spots can each be approximated from the measurements made at two adjacent pairs of microphones spaced by various angular distances, using a range of derivation and interpolation techniques. Then the perceptual accuracy of each approximation was tested by comparing the approximated values to those measured at the target ear positions. The quasi-anechoic binaural impulse responses obtained in the previous experiment were used to synthesise a range of binaural signals from which the three parameters were derived. The variations introduced in terms of the microphone spacing, the source specification, and the parameter derivation processes are described below.

Microphone spacing schemes

The tested microphone spacing schemes were based on a number of methods of estimating the ear spacing that would create errors in the ITD, ILD or IACC that would be below the JND thresholds used in the previous experiment. Firstly, an angular distance between microphones was calculated, based on the angular spatial sampling theorem [32], which would enable the reconstruction of the sound field. The maximum allowable angular distance between microphones for minimal aliasing error was calculated to be 11.3° for the

sphere head model (17.2cm diameter) for frequencies up to 10kHz. Secondly, attempts were made to increase this allowable angular distance whilst maintaining the resultant errors within the JND thresholds. The angular differences corresponding to the JNDs, or the minimum audible angle were added to the sampling theorem-based angular distance, as perceptually allowable margins. It was expected that if the microphone distance was increased by an amount that is smaller than these marginal angles, the resultant increase in estimation errors would not be perceivable. For instance, the ITD JND was converted to an angle by inverting the equations which were introduced and used by Algazi et al. [20] originally to approximate the ITD with the lateral source angle. Based on this, the ITD JND value relates to an angle of 1.4° and 2.2° for frequencies below and above 600Hz. For the ILD, the relationship between the ILD and the lateral source angle was found to vary with frequency and the reference source direction itself [20]. The smallest angle change about the centre corresponding to the ILD JND was found to be around 5.8° . In the case of the IACC, a short simulated experiment by the authors showed that the angular difference between the microphones needs to be within 2.5° for the corresponding IACC difference to be perceptually negligible for frequencies up to approximately 2kHz [27]. For higher frequency ranges, the allowable microphone gap was smaller.

These findings resulted in the approximate microphone intervals that may result in perceptually negligible differences ranging from 11.3° (sampling theorem-based spacing) to 17.1° (sampling theorem-based spacing plus the marginal angle corresponding to the ILD JND). Since the previous measurements of the quasi-anechoic binaural impulse responses were made in 2.5° intervals of orientation, multiples of 2.5° were used in this experiment as the tested gaps between the positions of microphone pairs, considering the range found above. Consequently, the following microphone intervals, around a single target position, were introduced:

- $\pm 5^\circ$ from the target position (similar to the spacing based on the sampling theorem)
- $+7.5^\circ$ and -5° from the target position (similar to the spacing based on the sampling theorem and additional angle corresponding to the ITD JND)
- $\pm 7.5^\circ$ from the target position (similar to the spacing based on the sampling theorem and additional angle corresponding to the ILD JND)
- $+12.5^\circ$ and -10° from the target position (similar to the spacing based on the sampling theorem and additional angle corresponding to the doubled ILD JND – i.e. applied on both sides).

Source simulation

Three source types were simulated for this experiment, using the binaural impulse responses obtained in the previous experiment – a single point source, two uncorrelated point sources positioned diametrically opposite in the horizontal plane, and spanned multiple uncorrelated sources, from 0° to all around the head, distributed at 20° intervals on each side. For the first two source types, the measurements were made for the azimuth head orientation angles from 0° to 357.5° at intervals of 2.5° . The third type is as seen in Figure 2(b), with the head fixed facing 0° azimuth.

Variations in parameter derivation processes

In order to find an optimal parameter measurement process, various calculation or interpolation techniques were tested at three different stages of parameter derivation. Firstly, two interpolation methods – linear interpolation as introduced in

[20], and interpolation using the sinc function following the sampling theorem [32] – were tested to estimate the ITD, ILD and IACC values.

Secondly, the parameters at the target position were estimated in two different ways, by altering the stage at which the interpolation was applied. One approach calculated the parameters directly from the binaural audio signals captured at the adjacent pairs and then interpolated these parameters to give the estimates of measurements at the target position. The other approach interpolated the binaural audio signals at the adjacent microphone pairs and the three parameters were calculated from this.

Thirdly, in the ITD calculation process, the effects of half-wave rectifying and low pass filtering of the binaural signals (as suggested in [25]) were examined, to take into account the breakdown of phase locking in the human auditory system at high frequencies. This procedure was found to be perceptually relevant, since at high frequencies the perceived source width was found to be dependent upon the IACC of the signal envelope instead of that of the signal itself. However, it was also found that this procedure sometimes alters the peak of the cross-correlation at which the ITD is found, as described above, resulting in incorrect ITD values at high frequencies [27]. Therefore, the results of ITD calculation with and without this low filtering of the binaural signals were compared.

Results and discussion

The three parameters ITD, ILD and IACC measured from signals captured at the target ear position were compared to those estimated from the signals captured at the adjacent microphone pairs separated by various angles as specified above. The comparison, for each of the variations in the calculation process, was made by firstly calculating the differences of the three parameters, and then by comparing them to their JNDs, calculated as described above.

Firstly, it was found from the overall results that as the microphone interval becomes wider, the approximation accuracy reduces, in that the frequency above which perceivable errors are found decreases. An example of these results can be seen in Figure 6. The plots show at which frequencies and for which head orientation angles the interpolation was perceptually negligible (dark) or not (bright). As in Figure 3, the frequency increases from the centre of the disc (100Hz) to the outmost points along the boundary of the disc (approximately 10.1kHz). The angles now correspond to the head orientation in azimuth for which the interpolation accuracy was observed. With the other settings fixed, wider spacing of microphone pairs made the ITD approximation less accurate at higher frequencies.

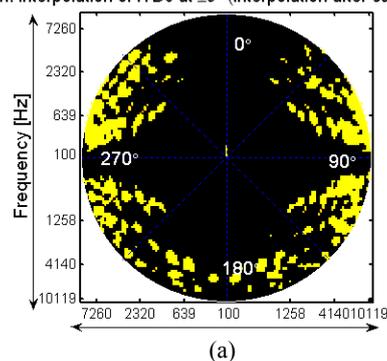
Secondly, using the sinc function in the interpolation was not found to be beneficial over the linear interpolation, as seen from Figure 7 in comparison to Figure 6(a) (with the other conditions fixed). This seems due to the fact only two adjacent measurements were used for the interpolation to maintain practicality and reduce the number of microphones required in the signal capture system.

Thirdly, it was found to be more accurate in general to interpolate the ITD, ILD and IACC after their calculation from the binaural signals, which can be seen by comparing Figure 8 to Figure 6(b). Finally, the low-pass filtering of captured binaural signals in the calculation of ITD from the cross-correlation was found to slightly improve the approximation performance in the case of a single point source, and thus to be relevant for the final measurement

system. Comparison of Figure 9 to Figure 6(a) shows an example.

In the optimal conditions found so far (linear interpolation after parameter calculation, and low-pass filtering of binaural signal in the calculation of ITD), the trade-off between the microphone spacing and the approximation accuracy was such that $\pm 5^\circ$ spacing gave reliable overall results up to about 2kHz, whereas $\pm 7.5^\circ$ spacing gave reliable results up to about 1.2kHz (some ITD comparison results for the opposite uncorrelated sources were not considered in this reasoning, due to these soundfields being perceptually erratic in general and to being less likely to occur in natural listening). This overall performance is in fact worse than what was expected from the discussion of the sampling theorem (negligible errors up to 10kHz), although the $\pm 5^\circ$ spacing is based on the sampling theorem itself. This seems primarily to be due to the fact that results were interpolated linearly between parameters calculated for signals captured at a pair of positions, whilst the sampling theorem requires sinc interpolation of binaural signals or parameters at a number of adjacent measurement spots. However, the development of a practical measurement system with a small number of microphones required only pairs of positions to be used in the interpolation.

ITD difference over tolerance, calculated for single source at 0° , between measurements at target ear positions and lin. interpolation of ITDs at $\pm 5^\circ$ (interpolation after calculation)



ITD difference over tolerance, calculated for single source at 0° , between measurements at target ear positions and lin. interpolation of ITDs at $\pm 12.5^\circ$ and -10° (interpolation after calc.)

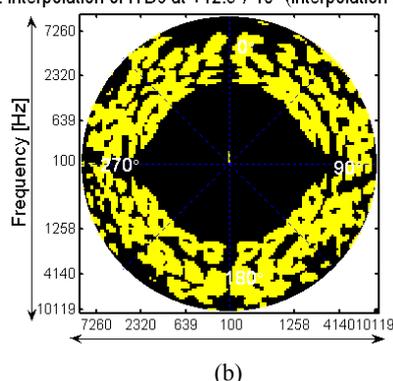


Figure 6. ITD difference for single point source, compared to the ITD JND, between the measurement at the target ear position with varying head orientation and the one *linearly interpolated* from those calculated at the microphone pairs (a) $\pm 5^\circ$ apart; and (b) $\pm 12.5^\circ$ and -10° apart (*low-pass filtering applied* to binaural signals in finding ITD). The dark areas indicate that the differences are perceptually negligible, and the bright areas indicate that they are not.

With $\pm 5^\circ$ as the allowable microphone spacing interval, and with 2kHz as the approximate upper frequency boundary for

perceptually accurate derivation of ITD, ILD and IACC through direct calculation or linear interpolation, it is now possible to design a prototype quasi-binaural signal capture device for head-movement-aware prediction of spatial impression and source direction. Figure 10 shows the suggested microphone layout, overlaid on the spherical histogram of the right ear positions found previously as in Figure 1.

ITD difference over tolerance, calculated for single source at 0°, between measurements at target ear positions and sinc interpolation of ITDs at ±5° (interpolation after calculation)

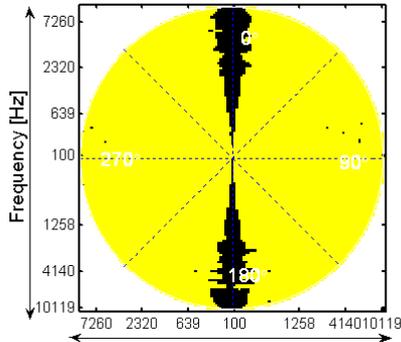


Figure 7. ITD difference for single point source, compared to the ITD JND, between the measurement at the target ear position and the ITD interpolated using the sinc function from those calculated at the adjacent microphone pairs ±5° apart (low-pass filtering applied to binaural signals in finding ITD).

ITD difference over tolerance, calculated for single source at 0°, between measurements at target ear positions and from interpolated signals at +12.5°/-10° (interpolation before calc.)

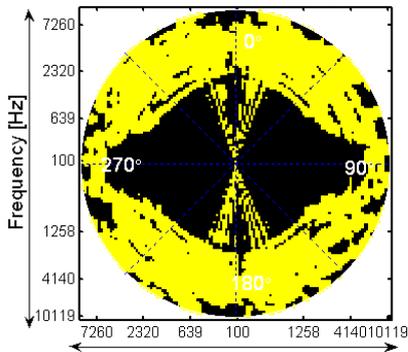


Figure 8. ITD difference for single point source, compared to the ITD JND, between the measurement at the target ear position and the ITD calculated from the linearly interpolated binaural signals at the adjacent microphone pairs +12.5° and -10° apart (low-pass filtering applied to binaural signals in finding ITD).

ITD difference over tolerance, calculated for single source at 0°, between measurements at target ear positions and lin. interpolation of ITDs at ±5° (interpolation after calculation)

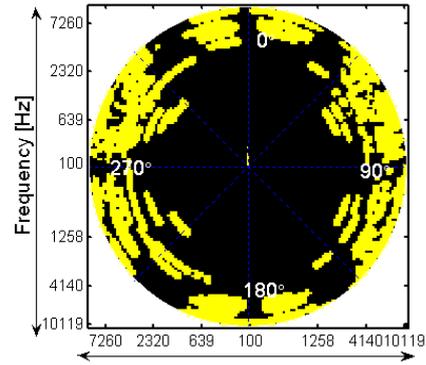


Figure 9. ITD difference for single point source, compared to the ITD JND, between the measurement at the target ear position with varying head orientation and the one linearly interpolated from those calculated at the microphone pairs ±5° apart (low-pass filtering not applied to binaural signals in finding ITD).

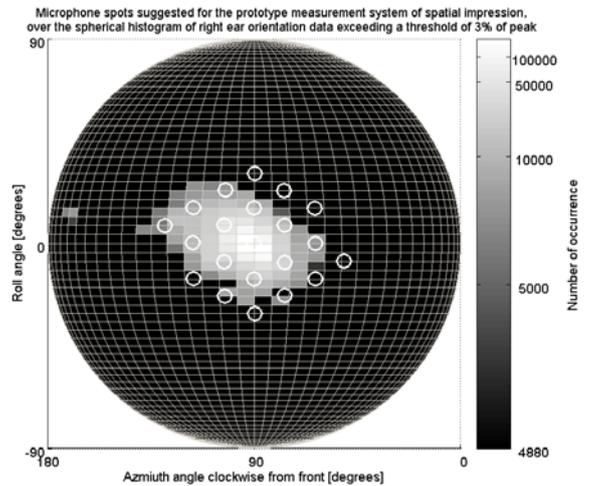


Figure 10. Suggested layout of microphones for the prototype measurement model of spatial impression incorporating head movements, drawn over the spherical histogram of Figure 1. The white circular marks indicate the suggested microphone spots. The small outstanding spot to the rear was excluded due to practicality and the small number of occurrences.

The white circles on the plot indicate the suggested microphone positions, 21 per side of head, drawn with 1-dimensional symmetry about the ear position for 0° head orientation (0° roll and ±90° azimuth on the sphere) taken into account. This is because when a right ear spot is selected, the corresponding left ear spot also has to be selected symmetrically on the opposite side. It is assumed that the ears are located at 0° roll and ±90° azimuth on the sphere for 0° orientation. The prototype signal capture system actually manufactured based on the design above is shown in Figure 11.



Figure 11. Developed prototype quasi-binaural signal capture system, based on the design suggested on Figure 10.

By intentionally considering the potential need for perceptually valid ITD, ILD or IACC measurements at any ear position over the ear coverage area traced from actual listeners' evaluation activities, this prototype provides the user with the flexibility of the amount of data to use depending on the needs and applications. For example, in some circumstances it may not be necessary to use the parameters measured with all the 21 pairs of microphones, or additional ITD, ILD or IACC data may be needed at an intermediate ear spot between any pair of microphones.

Summary

This experiment investigated a range of techniques to determine the number and pattern of microphones on the sphere for perceptually accurate ITD, ILD and IACC measurement incorporating the nature of head movement. The accuracy of the estimation of each parameter at a target position from measurements taken at adjacent microphone pairs was evaluated under a number of conditions including a range of microphone spacing intervals, simulated sound sources, and variations in the parameter calculation processes. It was found that in general linear interpolation of the ITD, ILD and IACC calculated from the captured binaural signals, along with the application of low pass filtering to the binaural signals when deriving ITD, gave the most accurate approximation results. Under these conditions and with relatively realistic source specifications, the microphone interval of $\pm 5^\circ$ was found to enable reliable ITD, ILD and IACC approximation for frequencies up to approximately 2kHz. This led to the design of a prototype sphere with 21 multiple microphones on each side, 10° apart from each other and distributed over the confined and sloped ear locus.

SUMMARY AND CONCLUSION

This research was aimed at incorporating the nature of head movement into the development of a signal capture system for objective evaluation of spatial impression. The findings in the described experiments lead to the following overall conclusions.

- Head movements occur in the evaluation of spatial impression, and thus should be incorporated in the development of an objective measurement and

evaluation system that mimics human listeners. The corresponding ear positions are confined around the initial ear position and tilted higher towards the rear and lower towards the front.

- A signal capture device has been developed that is a sphere attached to a torso, whose dimensions are based upon the average sizes of human head and torso, that contains multiple microphones which correspond to the varying ear positions caused by head movement. It can measure ITD and ILD to predict the source direction, and IACC to predict the perceived spatial impression with perceptually negligible errors.
- The number and layout of the microphones to be placed over the ear locus on the sphere depend on the frequency range of interest. A prototype was suggested using 21 microphones on each side, with a spacing of 10° , for perceptually reliable derivation of ITD, ILD and IACC up to approximately 2kHz.

FUTURE WORK AND APPLICATIONS

The prototype introduced as the result of this research will enable simultaneous capturing of binaural signals at various ear positions, and derivation of corresponding binaural parameters. In order for the developed system to successfully work as an evaluation tool of spatial acoustics, a clear connection needs to be established between the head-position-related variations of the binaural parameters and their perception as the next step, which will lead to an effective interpretation method of the measured parameters. Suitable user-friendly descriptive methods to present the expectedly complicated results of the interpretation of the binaural cues will be determined. Attempts will also be made towards optimisation of the evaluation system, by examining the possibilities to reducing the number of microphones necessary to indicate the predicted spatial attributes without sacrificing the perceptual accuracy. The complete system will provide the users with perceptually accurate estimation of perceived source direction, and of the source or environment width, in a variety of output formats from raw binaural parameters across a specified frequency range on an intuitive graphic display. It can be used for acoustic evaluations of a given enclosed space such as concert halls, listening rooms and cars in terms of the spatial impression delivered to the listeners. The quality of a reproduction system or process can also be evaluated with this system. For instance, the change in perceived spatial impression caused by a given processing method applied to a source material, or by employing a different type of reproduction system (in terms of loudspeaker layout or distribution of channels) can be predicted.

ACKNOWLEDGEMENT

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), UK, Grant No. EP/D049253.

REFERENCES

- 1 A.H. Marshall, "A Note of the Importance of Room Cross-Section in Concert Halls," *Journal of Sound and Vibration* vol. 5, no. 1, pp. 100-112 (1967).
- 2 M. Barron, "The Subjective Effects of First Reflections in Concert Halls – the Need for Lateral Reflections," *Journal of Sound and Vibration* vol. 15, no. 4, pp. 475-494 (1971).
- 3 L.L. Beranek, *Concert and Opera Halls – How They Sound*, Acoustical Society of America (1996).

- 4 J.S. Bradley and G.A. Soulodre, "Objective Measures of Listener Envelopment," *The Journal of the Acoustical Society of America* vol. 98, no. 5, pp. 2590-2597 (1995).
- 5 D. Griesinger, "The Psychoacoustics of Apparent Source Width, Spaciousness and Envelopment in Performance Spaces," *Acustica* vol. 93, no. 4, pp. 721-731 (1997).
- 6 T. Hidaka, T. Okano, and L. Beranek, "Interaural Cross Correlation (IACC) as a Measure of Spaciousness and Envelopment in Concert Halls," *The Journal of the Acoustical Society of America* vol. 92, no. 4, pp. 2469 (1992).
- 7 J. Blauert and W. Lindemann, "Spatial Mapping of Intracranial Auditory Events for Various Degrees of Interaural Coherence," *The Journal of the Acoustical Society of America* vol. 79, no. 3, pp. 806-813 (1986).
- 8 W.d.V. Keet, "The Influence of Early Lateral Reflections on the Spatial Impression," presented at The 6th International Congress on Acoustics, Tokyo (1968).
- 9 R.I. Chernyak and N.A. Dubrovsky, "Pattern of the Noise Images and the Binaural Summation of Loudness for the Different Interaural Correlation of Noise," presented at The 6th International Congress on Acoustics, Tokyo (1968).
- 10 K. Kurozumi and K. Ohgushi, "The Relationship between the Cross-Correlation Coefficient of Two-Channel Acoustic Signals and Sound Image Quality," *The Journal of the Acoustical Society of America* vol. 74, no. 6, pp. 1726-1733 (1983).
- 11 R. Mason, *Elicitation and Measurement of Auditory Spatial Attributes in Reproduced Sound*, PhD thesis, Institute of Sound Recording, University of Surrey (2002).
- 12 F.L. Wightman and D.J. Kistler, "Resolution of Front-Back Ambiguity in Spatial Hearing by Listener and Source Movement," *The Journal of the Acoustical Society of America* vol. 105, no. 5, pp. 2841-2853 (1999).
- 13 P.T. Young, "The Role of Head Movements in Auditory Localization," *Journal of Experimental Psychology* vol. 14, pp. 95-124 (1931).
- 14 D.R. Begault, *3-D Sound for Virtual Reality and Multimedia*, Academic Press (1994).
- 15 W.R. Thurlow and P.S. Runge, "Effect of Induced Head Movements on Localization of Direction of Sounds," *The Journal of the Acoustical Society of America* vol. 42, pp. 480-488 (1967).
- 16 S. Perrett and W. Noble, "The Contribution of Head Motion Cues to Localization of Low-Pass Noise," *Perception and Psychophysics* vol. 59, pp. 1018-1026 (1997).
- 17 S. Perrett and W. Noble, "The Effect of Head Rotations on Vertical Plane Sound Localization," *The Journal of the Acoustical Society of America* vol. 102, no. 4, pp. 2325-2332 (1997).
- 18 C. Kim, R. Mason, and T. Brookes, "An Investigation into Head Movements Made When Evaluating Various Attributes of Sound," presented at The 122nd AES Convention, Vienna, Austria (2007).
- 19 C. Kim, R. Mason, and T. Brookes, "Improvements to a Spherical Binaural Capture Model for Objective Measurement of Spatial Impression with Consideration of Head Movements," presented at The 125th AES Convention, San Francisco, USA (2008).
- 20 V.R. Algazi, R.O. Duda, and D.M. Thompson, "Motion-Tracked Binaural Sound," *Journal of the Audio Engineering Society* vol. 52, no. 11, pp. 1142-1156 (2004).
- 21 M.D. Burkhard and R.M. Sachs, "Anthropometric Manikin for Acoustic Research," *The Journal of the Acoustical Society of America* vol. 58, no. 1, pp. 214-222 (1975).
- 22 E. Benjamin, "Extending Quasi-Anechoic Electroacoustic Measurements to Low Frequencies," presented at The 117th AES Convention, San Francisco, CA, USA (2004).
- 23 IEC/TR 60959, *Provisional Head and Torso Simulator for Acoustic Measurements on Air Conduction Hearing Aids*, International Electrotechnical Commission (1990).
- 24 M. Slaney, "An Efficient Implementation of the Patterson-Holdsworth Auditory Filter Bank," in *Apple Computer Technical Report*, Apple Computer Inc. (1993).
- 25 R. Mason, T. Brookes, and F. Rumsey, "Frequency Dependency of the Relationship between Perceived Auditory Source Width and the Interaural Cross-Correlation Coefficient for Time-Invariant Stimuli," *The Journal of the Acoustical Society of America* vol. 117, no. 3, pp. 1337-1350 (2005).
- 26 R. Mason, T. Brookes, and F. Rumsey, "Integration of Measurements of Interaural Cross-Correlation Coefficient and Interaural Time Difference within a Single Model of Perceived Source Width," presented at The 117th AES Convention, San Francisco, CA, USA (2004).
- 27 C. Kim, *Towards a Head-Movement-Aware Measurement of the Spatial Impression of Audio*, PhD thesis, Institute of Sound Recording, University of Surrey (2009).
- 28 M. Morimoto and K. Iida, "A Practical Evaluation Method of Auditory Source Width in Concert Halls," *The Journal of the Acoustical Society of Japan* vol. 16, no. 2, pp. 59-69 (1995).
- 29 M. Morimoto and Z. Maekawa, "Effects of Low Frequency Components on Auditory Spaciousness," *Acustica* vol. 66, pp. 190-196 (1988).
- 30 T. Okano, T. Hidaka, and L.L. Beranek, "Relations between the Apparent Source Width (ASW) of the Sound Field in a Concert Hall and Its Sound Pressure Level at Low Frequencies (GL), and Its Inter-Aural Cross Correlation Coefficient (IACC)," *The Journal of the Acoustical Society of America* vol. 96, no. 5, pp. 3268 (1994).
- 31 B. Supper, *An Onset-Guided Spatial Analyser for Binaural Audio*, PhD thesis, Institute of Sound Recording, University of Surrey (2005).
- 32 T. Ajdler, et al., "Sound Field Analysis Along a Circle and Its Applications to HRTF Interpolation," *Journal of the Audio Engineering Society* vol. 56, no. 3, pp. 156-175 (2008).