Interaural time difference fluctuations:
their measurement, subjective perceptual effect, and application in sound reproduction

Russell Mason and Francis Rumsey
Institute of Sound Recording, University of Surrey
Guildford, Surrey, GU2 7XH, UK

Two objective measurement techniques have been proposed that relate the fluctuations in interaural time difference to one or more attributes of subjective spatial perception. This paper reviews these measurements, discusses how these fluctuations may be created in a real acoustical environment, summarises the experiments carried out to elicit the subjective effect of the fluctuations, and suggests ways in which this research can be applied to sound reproduction.

INTRODUCTION

The spatial performance of sound reproduction systems is becoming increasingly important with the advent of new distribution media that can deliver an enhanced spatial experience to the consumer. This field of research is relatively new compared to the research into topics such as distortion and timbre. Therefore the relationship between the physical parameters of the sound fields and the spatial attributes perceived by the subject is not fully understood.

Spatial sound has been studied in more detail in concert hall acoustics, resulting in a number of measurements that relate to various aspects of spatial impression. However, there are limitations to these measurements and it has been found that they are not as successful when applied to sound reproduction. If a measurement accurately matches a particular subjective attribute, then it is logical to expect it to work for all types of sound fields. As this is not the case for these measurements, the underlying physical parameters that create certain subjective effects need to be investigated further.

Griesinger proposed that interaural time difference (ITD) fluctuations are responsible for creating a subjective spatial impression [1]. ITD fluctuations are changes over time of the relative phase between the two audio signals measured at the ears. If the relative phase of these audio signals changes slowly, the perceived position of the sound will appear to move. If the phase of the audio signals varies more rapidly, the auditory system will fail to accurately track the movement, resulting in a perceptual effect termed 'localisation lag' [2]. Grantham and Wightman described this subjective effect as width or diffuseness [3].

This paper summarises a number of aspects related to ITD fluctuations. Firstly, the two proposed measurement techniques that aim to quantify the magnitude of ITD fluctuations are reviewed. The two techniques are compared and the main similarities and differences are noted. Secondly, the manner in which these fluctuations are created in real acoustic environments is investigated and the relative influence of the reflection pattern and the source signal is explored. Thirdly, the results of three experiments eliciting the subjective effect of various ITD fluctuations are summarised. Finally, the research is applied to various components in the sound recording and reproduction chain and the effect of each is discussed.

MEASUREMENT TECHNIQUES

Two measurement techniques have been proposed in recent years that attempt to measure the changes in interaural time difference (ITD) over time in order to create a measurement that relates to auditory spatial perception. These techniques are reviewed here and the similarities and differences are highlighted.

Diffuse Field Transfer Function (DFT)

Griesinger derived a measure of the envelopment at the listening position in a room that he termed the Diffuse Field Transfer function (DFT). The DFT is determined by a series of functions carried out in MATLAB software and is described in [4, 5].

The measure derives the ITDs between the two binaural channels by comparing the positive zero-crossing points of the audio signals. The resulting values are then averaged over a short period of time to give the running average. The running average is then weighted by signal amplitude in order to reduce the noise in the measure. This weighted running average is itself averaged to give a mean value for the entire source sample. The mean value is subtracted from the running average to give the interaural fluctuations over time. These fluctuations are filtered with a 3 to 17 Hz bandpass filter to distinguish the fluctuation frequencies that are most salient for envelopment. The strength of these fluctuations is the DFT.

Interaural Cross-Correlation Fluctuation Function (IACCFF)

An alternative measurement technique has been developed by the authors, which has been termed the interaural cross-correlation fluctuation function (IACCFF) [6]. This measurement attempts to quantify the fluctuations in ITD over time by using a concurrent series of interaural cross-correlation (IACC) calculations. The measure makes use of the property of the IACC calculation that the position of the maximum value of the IACC represents the ITD [7].

The IACCFF operates by initially splitting the binaural signal into 21 frequency bands covering the audio frequency range of approximately 50 to 2500 Hz. A concurrent series of IACC measurements is then used to determine the changes in the ITD in each of the frequency bands. This results in a plot of the changes in ITD over time. In order to reduce the noise in the measurement, the output is weighted by the audio amplitude. The result is then filtered with a 10 to 125 Hz bandpass filter to select the most salient fluctuation frequencies. The resulting ITD fluctuations with weighting and filtering are converted into the frequency domain using a Fourier transform. This depicts fluctuation frequency against the fluctuation magnitude of the measured signal. A mean of the magnitude can then be taken across the range of relevant fluctuation frequencies to give the IACCFF output result.

Discussion

The two measurements outlined above attempt to measure the same acoustical parameter, the magnitude of ITD fluctuations. However, they differ significantly in their calculation methods.

Some of the important similarities and differences between the two measurements are discussed below.

ITD fluctuation measurement technique

The main difference between the two measurements is the technique used to determine the ITD fluctuations in the audio signal. The DFT examines the positive zero-crossing points of the audio signal whereas the IACCFF uses a series of IACC measurements. There are advantages and disadvantages to each of these approaches.

The use of zero-crossing points is computationally efficient compared to the use of IACC calculations. It is also possible that it is more valid on a psychophysical basis as a neural spike is created by a auditory neurone at approximately the same phase of each cycle of a stimulus [8]. The use of a series of IACC measurements is computationally expensive, requiring a much larger amount of processing than the examination of zero-crossing points. In addition, the resolution of the measurement is limited by the window length of the IACC, which is required to be at least five times the period of the lowest audio frequency in order to avoid calculation error due to end effects [9].

However, the authors consider that the use of an IACC measurement is potentially more robust than the examination of zero-crossing points as the maximum ITD measurable by the latter is dependent on the frequency content of the audio signal. In addition, the use of IACC measurements may be less susceptible to noise for low level audio signals as the measurement is taken over a relatively long time window as opposed to the discrete single comparisons of zero-crossing times that are then averaged in the DFT.

However, it remains to be seen which of these techniques is more successful in measuring the magnitude of ITD fluctuations in a wide range of situations.

Noise

As mentioned above, both of the measurement techniques can be affected by the presence of noise in the binaural audio signal to be measured. If the noise is not identical in both channels, this can have a significant effect on the results, especially if measuring quiet or decaying signals. If the noise is decorrelated between the channels and random in nature, this will cause a large amount of ITD fluctuations to be measured which is difficult to differentiate from the ITD fluctuations in the signal that are the focus of the measurement. This is also important if the measurement calculates the fluctuations in a number of frequency bands and the source signal has a narrow frequency range. Such a situation would result in a number of the frequency bands containing mostly noise, which would affect the result of the analysis unless they are somehow disregarded.

One way to alleviate this is to weight the results by some form of signal amplitude or level measurement, such that instances of quiet noise with respect to the louder signal are effectively disregarded. This would improve the accuracy of the measurement with low-level noise, however it would not be able to correct a binaural recording that was excessively noisy.

An additional advantage of weighting the results by signal amplitude is that this will highlight fluctuations that are contained in the louder parts of the audio signal and which therefore may be more clearly perceivable. This assumption is yet to be tested in detail and will be investigated in due course.

Both of the measurement techniques outlined above make use of weighting by simple signal amplitude. Investigation is needed to determine whether a loudness measurement more similar to human perception would produce a result that matches the spatial perception more accurately.

Audio bandwidth

Both of the measurements summarised above only calculate the ITD fluctuations in audio signals up to a few kilohertz. This is based on research which has shown that audio frequencies above a few kilohertz do not affect the perception of spatial attributes, at least for time-based cues [10, 11, 12]. This is in agreement with the fact that the ear does not rely on phase or timing information for localisation above these frequencies, due to the ambiguous results caused by the short wavelengths compared to the interaural spacing [13] and the breakdown of phase-locking in the auditory neurones above these frequencies [8].

Frequency range of fluctuations

A difference between the two measurements outlined above is the range of fluctuation frequencies that are considered to be important for creating the spatial effect. The DFT measures across a range of 3 to 17 Hz whereas the IACCFF measures from 10 to 125 Hz. At the lower end of the fluctuation frequency scale it is likely that there is not a fixed frequency above which movement is no longer perceived due to binaural sluggishness, rather that there is a gradual transition from perceived movement to a perceived stationary impression as the fluctuation frequency rises. The work reported in [14] indicated that at least for some audio signals, movement was still perceived with ITD fluctuations of 10 Hz. However, this is not necessarily the case for more complex signals and for signals with ITDs that do not fluctuate sinusoidally.

The exact nature of the transition from a perception of movement to a stationary impression as the fluctuation frequency increases has not been investigated in detail for a wide range of source signals. Therefore it is uncertain which is the optimum lower fluctuation frequency to measure and this requires further investigation.

The work of Grantham and Wightman showed that a spatial effect was created by ITD fluctuations up to a fluctuation frequency of 500 Hz [3]. However, if a measurement attempts to quantify the fluctuations at these frequencies, there is the risk of the frequency components of the audio signal leaking into the measurement of ITD fluctuations, especially if a simple amplitude weighting is used to limit the effect of noise. In addition, the necessary window length of the IACC measurement as discussed above means that the resolution of the measurement at higher fluctuation frequencies is limited. Therefore the maximum measured fluctuation frequency is restricted to a lower frequency than that used by Grantham and Wightman.

It is expected that most source signals and acoustical environments (either real or reproduced) will create fluctuations across a wide range of fluctuation frequencies, therefore making the decision of the frequency range across which the fluctuations are measured less critical. However, the optimum range for a variety of source signals needs to be investigated.

Summary

This section of the paper has reviewed two measurements that aim to calculate the magnitude of ITD fluctuations in a given audio signal. The two techniques have been compared and the relative advantages and disadvantages of each were explored. This has uncovered a number of areas in which additional investigation is needed to understand the fluctuations more fully and to further develop the measurements. However, the measurements have been found to be similar in a number of ways and follow the same basic principal of calculating the magnitude of ITD fluctuations.

CREATION OF FLUCTUATIONS IN ACOUSTIC ENVIRONMENTS

ITD fluctuations are created in real acoustical environments, caused by the interaction of a direct source signal with the reflections from a number of boundaries or objects within that acoustical environment [15]. This is demonstrated in the following section using various levels of complexity of reflection patterns from a direct sound with a single reflection to an acoustical simulation of a simple room. The interaction of various reflection patterns and source signals and the differing ITD fluctuation characteristics caused by each are also investigated.

Artificial source signals with a single reflection

The creation of ITD fluctuations can be demonstrated with a direct sound and a single reflection. A simulation was created of a source 15 metres directly in front of a binaural receiver, with a reflection modelled as a wall parallel to the path of the direct sound at a distance of 5 metres to the right of the receiver. This is shown in Figure 1.

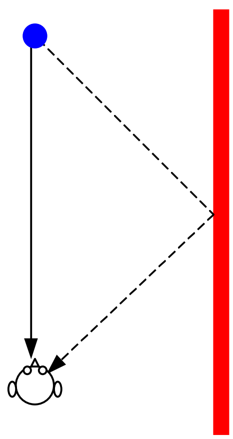

Figure 1: The arrangement of the source, wall and binaural receiver used to create a model of a direct sound from 15 metres directly in front of the receiver and a single reflection from a wall 5 metres to the right (not to scale).

The model was not intended as an accurate simulation, but aimed to approximate the salient cues for the purpose of demonstration. The source was assumed to be a perfect omnidirectional source, therefore decaying over distance according to the inverse square law [16]. The wall absorption was similar to painted concrete, with a fixed absorption coefficient of 0.1 at all frequencies [17]. The HRTFs used to simulate the binaural receiver were taken from the set of anechoic KEMAR measurements provided by Gardner and Martin [18]. As these were sampled at a resolution of 5° azimuth around the horizontal plane, the calculated reflection angle was rounded to fit this for simplicity of implementation.

Single sine tone

If a simple signal such as a single sine tone is passed through the above model, the direct sound and the reflection interact at the ears. The two signals, arriving from two directions and with different delays due to the time taken to travel the separate distances, interfere with each other. This causes summation or cancellation of the two signals reaching the ears, which may be different at each ear. The result is a phase and amplitude difference between the signals arriving at the ears. Figure 2 shows a 500 Hz sine tone passed through the model shown in Figure 1.

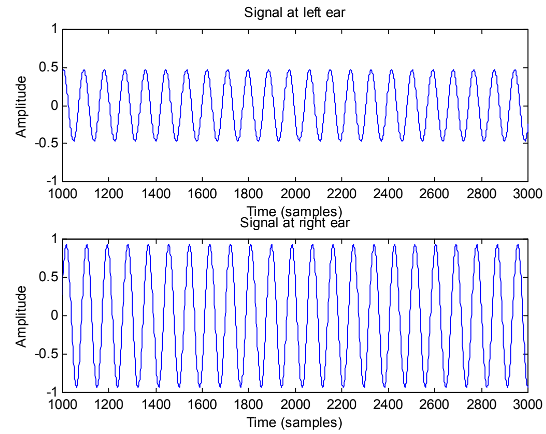

Figure 2: Plot of the audio signals arriving at each ear from a continuous 500 Hz sine tone with a direct sound from 15 metres in front of the receiver and a single reflection from a wall 5 metres to the right of the receiver.

As can be seen, the interaction caused by the direct sound and the single reflection causes a fixed phase and amplitude difference for a sound source consisting of a single sine wave. The magnitude of the phase and amplitude differences would not be the same in all cases, as this is dependent on the frequency of the source signal, the delay caused by the additional path length of the reflection, and the angle of incidence of the reflection. However, for a simple signal such as this, no ITD fluctuations are created for any configuration of these three variables.

Multiple sine tones

If a more complex source signal is used, then the frequencies present in the signal interact with each other. This interaction may cause phase and amplitude differences at the entrances to the ears which vary over time. This can be seen when a source signal of three sine tones of 480, 500 and 520 Hz is passed through the model shown in Figure 1. The resulting audio signals reaching the two ears are shown in Figure 3.

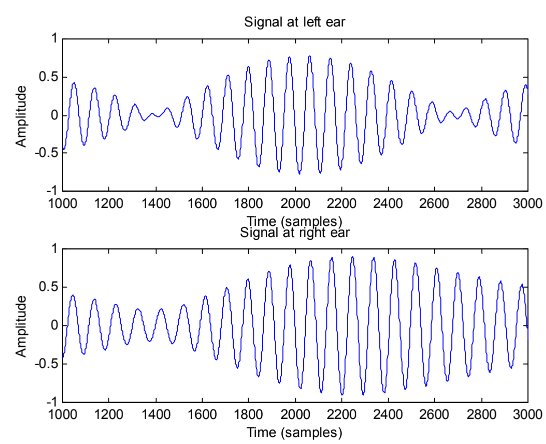

Figure 3: Plot of the audio signals arriving at each ear from three continuous sine tones of 480, 500 and 520 Hz with a direct sound from 15 metres in front of the receiver and a single reflection from a wall 5 metres to the right of the receiver.

It is apparent that both the interaural amplitude and interaural phase differences change over time[1]. The variation in interaural phase difference over time is apparent in Figure 3 as alternate leading and lagging in the phase of the two audio signals.

This fluctuation in interaural phase or time difference can be measured by using a consecutive series of IACC calculations as described in the previous section. The output of this measurement is shown in Figure 4. The ITD is shown on the y-axis for each calculation made over time on the x-axis. For clarity the fluctuations are filtered with a bandpass filter from 10 to 125 Hz and only the frequency band closest to the frequency of the sine tones is shown.

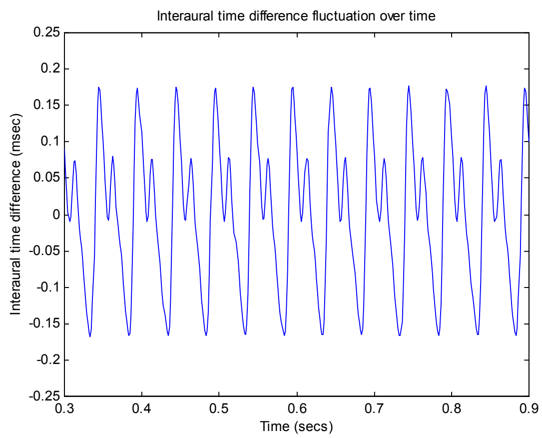

Figure 4: Plot of the interaural time difference fluctuations over time in the signal shown in Figure 3 calculated using a concurrent series of IACC calculations.

More information can be obtained if the amplitude of ITD fluctuations at different fluctuation frequencies is plotted by the use of a fast fourier transform (FFT). Again, for clarity, the FFT plot is 1/3rd octave smoothed. This is shown in Figure 5.

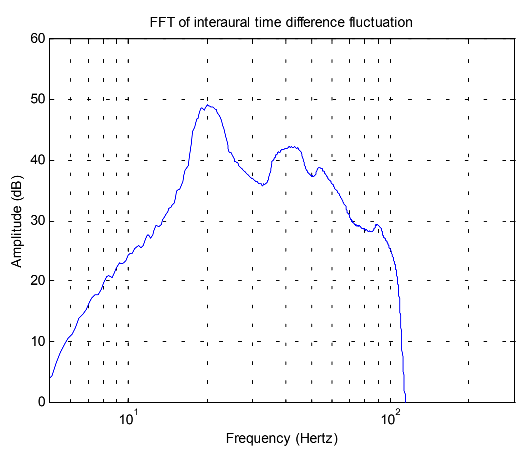

Figure 5: FFT plot of the interaural time difference fluctuations over time shown in Figure 4.

It can be seen that the predominant frequency of the ITD fluctuations is 20 Hz. This is the same frequency as the spacing between the three sine tones, indicating that the source signal influences the pattern of the ITD fluctuations.

Changing the distance between the side wall and the receiver (therefore increasing the delay and the angle of incidence of the reflection from the median plane) changes the pattern of the ITD fluctuations and therefore the resulting FFT. Figure 6 to Figure 11 show the fluctuations created by the three sine tones of 480, 500 and 520 Hz passed through a similar system of direct sound and single reflection as shown in Figure 1, with the distance between the receiver and the side wall increasing by 5 metres in each case.

It is apparent that in most cases the frequency of the ITD fluctuations with the largest FFT magnitude is 20 Hz. This indicates that the source signal is as important as the characteristics of the single reflection for creating ITD fluctuations.

It is also apparent that over the range of distances to the side wall, the magnitude of the ITD fluctuations do not rise or fall monotonically. The reason that the reflection pattern with a wall 15m to the right has a lower fluctuation magnitude than the others is most likely to be due to the delay of that particular reflection. In this case the reflection arrives at the ears of the receiver approximately 54ms after the direct sound. As this is almost equal to the period of the frequency spacing of the three sine tones, the created fluctuations are not as strong. This is further indication that the source signal affects certain characteristics of the ITD fluctuations.

Musical source signals with a single reflection

As the source signal has a significant effect on the resulting ITD fluctuations for a direct sound with a single reflection, then it is important to investigate the ITD fluctuations created by musical source signals. The musical signals differ from the sine tones used above in a number of ways. Firstly, they have a more complex frequency content than the sine tones. Secondly, the spectral content and pitch of the musical signals may change over the length of a note whereas the sine tones used above were constant. Thirdly, the amplitude of the musical signal may change over the length of a note in contrast to the continuous sine tones. Finally, a musical signal may contain a continuous musical phrase containing a number of different notes, each with individual frequency, envelope, amplitude and spectral characteristics.

Only single notes or chords were used for the purposes of this paper in order to limit the complexity of the analysis. A cello and an acoustic guitar were chosen due to their different temporal characteristics, with the cello being a relatively continuous bowed single note, and the acoustic guitar being a plucked chord that is allowed to decay. These instruments needed to be recorded in mono in an anechoic environment to prevent any recorded spatial or reverberant artefacts from affecting the results. The anechoic recordings used were from the collection on the Bang and Olufsen CD documented in [19]. The selected notes were edited from the musical phrases, and then processed through the same single reflection models used for the sine tone examples above.

Cello

The first note of the anechoic recording of the cello (Theme by Weber, Track 20 of the B&O CD) was used for this experiment. The sound is shown in Figure 12 as both a waveform and a spectrogram. The spectrogram is a plot of the amplitude by audio frequency over time. The darker shades in the spectrogram indicate larger amplitudes of signal at the associated frequency and time.

[1] Whilst the interaural amplitude differences may be a factor in the perception of some auditory spatial attributes, they are beyond the scope of this paper and will be investigated at a later date.

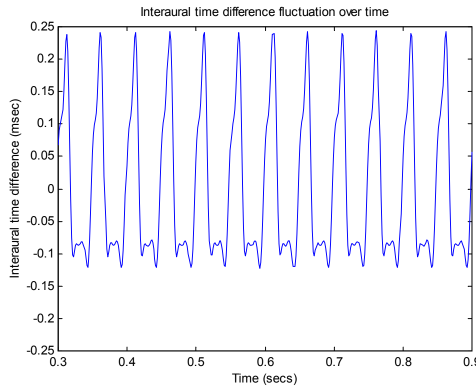

Figure 6: Plot of the ITD fluctuations over time in the signal arriving at each ear from three continuous sine tones of 480, 500 and 520 Hz with a direct sound from 15 metres in front of the receiver and a single reflection from a wall 10 metres to the right of the receiver.

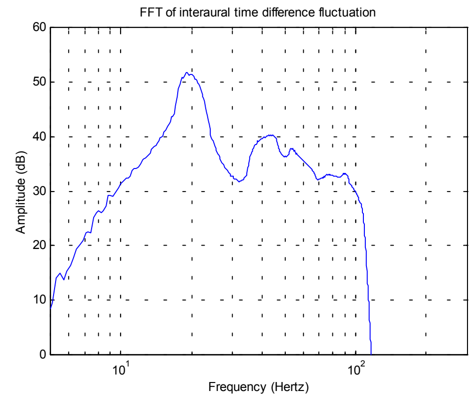

Figure 7: FFT plot of the ITD fluctuations over time shown in Figure 6.

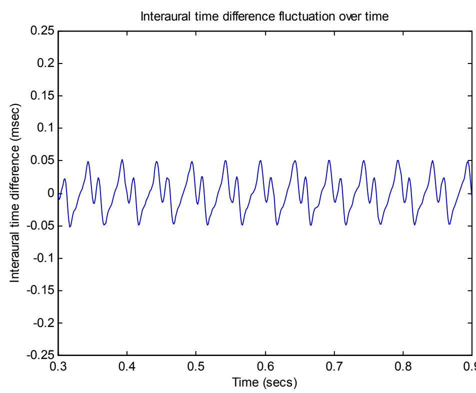

Figure 8: Plot of the ITD fluctuations over time in the signal arriving at each ear from three continuous sine tones of 480, 500 and 520 Hz with a direct sound from 15 metres in front of the receiver and a single reflection from a wall 15 metres to the right of the receiver.

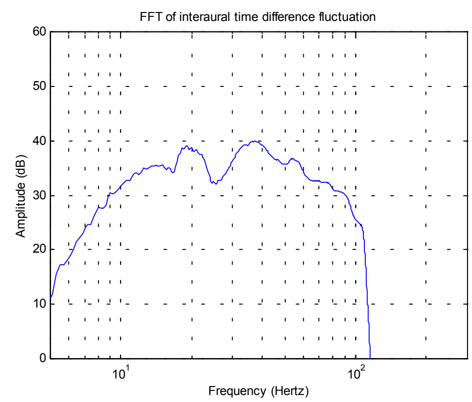

Figure 9: FFT plot of the ITD fluctuations over time shown in Figure 8.

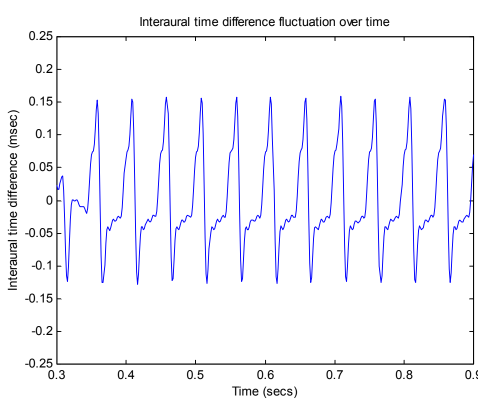

Figure 10: Plot of the ITD fluctuations over time in the signal arriving at each ear from three continuous sine tones of 480, 500 and 520 Hz with a direct sound from 15 metres in front of the receiver and a single reflection from a wall 20 metres to the right of the receiver.

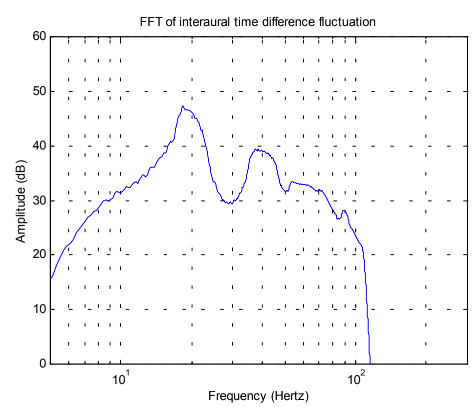

Figure 11: FFT plot of the ITD fluctuations over time shown in Figure 10.

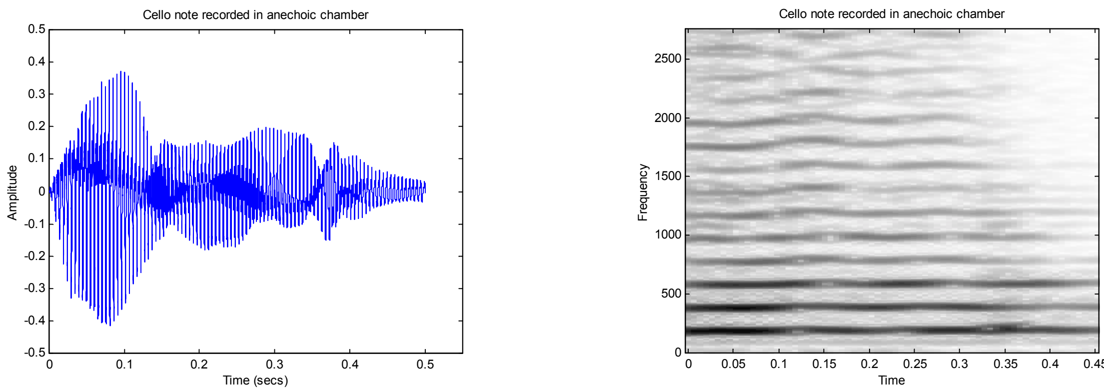

Figure 12: Waveform and spectrogram of the anechoic recording of a cello note.

The waveform shows an initial attack to the sound, followed by a more continuous section before a final decay. Throughout the note, the string is under continuous excitation by the bow, resulting in a changing amplitude and spectral content. This is due to the inconsistent pressure caused by the irregularity of the bow along its length and the action of the musician. In addition, it appears from the spectrogram that the pitch of the signal changes due to vibrato. The more complex waveform and these variations over time are in

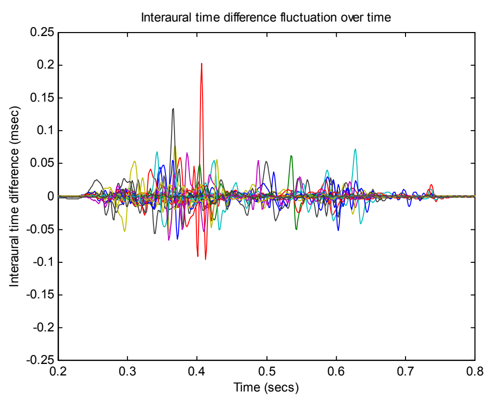

Figure 13: Plot of the ITD fluctuations over time in the signal arriving at each ear from an anechoic recording of a single cello note with the direct sound from 15 metres in front of the receiver and a single reflection from a wall 5 metres to the right of the receiver weighted by audio amplitude.

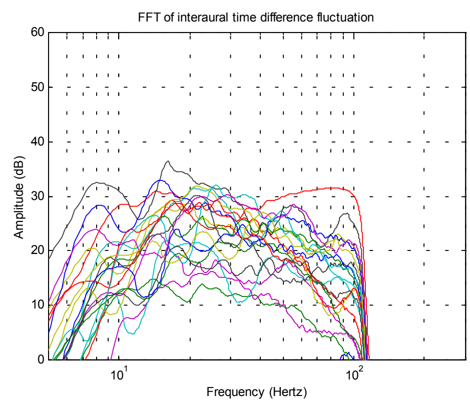

Figure 14: FFT plot of the ITD fluctuations over time shown in Figure 13.

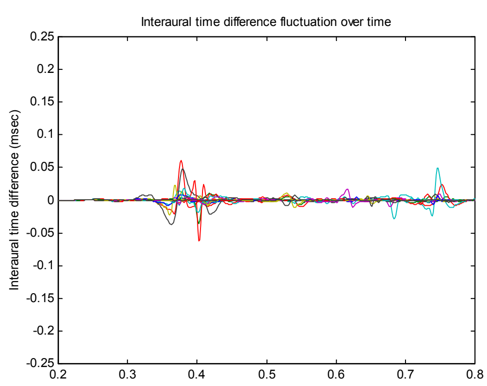

Figure 15: Plot of the ITD fluctuations over time in the signal arriving at each ear from an anechoic recording of a single cello note with the direct sound from 15 metres in front of the receiver and a single reflection from a wall 20 metres to the right of the receiver weighted by audio amplitude.

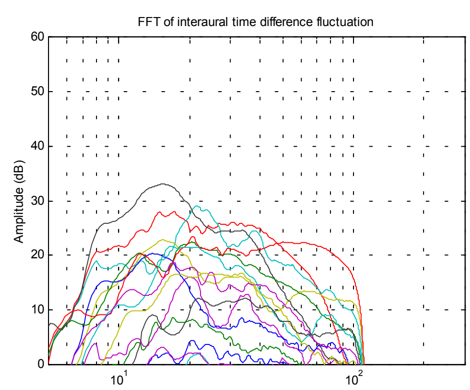

Figure 16: FFT plot of the ITD fluctuations over time shown in Figure 15.

contrast to the continuous sine tones used in the previous section and may cause the resulting the ITD fluctuations to be different.

The anechoic cello note was passed through the model used in the previous section, with the direct sound from 15 metres in front of the receiver and a single reflection from a painted concrete wall a given distance to the right of the receiver. The source signal was a much wider bandwidth than the three sine tones used in the previous section, therefore the ITD fluctuations of a number of frequency bands were measured. In addition, the source signal was not of a constant level which meant that the loudness of the sound over the time and in each of the frequency bands needed to be taken into account. This was done by the amplitude weighting employed by the IACCFF measurement and described above.

For comparison, the resulting ITD fluctuations from two different reflection patterns were measured. Figure 13 and Figure 14 show the results of a single reflection from a wall 5 metres to the right of the receiver combined with the direct sound from 15 metres away. Figure 15 and Figure 16 show the results when the single reflection is from a wall 20 metres to the right. Again, in both cases the results are filtered from 10 to 125 Hz for clarity.

The most obvious difference between these results and previous measurements of the sine tones is that the ITD fluctuations are not steady-state but vary greatly over the length of the sound. This is to be expected as the source signals differ in the same manner. There is also a large difference between the two reflection patterns used. It is apparent that the reflection from the furthest wall results in a lower magnitude of ITD fluctuations for this source signal. This can be seen in both the plots of ITD fluctuations and the associated FFT plots. This is most likely to be caused by the amplitude of the further reflection being lower due to the further distance travelled.

Acoustic guitar

The final chord from the anechoic recording of the acoustic guitar (Etude in A minor by F. Sor, Track 14 of the B&O CD) was used for this experiment. The waveform and spectrogram are shown in Figure 17.

There is an initial attack to the sound as the chord is plucked, followed by a decay of almost consistent spectral content and pitch. This can be seen in more detail in the spectrogram. It is apparent that the complexity of the spectral content is more similar to the cello note than the three sine tones. However, the constant spectral content and pitch is unlike the variation in the cello note and is more similar to the sine tones, albeit with the addition of an initial attack and a decay over the length of the chord. Therefore it is expected that the measured ITD fluctuations will contain features similar to both the cello note and the sine tones.

The signal was passed through the same model as used in the previous section, with a direct sound from 15 metres directly in front of the receiver and a single reflection from a wall a given distance from the right of the receiver. The ITD fluctuations of the resulting signal were then measured over a number of frequency bands, and the results weighted by audio amplitude and filtered between 10 and 125 Hz for ease of viewing.

For comparison, the ITD fluctuations of two reflection patterns were measured. The first, with a distance of 5 metres between the receiver and the wall is shown in Figure 18 and Figure 19 and the second, with a distance of 20 metres between the receiver and the wall is shown in Figure 20 and Figure 21.

The fluctuations for the acoustic guitar chord are more similar to the continuous sine tones than the cello note in that the fluctuations are continuous throughout the duration, albeit with a small decay. This is most likely to be due to the differences described above, such that the guitar chord is more continuous in terms of spectral content and pitch. Also apparent is that the frequency band with the largest magnitude of fluctuations changes during the length of the sound. This is likely to be due to the different spectral content of the initial attack (the pluck of the strings) compared to the resonating decay. The difference between the two reflection patterns is similar to the difference seen in the cello example above. For the acoustic guitar chord the magnitude of fluctuations for the model with the reflection from 20 metres to the right is again lower than for the reflection from 5 metres to the right, for the same reasons as outlined above.

Discussion

It is clear from the above examples that ITD fluctuations are created by the interaction of a direct sound and a least one reflection. It is also apparent that for a single reflection the characteristics of the source signal are just as important as the reflection delay, position and amplitude in determining certain aspects of the ITD fluctuations. The small selection of source signals tested show that the fluctuations created can have very different characteristics over time, and the subjective effect of these needs to be determined. In order to investigate the complex interaction in more detail, further analysis is needed to uncover the salient factors of the reflection pattern and source signal which govern the specific frequency and magnitude of ITD fluctuations

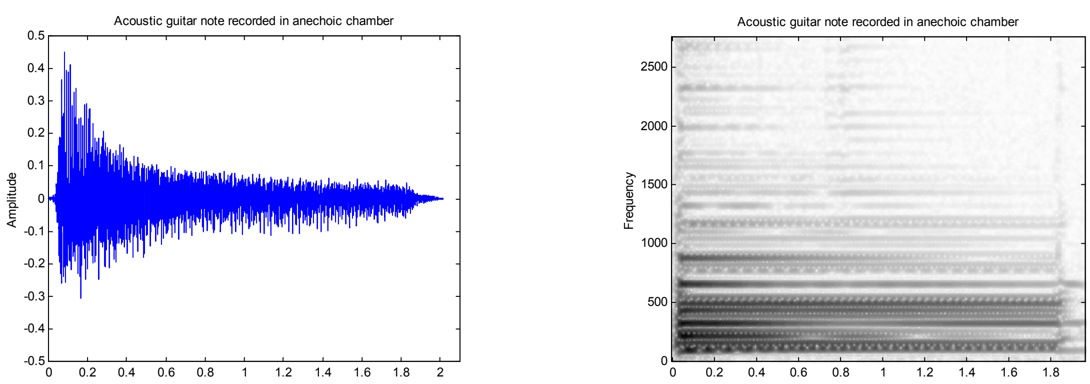

Figure 17: Waveform and spectrogram of the anechoic recording of an acoustical guitar chord.

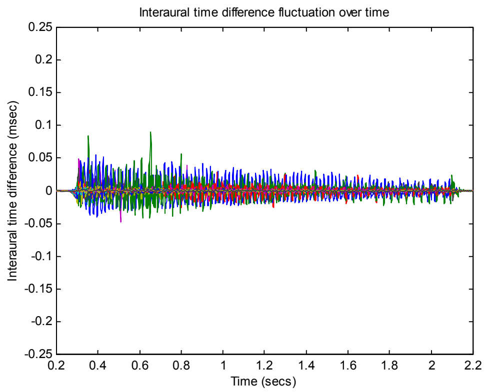

Figure 18: Plot of the ITD fluctuations over time in the signal arriving at each ear from an anechoic recording of a single acoustic guitar chord with the direct sound from 15 metres in front of the receiver and a single reflection from a wall 5 metres to the right of the receiver weighted by audio amplitude.

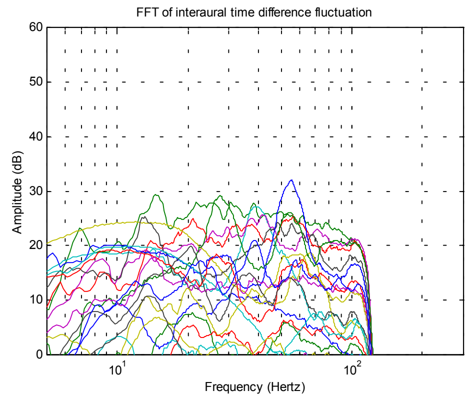

Figure 19: FFT plot of the ITD fluctuations over time shown in Figure 18.

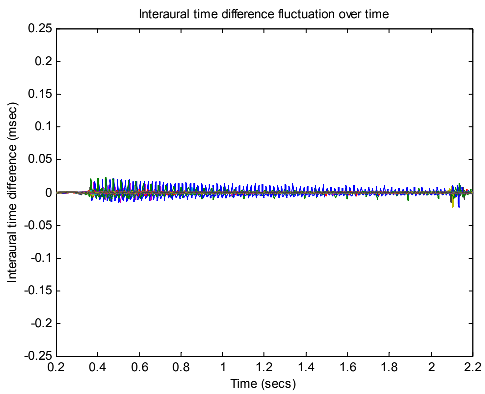

Figure 20: Plot of the ITD fluctuations over time in the signal arriving at each ear from an anechoic recording of a single acoustic guitar chord with the direct sound from 15 metres in front of the receiver and a single reflection from a wall 20 metres to the right of the receiver weighted by audio amplitude.

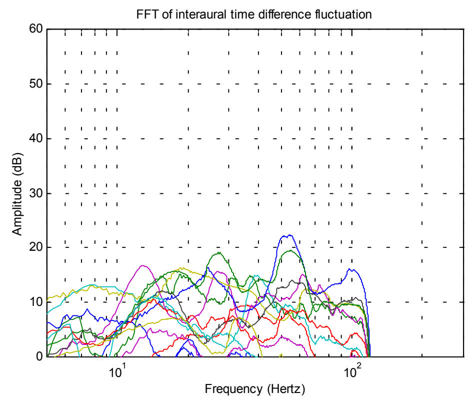

Figure 21: FFT plot of the ITD fluctuations over time shown in Figure 20.

Artificial source signals with a pair of symmetrical reflections

To further investigate the creation of ITD fluctuations, the above example can be repeated with more reflections. In this case, two reflections were combined with a direct sound. The simulation was created in a similar manner to that used previously, with the direct sound from 15 metres in front of the binaural receiver, but in this case with two symmetrical reflections, from walls 5 metres to the left and right of the receiver. This is shown in Figure 22.

As above, an artificial signal of three sine tones of 480, 500 and 520 Hz was passed through the model. The ITD fluctuations of this were measured in the same manner as the examples above, and the results of the frequency band closest to the frequency of the sine tones is shown in Figure 23.

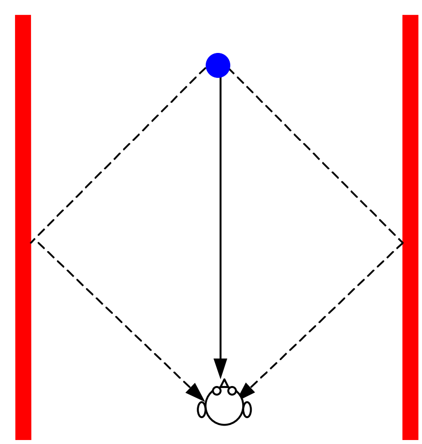

Figure 22: The approximate arrangement of the source, wall and binaural receiver used to create a model of a direct sound from 15 metres directly in front of the receiver and a pair of symmetrical reflections from walls 5 metres to the left and right (not to scale).

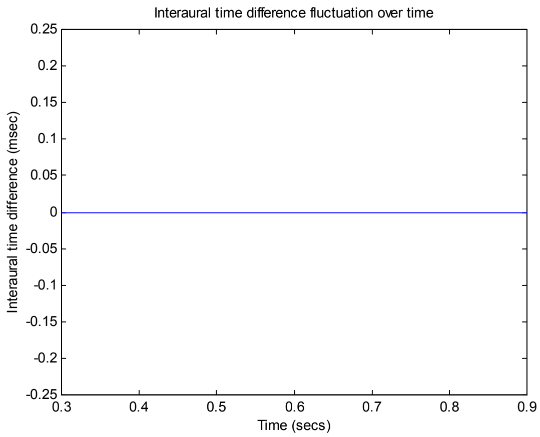

Figure 23: Plot of the ITD fluctuations over time in the signal arriving at each ear from three continuous sine tones of 480, 500 and 520 Hz with a direct sound from 15 metres in front of the receiver and a pair of symmetrical reflections from walls 5 metres either side of the receiver calculated using a concurrent series of IACC calculations.

It is clear from Figure 23 that no ITD fluctuations are created by the interaction of the direct sound with a symmetrical pair of reflections. This is to be expected, as the interaction between the direct sound and reflections in this model will be symmetrical, causing no differences between the signals arriving at the ears.

However, when attempting to apply this result to the real world, a number of important factors must be considered. Firstly, the human head is unlikely to be completely symmetrical, though it is uncertain whether the asymmetry is significant enough to cause ITD fluctuations to be created. Secondly, it is unlikely that the human head will be located at the exact centre between two reflective surfaces, and it is even less likely that it will remain completely stationary. Finally, whilst the fluctuations are not created by a pair of symmetrical reflections from the side, it remains to be seen whether additional lateral reflections or reflections from other directions interact further to create ITD fluctuations.

Simulated acoustical environment with off-centre source and receiver

Whilst the examples above with a single reflection indicate that ITD fluctuations are created by the interaction of a direct sound and at least one reflection, examination of more complex reflection patterns are needed to examine how this relates to real acoustical environments. For this purpose, a simulation of a simple room was created in CATT-Acoustic.

The acoustical space was a simple shoe-box room with dimensions and reverberation time similar to the Grosser Musikvereinsaal in Vienna [20]. The simulated sound source was omnidirectional and was deliberately positioned off-centre to avoid any potential problems caused by symmetrical lateral reflections as highlighted in the previous section. The sound source was located 6 metres to the left of centre, 4 metres from the back wall, and 1.8 metres from the floor. The receiver was a simulated binaural receiver positioned 11 metres directly in front of and facing the source. This can be seen in Figure 24.

The impulse response was calculated in CATT-Acoustic for convolution with the test and musical source signals to create a simulation of the sound being emitted by the source and captured by the binaural receiver. The results for each source signal are shown below.

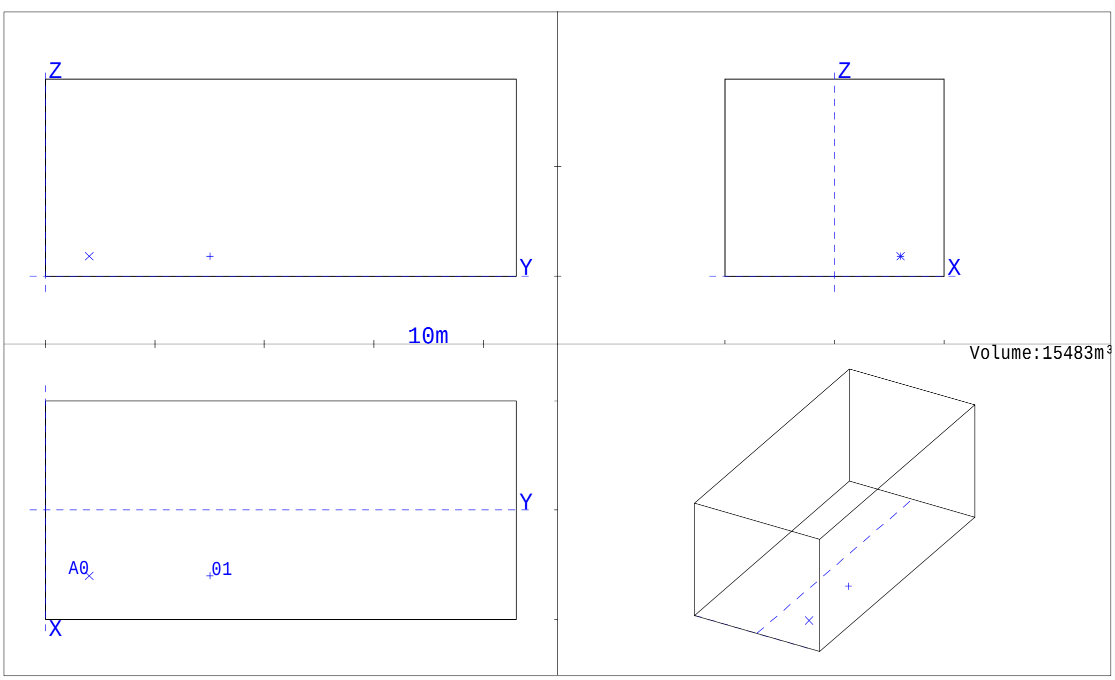

Figure 1: Side elevation, end elevation, plan and three-dimensional views of the simple shoe-box room used in the simulation, showing the omnidirectional sound source (labelled A0 and marked with ×) and the binaural receiver (labelled 01 and marked with +) positions.

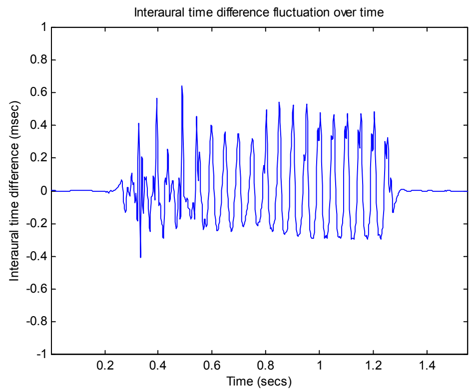

Figure 26: Plot of the ITD fluctuations over time in the active sound source segment of a signal arriving at each ear from three continuous sine tones of 480, 500 and 520 Hz convolved with the impulse response of the room simulation.

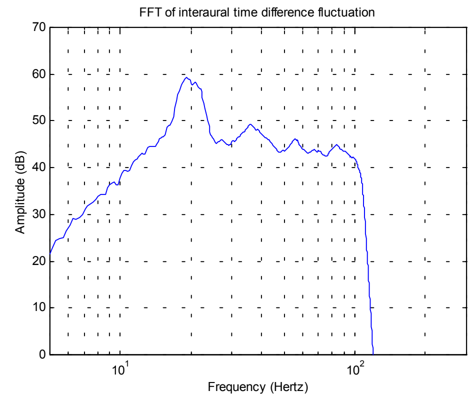

Figure 27: FFT plot of the ITD fluctuations over time shown in Figure 26.

Multiple sine tones

The first signal to be measured was the three sine tones of 480, 500 and 520 Hz. This was convolved with the impulse response of the simulated room. The resulting binaural audio signal was measured in the same manner as the previous section, using a series of IACC calculations to quantify the magnitude and frequency of ITD fluctuations over a fluctuation frequency range of 10 to 125 Hz. The measured ITD fluctuations are shown in Figure 25.

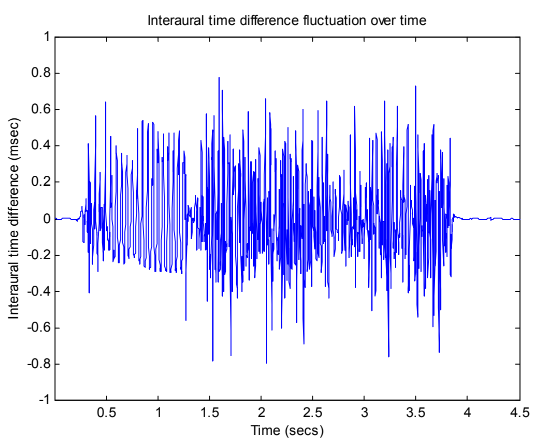

Figure 25: Plot of the ITD fluctuations over time calculated in the signal arriving at each ear from three continuous sine tones of 480, 500 and 520 Hz convolved with the impulse response of the room simulation.

It can be seen in Figure 25 that in the active sound source segment[2] (c. 0.25 to 1.3 secs) the fluctuations are periodic and of relatively constant amplitude. In the reverberant segment[3] (from c. 1.3 secs onwards) the fluctuations become much more random in terms of frequency and amplitude. As these sections are so different, it may be beneficial to analyse them separately. This is shown in Figure 26 to Figure 29.

It is apparent from Figure 26 that the fluctuations are fairly erratic over the first part of the active sound source segment. This is caused by the early reflections of the beginning of the sine tone reaching the receiver and each individually affecting the ITD fluctuations as the reflection pattern develops. After approximately 0.6 seconds in the plot (and approximately 0.3 seconds after the start of the sine tone) the fluctuation pattern is more constant, almost reaching a steady state, as the additional reflections after this point are much lower in level and therefore contribute less to the resulting signal.

It can be seen in Figure 28 that the fluctuations become more erratic in the reverberant segment once the direct sound ends. This is caused by the system of reflections no longer being driven by the active sound source emitting a steady-state signal. In addition, the steady-state direct sound is no longer reaching the receiver, leaving only the complex decaying reverberation.

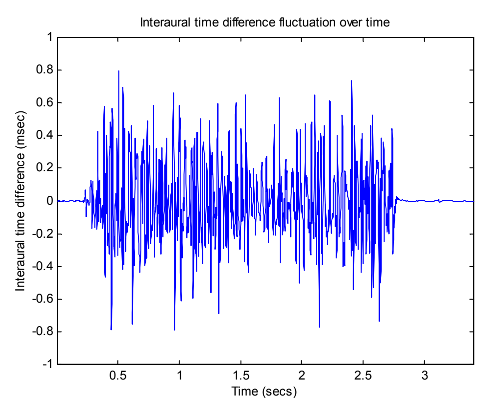

Figure 28: Plot of the ITD fluctuations over time in the reverberant segment of a signal arriving at each ear from three continuous sine tones of 480, 500 and 520 Hz convolved with the impulse response of the room simulation.

[2] For the purpose of simplicity in the paper, the phrase 'active sound source segment' refers to the segment of time when there is a direct sound component reaching the receiver. During this time, the sound captured by the receiver is made up of both the direct sound and reverberation and in this case is the segment of time where the sine tones are playing.

[3] Again, for simplicity in the paper, the phrase 'reverberant segment' refers to the segment of time where the direct sound has ended and the sound reaching the receiver is purely reverberant.

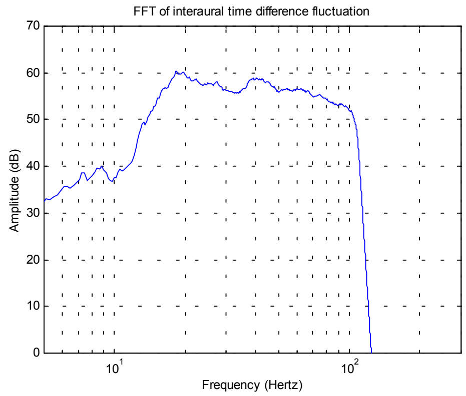

Figure 29: FFT plot of the ITD fluctuations over time shown in Figure 28.

Also of interest is that the interaction frequency of the sine tones can be seen in the FFT of the measured ITD fluctuations of the active sound source segment, though this is not apparent in the measurement of the reverberant segment. This appears to indicate that the fluctuations in the reverberant segment are less related to the original source sound than in the active sound source segment.

As the measurements shown in Figure 26 to Figure 29 have no weighting by audio amplitude, the result is not an accurate representation of the perceptual effect of the fluctuations, assuming that the ITD fluctuations in quieter audio signals will not be as clearly perceived. This would be more accurately displayed by the use of weighting by audio amplitude, however, this representation is useful to indicate the extent of the fluctuations contained in the decaying signal, which would otherwise be concealed by the amplitude weighting.

Musical source signals - Cello

For musical signals, the difference between the ITD fluctuations in the active sound source segment and the reverberant segment is less obvious. This is due to the fact that during the note, the source signal changes in frequency, spectral content and amplitude and is not steady-state as the sine tones used in the previous example. The measurement results of the anechoic recording of the cello note convolved with the impulse response of the simple shoe-box hall are shown in Figure 30 to Figure 33, separated into the active sound source and reverberant segments. Due to the nature of the complex frequency and amplitude characteristics of these signals, these measurements were made with weighting by audio amplitude in order to reduce the effect of noise in frequency bands that did not contain audio signal, as mentioned above.

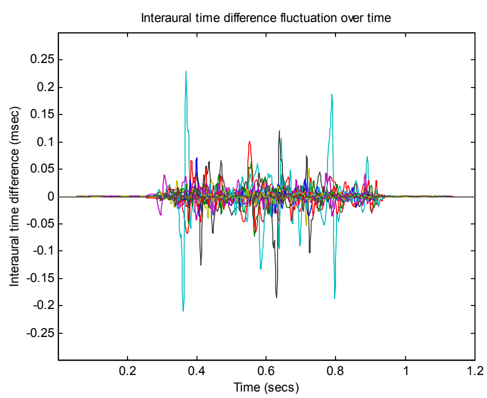

Figure 30: Plot of the ITD fluctuations over time in the active sound source segment of a signal arriving at each ear from an anechoic recording of a single cello note convolved with the impulse response of the room simulation.

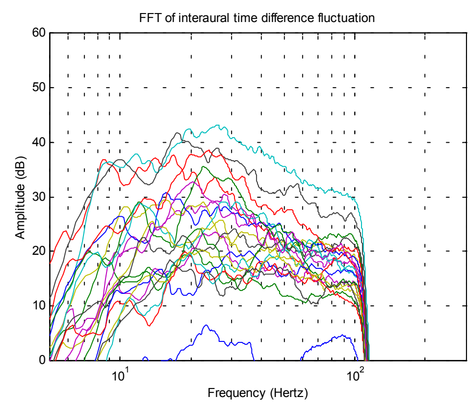

Figure 31: FFT plot of the ITD fluctuations over time shown in Figure 30.

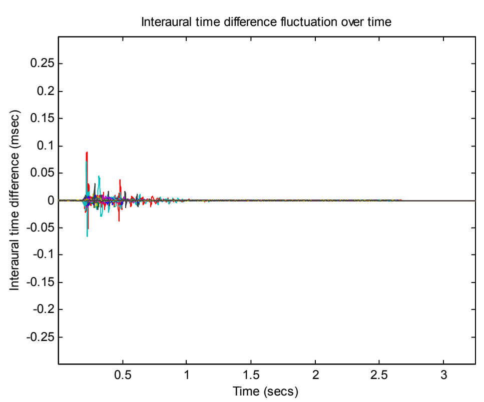

Figure 32: Plot of the ITD fluctuations over time in the reverberant segment of a signal arriving at each ear from an anechoic recording of a single cello note convolved with the impulse response of the room simulation.

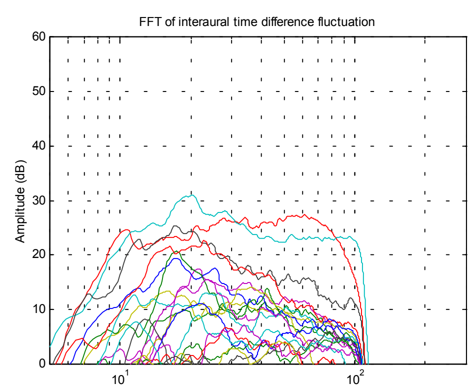

Figure 33: FFT plot of the ITD fluctuations over time shown in Figure 32.

For the active sound source segment, the fluctuations appear to be quite similar to those caused by a single reflection. There are differences, in both the magnitude of fluctuations and the pattern in each frequency band, however there is no remarkable difference in the overall form of the fluctuations.

The difference in the reverberant sound segment is more obvious in this example due to the weighting by audio amplitude. This superimposes the decay in amplitude on the measured fluctuations, therefore giving the misleading impression that the fluctuations are decaying. This weighting by audio amplitude is necessary to reduce the effect of noise on the measurement as discussed above. This weighting also causes the magnitude of the FFT to be lower, therefore the measured ITD fluctuations in the two segments cannot be compared directly.

Musical source signals - Guitar

The measurements were repeated for the anechoic recording of the acoustic guitar chord convolved with the impulse response of the simple shoe-box room simulation. Again, the analysis was carried out independently on the active sound source and reverberant segments of the resulting binaural signal in order to examine the effect of each separately. The measurements were made with weighting by amplitude for the same reasons as the cello example. The ITD fluctuation plots and associated FFT plots are shown in Figure 34 to Figure 37.

For the active sound source segment it is apparent that ITD fluctuations are very different in the first section (up to approximately 1 second) of the chord as opposed to the following section (after approximately 1 second). The differences include a large variation in the magnitude of the fluctuations in addition to different frequency bands containing the most prominent fluctuations. The fluctuations being in different frequency bands can also be seen in the example of the guitar chord with a single reflection in Figure 18 which is discussed above.

The increased fluctuation magnitude over this section is likely to be due to the attack of the chord. This is a louder sound with a less tonal frequency characteristic. These features mean that the early reflections are excited over a wider frequency range, causing a larger amount of ITD fluctuations. In addition, the faster rise time of the guitar chord compared to the cello note may also cause larger fluctuations to be created. This will be investigated further in due course. As the attack ends relatively quickly, its early reflections decay, leaving the resonating strings and its associated reverberation.

In the section of time for the active sound source segment measurement after approximately 1 second, the ITD fluctuations are again relatively periodic and constant in amplitude. This is similar to the measurement of the guitar note with a single reflection, and more similar to the three sine tones than the cello

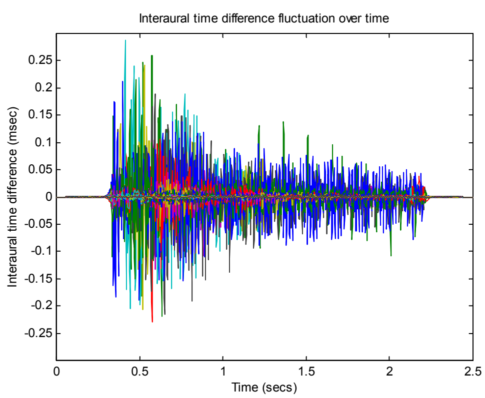

Figure 34: Plot of the ITD fluctuations over time in the active sound source segment of a signal arriving at each ear from an anechoic recording of a single acoustic guitar chord convolved with the impulse response of the room simulation.

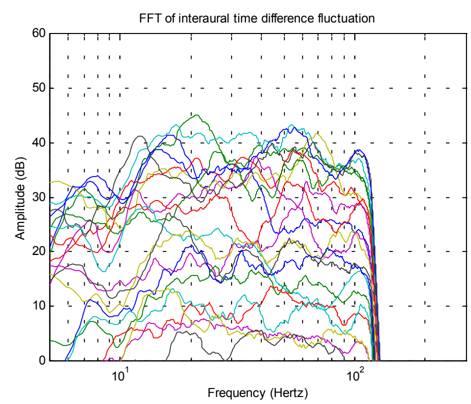

Figure 35: FFT plot of the ITD fluctuations over time shown in Figure 34.

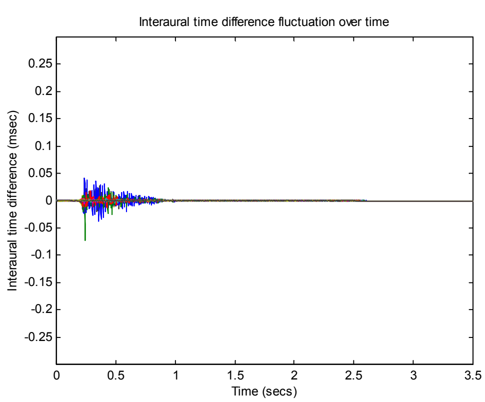

Figure 36: Plot of the ITD fluctuations over time in the reverberant segment of a signal arriving at each ear from an anechoic recording of a single acoustic guitar chord convolved with the impulse response of the room simulation.

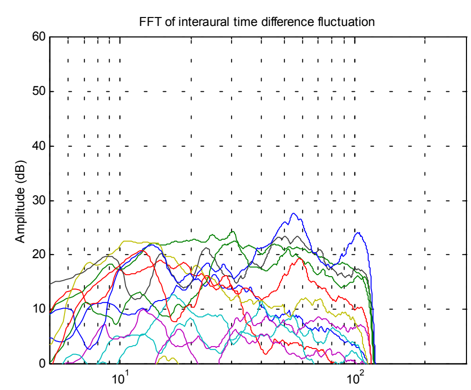

Figure 37: FFT plot of the ITD fluctuations over time shown in Figure 36.

due to the near-constant pitch, amplitude and spectral content.

The appearance of the measurement of the reverberant segment is again strongly affected by the weighting by the amplitude of the audio signal. Due to this, the results in Figure 36 and Figure 37 do not appear to be significantly different to the results for the reverberant segment of the cello note.

Discussion

This section of the paper has outlined how ITD fluctuations are caused by the interaction of a direct sound and at least one acoustic reflection. It is apparent from the results that both the source signal and the reflection pattern have a significant effect on the characteristics of the ITD fluctuations that are created. This has been found to be true for both single reflections and complex reflection patterns in a simulated room.

Several aspects of the source signal have been found to be important. These include how the pitch, amplitude and spectral content of the sound change over time. Other specific examples of attributes that affect the created ITD fluctuations are non-tonal attacks such as plucking a string and frequency modulation such as vibrato. It has also been found that these changes in the source signal can cause the properties of the ITD fluctuations to change significantly over a short space of time. This may prove to be an important factor in the measurement of these fluctuations and therefore a dynamic measurement of the fluctuations may be preferable. In addition, the perceptibility of these changes needs to be studied.

The relationship between the audio frequency and the delay time of the reflections has also been found to be important. Whilst this may not be a major factor in room acoustics where there are a large number of reflections with different delays, it is important for situations with a small number of reflections.

It has also been found that for a pair of symmetrical reflections, no ITD fluctuations are created due to the interactions being identical at each ear. Whether this is the case for a large number of reflections or when reflections are also incident from angles close to the median plane needs to be investigated in detail. However, in natural listening situations it is unlikely that the source and binaural receiver will remain stationary exactly on a central line. Nonetheless, it may be an important consideration for creating deliberate effects and when generating artificial reflection patterns.

It has also been found that depending on the source signal, the ITD fluctuations in the active sound source segment have different characteristics compared to the reverberant sound source segment. This is clearer for source signals that vary little over the length of the sound, but is present for all the source signals tested. Because of this difference, it may prove useful to analyse these segments separately in the measurement, and also to analyse the associated subjective effect separately.

Further investigation of the effect of source signals and reflection patterns on the creation of ITD fluctuations is required. However, it has been shown that the fluctuations are created by the interaction of a direct sound and its reflections, and therefore the calculation of the magnitude of these fluctuations may be a viable basis for an acoustic measurement.

SUBJECTIVE EFFECT OF ITD FLUCTUATIONS

As mentioned above, if the relative interaural phase of an audio signal fluctuates slowly, the subjective effect will be a change in the perceived position of a sound. However, if the fluctuations occur at a frequency above a few Hertz, the perception will no longer be movement due to the perceptual effect of 'binaural sluggishness' or 'localisation lag' [2]. Grantham and Wightman researched this effect, and found that ITD fluctuations at a rate of greater than approximately 20 Hz caused the perception of width or diffuseness instead of movement [3]. Griesinger described the subjective effect as a 'stationary source in the presence of a surround' [21]. However, the precise subjective effect caused by ITD fluctuations had not been investigated in detail, and was therefore elicited in a series of experiments.

For these experiments, artificially created noise signals with predetermined sinusoidal ITD fluctuations were used, comprising of various fluctuation rates and magnitudes. Artificial signals were used in order to allow accurate control of the physical parameters of the fluctuations without introducing confounding variables such as characteristics of the musical extract, bias caused by preconceptions of the spatial attributes of natural sound sources, the recording technique or additional processing of pre-recorded extracts.

The stimuli were 2-channel noise-like samples similar to those used by Grantham and Wightman [3]. Each stimulus consisted of a series of pairs of sine tones that were frequency modulated with a sinusoidal fluctuation signal that was 180º out of phase in one of the pair with respect to the other. When one of the pair of tones was fed to the left ear, and the other to the right ear, the frequency modulation resulted in a sinusoidal fluctuation in the ITD between the two signals. This was created by the relative frequency of the sine tone rising and falling in one of the pair of sine tones with respect to the other, subsequently causing leading phase and then lagging in each cycle.

The equation for each pair of sine tones is shown below[4].

$$l = \sin[2\pi f_c t + \theta_c + m\sin(2\pi f_m t)]$$
$$r = \sin[2\pi f_c t + \theta_c - m\sin(2\pi f_m t)]$$

where l is the left ear signal
r is the right ear signal
f_c is the audio frequency
θ_c is a random phase component (identical in each ear)
m is the fluctuation magnitude
f_m is the fluctuation frequency

When a large number of pairs of frequency modulated sine tones were reproduced simultaneously with a range of audio frequencies (f_c) and random starting phases (θ_c), the subjective effect was similar to a noise signal. In order to give more equal loudness precedence to the lower audio frequencies which are most important for creating spatial attributes such as envelopment [22], the stimuli were created with a pink noise-like frequency response (equal power in each octave band).

Values of f_c were used from 40 to 2560 Hz, a range of 6 octaves. The upper limit was based on previous research that suggests that audio frequencies above a few kilohertz do not alter the magnitude of the spatial effect [10]. Various values of fluctuation magnitude (m) and fluctuation frequency (f_m) were used as required for each experiment.

Continuous stimuli delivered over headphones

The first experiment (described in [14]) investigated the subjective effect of continuous stimuli with specific ITD fluctuations created as described above. In order to control accurately the time difference fluctuations reaching the ears, these stimuli were delivered to the subjects over headphones. There were nine stimuli used, consisting of three levels of fluctuation magnitude (m) of

[4] For a more detailed discussion of the method of creating the noise samples with sinusoidal ITD fluctuations, the reader is referred to [3, 14]

0.075, 0.15 and 0.3 and three levels of fluctuation frequency (f_m) of 5, 10 and 100 Hz.

The experimental method used was a graphical sketch-map technique. This required the subjects to draw the perceived spatial attributes of the stimulus using a graphics tablet on a computer-based experimental interface. The interface included a plan view and a side view, and both moving and stationary scene components[5] could be depicted.

The results showed that for the stimuli with a fluctuation frequency of 5 Hz, the subjects perceived a moving scene component. The movement was perceived to be less clear for the stimuli with a fluctuation frequency of 10 Hz, and no movement was perceived for the stimuli with a fluctuation frequency of 100 Hz. In all of these cases there was a stationary scene component depicted in addition to any moving scene component.

For the lowest level of fluctuation magnitude (m = 0.075), no movement was perceived for any fluctuation frequency and the stationary scene component depicted by all the subjects was relatively small and located at the centre of the head. For the higher two levels of fluctuation magnitude (m = 0.15 and m = 0.3), the increase in fluctuation magnitude resulted in either the trajectory of movement being represented wider for the stimuli for which moving scene components were depicted, or the stationary scene component being represented larger for the stimuli for which no movement was depicted.

However, there were differences in the depicted results between the individual subjects. Firstly, the trajectory of the movement depicted varied between the subjects. Most of the subjects depicted the movement as being directly between the ears, though some indicated that the scene component was moving outside the head or around the head. Secondly, the size and position of the depicted stationary scene components varied between the subjects, although the trend of increasing size with increasing fluctuation magnitude was consistent for all the subjects.

The use of headphone reproduction for this experiment meant that the ITD fluctuations could be accurately controlled, but it also limited the external validity. A majority of the scene components were depicted as located either completely or partially within the head, which is unlike natural listening and sound reproduction where scene components are generally perceived to be outside the head. Therefore, to widen the scope of the experiment, the subjective effect of ITD fluctuations with scene components located outside the head needed to be investigated.

Continuous stimuli delivered over loudspeakers

To generate scene components that were located outside the head, an experiment was carried out using similar continuous noise stimuli with predetermined sinusoidal time difference fluctuations, however in this case they were presented over loudspeakers. This is described in detail in [23]. Whilst the time difference fluctuations created in this case were inter-loudspeaker and not interaural, they resulted in interaural time difference fluctuations for a subject at the listening position. The stimuli were similar to those used in the experiment described in the previous section, although in this case there was a fixed fluctuation frequency (f_m) of 100 Hz and three levels of fluctuation magnitude (m) of 0.1, 0.7 and 1.3. Measurement of the resulting signals at the listening position showed that ITD fluctuations were created that varied in magnitude in the same manner as the stimuli.

The stimuli were presented over two different pairs of loudspeakers. The first pair was at ±30º from the median plane, and the second was at ±90º. This enabled comparison of the subjective effect of the fluctuations reproduced from two positions.

The subjective effect of these stimuli was elicited using both non-verbal and verbal methods. The first elicitation section used a combination of non-verbal sketch-map and verbal descriptor methods. The non-verbal method required the subjects to draw their spatial perception of the stimuli on a plan view. This was conducted using pencil and paper in order that the sketching was as natural as possible for the subject. The subjects were provided with a number of colours to enable them to indicate different aspects of the scene. The verbal descriptors were used to complement this, with the subjects free to use any terms they required to describe the stimuli.

The second elicitation section used a method similar to the Repertory Grid technique. The stimuli were presented in pairs and the subjects were asked to give pairs of descriptors that described the differences between them. The elicited pairs of descriptors were then used as the end points of scales for a grading experiment. For this, each subject evaluated the same stimuli on the scales that they had created in the elicitation stage of the experiment.

The results of the first elicitation section showed that for the increasing levels of fluctuation magnitude, the main difference in the graphical results was an increase in the depicted width and depth of the stimuli. As there was little indication of any perception of an environment in either the verbal or non-verbal results, it was concluded that the change in perceived width and depth was related to the dimensions of the perceived sound source. The main difference observed between the continuous stimuli presented over the separate pairs of loudspeakers was that the stimuli presented over the loudspeakers at ±30º were perceived to be further away in front of the subject than the stimuli presented over the loudspeakers at ±90º.

This was supported by the results from the second elicitation section. A content analysis showed that the most common terms were those related to width and envelopment. It is interesting that these two terms appeared to be used similarly by the subjects, indicating that in this case they may refer to the same underlying attribute. Whilst the term envelopment is more commonly used to describe the perception of an acoustic environment, the fact that there was little mention of an environment in either elicitation section indicated that in this case the term was used to describe an aspect of the sound source. There was an important difference between the results of the first and second elicitation sections – in the second elicitation section there was only one mention of a depth attribute, indicating that width was the principal factor and that width and depth had possibly been confused in the first elicitation section.

Decaying stimuli delivered over loudspeakers

A further subjective experiment was carried out to evaluate the subjective effect of stimuli with specific time difference fluctuations presented over loudspeakers, this time with a decay in amplitude which imitated the decay of reverberation. This is described in detail in [24]. The stimuli were similar to those used in the experiment described in the previous section, but in this case the continuous noise signals were modified by one of two amplitude envelopes simulating the reverberation time of two separate acoustical spaces. This resulted in stimuli with a fluctuating inter-channel time difference that decayed in amplitude in a manner similar to each of the simulated acoustic spaces. The

[5] For this paper the term 'scene component' has been used instead of the more common terms of 'sound source' or 'sound object'. This is to differentiate that in reproduced sound the source of the sound is in fact usually loudspeakers or headphones, and that for more abstract signals such as noise, separate components may be perceivable with different attributes, though they are part of the same 'object'.

fluctuating signals were presented over loudspeakers located at ±90º from the median plane. Again, the time difference fluctuations created were inter-loudspeaker and not interaural, but they resulted in interaural time difference fluctuations at the listening position that varied in magnitude in a similar manner to the stimuli. A fixed fluctuation frequency (f_m) of 100 Hz and three levels of fluctuation magnitude (m) of 0.1, 0.7 and 1.3 were used. These values were the same as used in the previous experiment to enable comparison of the results.

In addition to the fluctuated signal produced from the side loudspeakers, a mono decaying signal and an impulse were generated from a loudspeaker located directly in front of the subject. This resulted in an initial impulse followed by a decay, similar to an impulse sounded in a reverberant room.

The elicitation methods used were the same as the experiment described in the previous section – firstly, a non-verbal graphical response method with supporting verbal descriptors, and secondly, a relative descriptor method similar to Repertory Grid.

The results of the first elicitation section were that for the various levels of fluctuation magnitude, the main difference observed in the graphical results was a change in the width and depth of the depicted scene. As the majority of cases depicted both a perceived sound source and an acoustical environment, these were analysed separately. From this, it was apparent that the size of the perceived sound source did not change with increasing fluctuation magnitude, but the perceived width and depth of the acoustical environment increased.

These results were supported by the second elicitation section. Content analysis of the pairs of elicited descriptors showed that the most common terms were those related to width, spaciousness and envelopment, although in this case the source and the environment were not differentiated. It is interesting to note the similar use of width, spaciousness and envelopment in this experiment, indicating that in this case they may be a similar underlying attribute. Width and envelopment were mentioned in the previous experiment with continuous stimuli as being attributes of the sound source. It appears that in this case they are attributes of the environment, as indicated by the results in the first elicitation section, and by the additional term spaciousness, which is commonly used in concert hall acoustics and in this case appears to be used in a similar manner to width and envelopment.

The authors consider that the terms envelopment and spaciousness are more qualitative in nature, and the term environment width is more quantitative. An example of this is that width can be measured from a sketch map plan (which is inherently spatial) whereas envelopment and spaciousness cannot be directly measured but may be interpreted from the measured results.

As for the previous experiment, there was no mention of a depth attribute in the second elicitation section, indicating that width was the principal factor and that in the graphical elicitation section the two spatial attributes, width and depth, had possibly been confused.

Discussion

It is apparent from the results summarised above that increasing the magnitude of ITD fluctuations causes an increase in the perceived width of a certain attribute of a scene, depending on the nature of the signal. For scene components with low frequency ITD fluctuations that are perceived to be moving, increasing the magnitude of the fluctuations appears to result in the trajectory of movement widening. For continuous scene components with ITD fluctuations of a higher fluctuation frequency that are perceived as a stationary sound source, increasing the magnitude of the fluctuations appears to result in the perceived size of the source widening. For decaying scene components with ITD fluctuations of a higher fluctuation frequency that are perceived as a stationary reverberance or environment, increasing the magnitude of the fluctuations appears to result in the perceived size of the environment widening.

This can be applied to sound reproduction as follows. If ITD fluctuations are created in a sound that is perceived to be a sound source component then these will create a certain perceived width. Also, if ITD fluctuations are created in a sound that is perceived to be reverberance or an acoustical environment, then these will create a certain perceived width of environment or spaciousness. Therefore the problem is separating the measured fluctuations that are related to the sound source or the environment in a meaningful way, that is consistent in order to allow comparison between results, and that ideally can be automated within a measurement.

The method of separating source-related and environment-related attributes commonly used in concert hall acoustics is to measure the impulse response of the room, and separate the early and late parts of the decay at a time of 80ms after the arrival of the direct sound at the receiver. The characteristics of the early reflections before 80ms are then considered to be an indication of the width of the sound source, and the remaining reverberation after 80ms is considered to be an indication of the envelopment of the hall [20, 25, 26]. However, for a continuous sound, the reflections arriving at the receiver much later than 80ms after the direct sound can significantly affect the created ITD fluctuations, as is evident in Figure 26. As this is a continuous sound that may mask the reverberation of the acoustical environment, it is likely that these fluctuations will be perceived as an attribute of the sound source. Therefore it appears that the separation of the impulse response at 80ms is not sufficient.

This separation of aspects of the auditory scene and their associated physical parameters into separate components or streams is well documented [27]. Griesinger has theorised that there are three groups in concert hall listening, the early, continuous and background spatial impressions [21]. This may be further subdivided depending on the number and type of source signals present. As there are many factors to consider in the grouping of streams and the grouping may in fact be context dependent, it is a difficult task to model this in a measurement.

It may be a simpler task for a system to measure specific source signals that are known to be perceived a certain way and measure aspects of these. For instance, to quantify the subjective width of the sound source, a continuous sound source such as a sustained note could be measured over the duration of the active sound as used above. The result of this measurement would then be related to the attributes of the sound source, as long as it is known that the subjective effect of all the related auditory attributes over this time will be perceptually grouped as part of the sound source. To quantify the subjective width or spaciousness of the acoustical environment or reverberation, either the decay of the continuous sound or an impulsive signal could be measured as used above. The result of this measurement would then be related to the attributes of the environment, as long as it is known that the subjective effect of all the related auditory attributes over this time will be perceptually grouped as part of the environment.

Whilst it may be convenient to measure the fluctuations in any sound that is occurring as part of a musical performance or background sound, currently the grouping of these sounds and the separate measurement of each of these groups is a task which cannot be automated. Therefore pre-prepared source signals may be preferable.

The stimuli used in the subjective experiments described above were artificial noise signals that were presented over headphones or loudspeakers and were either continuous or decaying. There are a large amount of differences between these signals and the acoustic environments, reproduction systems and programme material to which a measurement based on ITD fluctuations may be applied.

The first major difference is that the separation of continuous or decaying sounds is a simplification of the complex temporal changes caused by a musical phrase or speech in an acoustical environment such as a concert hall. Whilst this study has provided a useful indication of two specific conditions in a simplified manner, further detailed investigation is needed to uncover the subjective effect of more complex signals.

The second major difference is that the noise signals are different to the programme material that may be produced in an acoustic environment or through a reproduction system to which a measurement based on ITD fluctuations may be applied. This includes the programme material being mostly tonal as opposed to the wide frequency range of the noise, and the programme material containing dynamic changes both within words or musical notes and across phrases. Finally, the noise signals contained sinusoidal ITD fluctuations, which are dissimilar to the fluctuations measured for the various sources in the simulated acoustical spaces shown above. However, despite these differences, the authors believe that the results from these relatively simple stimuli can be applied to certain aspects of conventional programme material.

APPLICATION OF INTERAURAL TIME DIFFERENCE FLUCTUATIONS IN SOUND RECORDING AND REPRODUCTION

There are a large number of parameters in sound recording and reproduction that can affect the final result perceived by the listener. This section of the paper considers the application of the ITD fluctuation research to the three main sections of the process as outlined in Figure 38. These are the sound recording, any additional processing and the sound reproduction.

Within this paper it would be impossible to consider every form of sound recording and reproduction. Therefore a single scenario is considered, attempting to create for the listener a sense that 'they

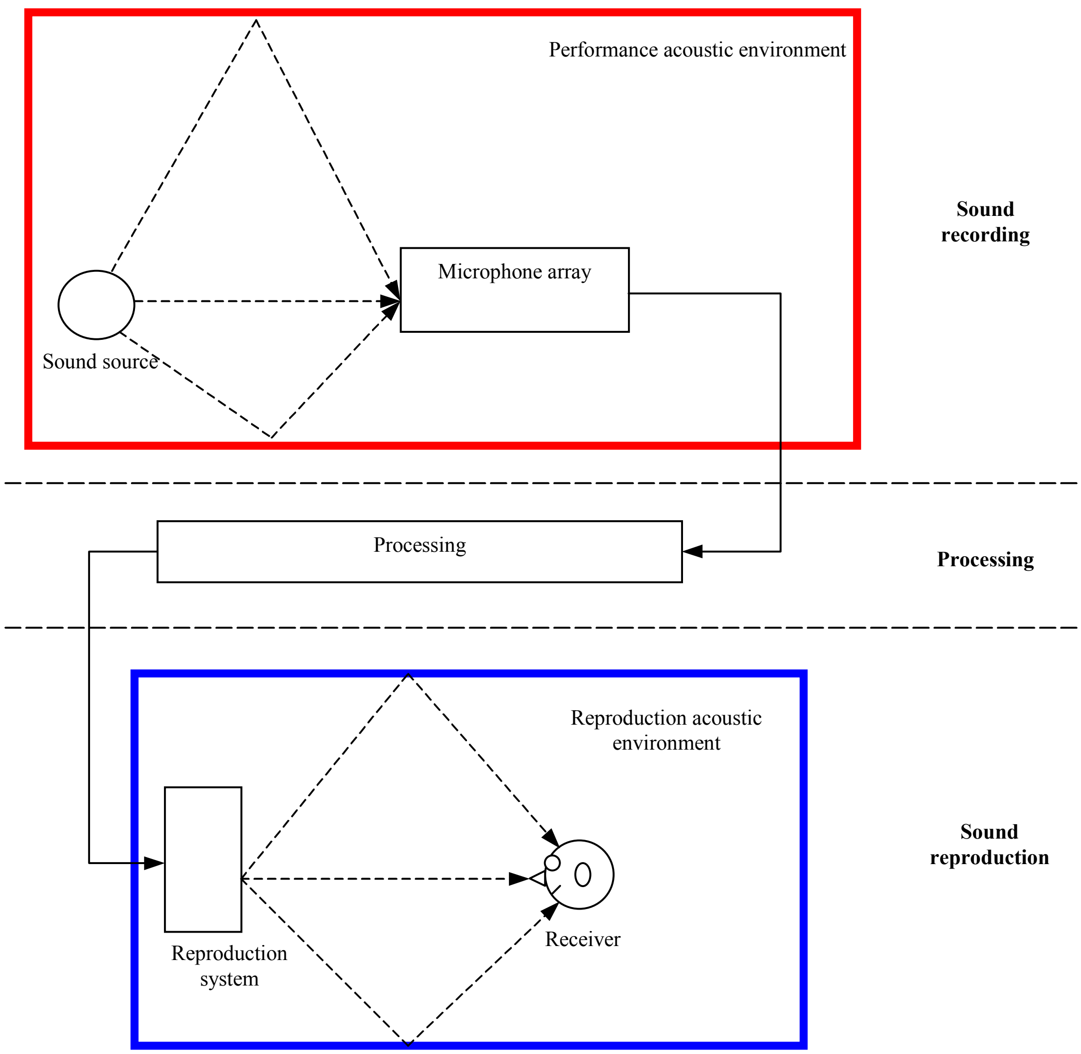

Figure 38: Block diagram of the principle factors in the sound recording and reproduction chain that may affect the perceived overall sound quality.

are there' in the performance acoustic environment. Whilst this is not the only type of illusion that may be created with a sound reproduction system, it is a common example used for demonstrating the performance of sound recording and reproduction, and the principles may be expanded to more complex situations. The aim of this investigation is not to accurately recreate the experience of being in the performance space, but to examine how each of the factors in the recording and reproduction chain may affect the ITD fluctuations created at the ears of the listener. This is then related to the subjective effect of source width or environment width as discussed above.

A number of the many variables in sound recording and reproduction have been fixed to limit the complexity of the example. The sound is limited to a single musical source positioned directly in front of the listener, with an accompanying impression of the acoustics of the performance space. The microphone technique is assumed to consist of a simple main array and processing is assumed to be minimal. The distribution medium is assumed to be discrete 5-channel surround sound.

Sound recording

There are a large number of attributes within the recording part of the sound recording and reproduction chain that affect the final perceived result for the listener. The three variables in the sound recording that are considered here are the sound source, the performance acoustical environment, and the microphone technique.

Sound source

The main factors related to the sound source that may affect the creation of ITD fluctuations are the directivity and position of the source, and the content of the sound source signal.

The directivity of the source affects the pattern of reflections that reach the receiver, which in turn can affect the fluctuations created at the receiver. This is complicated by the fact that musical instruments can have very different directivities at different frequencies, sometimes changing a large amount over a small frequency range. In addition, performers often move their instruments a certain amount in a performance, altering the reflection pattern further. It is unlikely that the reflection pattern created by the directivity or motion of the sound source could be used as a creative effect to alter the created ITD fluctuations, however it may be a factor in explaining why the spatial attributes of certain musical instruments are perceived differently.

The position of the sound source could be altered deliberately to increase or limit ITD fluctuations. It was found above that simple symmetrical reflection patterns created no ITD fluctuations due to the similarity of the signals reaching each ear. This was not expanded to examine more complex reflection patterns such as performance spaces, however it may be that moving a sound source from the centre line of a hall may increase the magnitude or change the characteristics of the ITD fluctuations created at a receiver.

Finally, it was demonstrated above that the characteristics of the sound source signal have a large effect on the attributes of the ITD fluctuations that are created. This includes aspects such as the spectral content, the attack and decay, the pitch and the amplitude of the sound and how these change over the duration of a note or musical phrase. Musical features such as vibrato may increase the ITD fluctuations created, however apart from this it may be difficult to use as a creative effect. Nevertheless, the fluctuations created by the attributes of the musical instrument may be a factor in the inherent perceived spatial attributes of that instrument.

Acoustical environment

The acoustical environment in which a sound recording takes place may be deliberately chosen, or may be imposed by the particular performance. However, this is an important factor in creating ITD fluctuations. The design of concert hall acoustics has been developed over many years, and the importance of spatial impression was recognised and researched in this field long before the current interest in the spatial attributes of sound reproduction.

There are measurements that have been proposed to quantify the spatial performance of concert halls, such as LF and IACC [20]. It may be that the measurement of ITD fluctuations will be an additional useful tool to assist in designing and quantifying the quality of concert halls and performance spaces. This will be investigated further in the future. However, there will not be a uniform magnitude of ITD fluctuations created at all positions in the acoustical space, and this should be considered as part of the microphone and source placement.

Microphone technique

The microphone technique is where the recording engineer potentially has most control over the recorded sound, and is currently the focus of a great deal of research on surround sound. Maximising or minimising the ITD fluctuations created at the ears of the listener may be used creatively to alter the related perceptual effects of source width and environment width.

It is a convenient categorisation to consider that the front and rear microphone techniques affect the source and environment attributes individually. This means that the pattern of the microphones feeding the front loudspeakers of a 5-channel reproduction system will only affect the attributes of the source, and the microphones feeding the rear loudspeakers of a 5-channel reproduction system will only affect the attributes of the environment. Whilst it has been shown that these factors interact with each other and therefore should not be considered separately in this manner, it is a useful grouping in order to analyse certain attributes of microphone techniques.

To affect the perceived attributes of the source, the microphone technique feeding the front channels can be deliberately chosen to minimise or maximise the ITD fluctuations created in the reproduction. To create the minimum amount of ITD fluctuations, a number of microphones can be fed to the central loudspeaker or a single microphone can be fed to either one or more of the front loudspeakers. The result is that the front channels of the recording are highly correlated with each other. If these are replayed over loudspeakers that are symmetrical about the median plane of the listener, then the signals arriving at each ear of the listener will be similar and therefore minimal ITD fluctuations will be created. In this situation, only factors after the recording stage, such as deliberate processing or the attributes of the reproduction system, will cause ITD fluctuations to be created at the ears of the listener.

To generate a large amount of ITD fluctuations, an array of spaced microphones can be used, each of which is reproduced from a different phantom or real location in the reproduced sound field. It may also be an advantage to make the array asymmetrical, both in terms of microphone position and the position of the microphones in the reproduced sound field as this will help to create more interaural differences and therefore a larger magnitude of ITD fluctuations.

The use of coincident or spaced microphone techniques can also affect the ITD fluctuations created at the ears of the listener. Due to the decreasing directivity of most microphones at low audio frequencies, coincident microphone techniques are increasingly correlated at low frequencies. This high correlation at low frequencies means that the signal is similar to the mono signal discussed above, and therefore will create minimal ITD fluctuations. Using a spaced microphone technique may minimise the correlation between the channels at certain audio frequencies, depending on the distance between the microphones as discussed by Griesinger [28]. This may therefore increase the magnitude of ITD fluctuations.

This discussion is primarily concerned with varying the magnitude of ITD fluctuations and disregards other attributes such as localisation accuracy and timbre, which are essential factors in the overall quality of the sound recording. However, this discussion is important as by understanding the physical parameters that affect how humans perceive recorded and reproduced sound, an optimum combination of the salient attributes can be reached. This is discussed further below.

The techniques used to maximise or minimise the ITD fluctuations for the front channels of a surround sound array may be applied to the rear channels. However in this case, assuming a concert hall layout with no discrete sources behind, the correct localisation of the rear portion of the sound may be less critical than for the front channels. This gives more flexibility in the design of the microphone array feeding the rear channels.

Processing

Whilst this paper has concentrated on the ITD fluctuations created by the interaction of a direct sound and a number of acoustic reflections, the fluctuations can also be created in sound reproduction by the addition of delayed signals to an original signal. For this technique, the delayed versions of the signal must be positioned in a different real or phantom location to the direct sound. In addition, for a small number of delayed signals, the pattern needs to be asymmetrical about the median plane of the listener as demonstrated above. This enables the creation of ITD fluctuations as a creative effect.

A similar effect is generated when using spot microphones in combination with a main microphone array. The sound from an instrument will reach the spot microphone first, followed by any spot microphones on other instruments, before finally reaching the main microphone array. If no artificial delay is used to compensate for this, then there will be a number of versions of the direct sound of the instrument reproduced in the listening room at slightly different times caused by the summation of the feeds from the various spot and main microphones. If the microphones are panned to different positions then this may create additional ITD fluctuations. In addition to this, Theile suggested adding a number of delayed versions of spot microphone feeds to preserve the spatial attributes of a main microphone array [29]. This may also be related to the creation of ITD fluctuations and may benefit from this research.

Artificial reverberation may also be used to increase the number of reflections or alter their pattern in a sound recording. This can be used creatively to increase the magnitude of the ITD fluctuations created at the ears of the listener, based on the principles discussed above such as asymmetric reflections in terms of delay time and intended perceived location. Optimisation of reverberation algorithms may also benefit from this research.

Sound reproduction

There are a large number of parameters of the sound reproduction that may affect the ITD fluctuations that are created at the ears of the listener. The possible influence of each of the main factors is discussed below. However, the characteristics of the programme material must be taken into account and whereas for one item of programme material a change may increase the magnitude of ITD fluctuations created at the listening position, for another item of programme material an identical change may reduce the magnitude of ITD fluctuations created. This is discussed in more detail below.

Loudspeaker layout

The layout of the loudspeakers is a major factor in the ability of the reproduction system to create ITD fluctuations. If all the loudspeakers are positioned in a similar location, or all five delivery channels are reproduced through one loudspeaker, the fluctuations created will be minimal and will be solely created by the attributes of the loudspeaker and the reproduction space.

Asymmetry in the reproduction arrangement may also cause an increased magnitude of fluctuations to be created, as demonstrated with the asymmetrical reflection patterns. However, this is dependent on the characteristics of the programme material as moving from the central ‘sweet spot’ may result in a lower magnitude of ITD fluctuations created at the listening position as found in [23].

Griesinger investigated the ITD fluctuations created by a number of different loudspeaker positions [4, 30, 5] and concluded that lateral loudspeakers create the largest magnitude of ITD fluctuations. However, more work is needed to fully understand the interaction of loudspeaker position with various items of programme material and what layout of loudspeakers creates the largest magnitude of fluctuations for a wide range of programme material.

Loudspeaker type

Various parameters of loudspeaker design may also have an affect on the creation of ITD fluctuations at the ears of the listener.

The type of driver used in a loudspeaker may affect the created ITD fluctuations. It may be that the extended resonance of distributed mode loudspeakers will create additional fluctuations, as may the wide dispersion pattern and the decorrelation of the off-axis radiation caused by the random resonance of the panel [31].

The directivity of the loudspeaker may have an affect in combination with the acoustics of the listening room. The more omnidirectional the loudspeaker, the more it will excite the listening room acoustic, and the more dependent the result will be on that particular acoustic [32]. The excitation of acoustic reflections in the listening room may create additional ITD fluctuations that are not part of the recorded signal. This especially applies to dipole loudspeakers with the null towards the listener, which are recommended for use for the rear channels.

Listening room acoustics

The acoustics of the listening room will also be a factor in the creation of ITD fluctuations. This situation is the same as the performance acoustic environment, with each of the loudspeakers being a source which produces a direct sound that reaches the listener together with the associated reflections of that source from the objects and boundaries of the listening room. These reflections and the content of the direct sound produced by the loudspeaker will interact in the same manner as those demonstrated above, therefore possibly affecting the magnitude and characteristics of the ITD fluctuations contained in the programme material.

As mentioned above, the extent of this effect will depend on the directivity of the loudspeakers, though it may be significant for all types of loudspeakers. The additional reflections are inevitable unless the listening room is anechoic or specifically treated. The magnitude of the additional fluctuations created by the listening room will depend on the absorption and diffusion of the acoustics, and the directivity of the loudspeakers. Griesinger demonstrated that the ITD fluctuations created by a simulation of a single driver in a small room lowered as the wall absorption coefficient rose [5]. Most critical listening rooms and control rooms are acoustically treated to control the early reflection pattern and the reverberation time. However consumer listening environments can be very different to this and research is therefore needed to understand how this affects the sound perceived by the listener.

The audibility of the additional fluctuations created by the listening room acoustics has not been quantified. The temporal characteristics of the recorded programme material and the ITD fluctuations created by the programme material may mask any fluctuations created by the listening room acoustics. In addition, the research of Olive and Toole showed that the threshold of perception of a single reflection was dependent on the characteristics of the programme material [33]. Therefore it is likely that the audibility of the additional fluctuations is dependent on the programme material. This requires further detailed investigation.

Optimum contribution of the sound reproduction system

This section has concentrated on how the magnitude and characteristics of ITD fluctuations can be manipulated by certain parameters in the sound reproduction system. However it is not yet clear what the optimum contribution of the sound reproduction system should be.

It may be preferable that the sound reproduction takes place in an anechoic space so as not to add any additional fluctuations to those created by the recording. However, this is impractical for the consumer listening environment, as the high cost is prohibitive. It may be preferable that the reproduction system (including the loudspeaker positions, loudspeaker type, acoustical environment and listening position) is made similar to the control room used for the recording and mastering of the programme material, therefore ensuring that the listener experiences the sound as it was intended by the producer or musician. However, similarity in all the factors would be difficult to attain and may have to be different for each production or record label. In addition, it may be that the situation in the control room can be improved upon, especially in view of the limiting conditions of some location recording venues.

Therefore the optimum contribution of the sound reproduction system to the magnitude of ITD fluctuations needs to be evaluated, possibly with the calculation of a range of acceptable conditions. In addition, the relative contribution of each of the factors needs to be analysed.

Discussion

As mentioned above, the purpose of this section of the paper was to discuss how each of the sections in the recording and reproduction chain can affect the ITD fluctuations created at the ears of the listener. It has not been considered how this relates to the overall perceived sound quality or the subjective preference of the listener.

Preference or perceived sound quality is generally accepted to be a combination of a large number of perceived attributes [34]. Of these, the relative importance of the perception of auditory source width or environment width compared to more established factors such as dynamic range, bandwidth and distortion has not been determined. However, spatial attributes have been found to be an important factor in subjective preference of sound reproduction systems [35].

It is likely that the preferred magnitude of ITD fluctuations will be dependent on the characteristics of the programme material. Whilst for some programme material it may be preferable to create a magnitude of ITD fluctuations as large as possible, this may not be suitable for other programme material. This may be further subdivided by considering the optimum magnitude of fluctuations associated with each of the separate sources or the acoustical environment. It is possible that the optimum magnitude of fluctuations will be different for each of these.

It is also possible for the magnitude of fluctuations to be too high, therefore resulting in a source that is perceived to be overly wide, or an acoustic environment that is perceived to be overly wide or spacious. This means that a sound reproduction with a larger magnitude of ITD fluctuations would not necessarily be preferred or considered of higher sound quality. The relationship between preference and ITD fluctuations together with other spatial attributes in sound recording requires further investigation.

The process of sound recording is often a compromise of a number of factors, and the same is true here. Whilst following some of the suggestions in this section may create a larger magnitude of ITD fluctuations at the listening position and therefore a greater perceived source width or spaciousness, these techniques may degrade other aspects of the audio quality such as sound source localisation and timbre. Ultimately it is the task of the recording engineer to make decisions about the required subjective result and to attempt to achieve this based on the knowledge of underlying physical cues discussed in this and other similar studies.

CONCLUSION

This paper has covered a number of important aspects of interaural time difference (ITD) fluctuations. Firstly, two measurements were reviewed which aim to quantify the magnitude of ITD fluctuations as an acoustic measure. These were compared and the relative advantages and disadvantages of each were explored. Secondly, the creation of ITD fluctuations by the interaction of a direct sound and one or more reflections was examined. This demonstrated a number of important details including the effect of various types of source signal and certain reflection patterns on the creation of ITD fluctuations.

Thirdly, the results of three experiments that investigated the subjective effect of various ITD fluctuations were summarised. These showed that the main perceptual effect of increasing the magnitude of fluctuations was an increase in the width of either the trajectory of movement, the perceived size of the sound source, or the perceived size or spaciousness of the acoustical environment, depending on the characteristics of the sound.

Finally, the process of applying this research to the various parameters of sound recording and reproduction were discussed, with suggestions given on how to maximise or minimise the magnitude of fluctuations created at the ears of the listener.

The paper also highlighted the large amount of research that still needs to be conducted in order to fully understand the details of ITD fluctuations. Further investigation of the effect of various physical parameters of aspects of the sound recording and reproduction chain on the ITD fluctuations created at the ears of the listener is of particular importance. The task of relating the perception of ITD fluctuations to subjective preference is also significant. The results of this research will be presented in future publications.

ACKNOWLEDGEMENTS

The authors would like to thank David Meares and BBC R&D for sponsorship and provision of equipment, the MEDUSA team for their thought-provoking discussion and guidance, the AES Educational Foundation for financial assistance and Lin Oskam for her patient proof reading.

REFERENCES

[1] Griesinger, D. 1992: 'IALF – Binaural Measures of Spatial Impression and Running Reverberance', **Audio Engineering Society Preprint**, 92nd Convention, preprint no. 3292.

[2] Blauert, J. 1972: 'On the Lag of Lateralization Caused by Interaural Time and Intensity Differences', **Audiology**, vol. 11, pp. 265-270.

[3] Grantham, W. D. and Wightman, F. L. 1978: 'Detectability of varying interaural temporal differences', **Journal of the Acoustical Society of America**, vol. 63, no. 2 (February), pp. 511-523.

[4] Griesinger, D. 1998: 'General Overview of Spatial Impression, Envelopment, Localization, and Externalization', **Proceedings of the 15th International Audio Engineering Society Conference**, Copenhagen, Denmark, pp. 136-149.

[5] Griesinger, D. 1999: 'Objective Measures of Spaciousness and Envelopment', **Proceedings of the 16th International Audio Engineering Society Conference**, Rovaniemi, Finland, pp.27-41.

[6] Mason, R. and Rumsey, F. 2000: 'An Assessment of the Spatial Performance of Virtual Home Theatre Algorithms by Subjective and Objective Methods', **Audio Engineering Society Preprint**, 108th Convention, preprint no. 5137.

[7] Blauert, J. and Cobben, W. 1978: 'Some Consideration of Binaural Cross Correlation Analysis', **Acustica**, vol. 39, pp. 96-104.

[8] Moore, B. C. J. 1997: **An Introduction to the Psychology of Hearing**, 4th edition, (London: Academic Press).

[9] Ifeachor, E. C. and Jervis, B. W. 1993: **Digital Signal Processing: A Practical Approach**, (Wokingham, UK: Addison-Wesley).

[10] Barron, M. and Marshall, A. H. 1981: 'Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure', **Journal of Sound and Vibration**, vol. 77, no. 2, pp.211-232.

[11] Hidaka, T., Beranek, L. L. and Okano, T. 1995: 'Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls', **Journal of the Acoustical Society of Americ**a, vol. 98, no. 2 (August), pp. 988-1007.

[12] Morimoto, Masayuki. 1995: 'The Relation Between Spatial and Cross-Correlation Measures', **Proceedings of the 15 th International Congress on Acoustic**s, Trondheim, Norway, pp. 581-584.

[13] Blauert, J. 1997: **Spatial Hearing**, 3rd edition, (Cambridge, USA: MIT Press).

[14] Mason, R., Rumsey, F. and de Bruyn, B. 2001: 'An investigation of interaural time difference fluctuations, part 1: the subjective spatial effect of fluctuations delivered over headphones', **Audio Engineering Society Preprint**, 110th Convention.

[15] Griesinger, D. 1992: 'Measures of Spatial Impression and Reverberance Based on the Physiology of Human Hearing', **Proceedings of the 11th International Audio Engineering Society Conference**, Portland, Oregon, USA, pp. 114-145.

[16] Howard, D. M. and Angus, J. 1996: **Acoustics and Psychoacoustics**, (Oxford: Focal Press).

[17] Everest, F. A. 1994: **The Master Handbook of Acoustics**, (New York: TAB Books).

[18] Gardner, B. and Martin, K. 1994: 'HRTF Measurements of a KEMAR Dummy-Head Microphone', **http://sound.media.mit.edu/~kdm/hrtf.html**.

[19] Hansen, V. and Munch, G. 1991: 'Making Recordings for Simulation Tests in the Archimedes Project', **Journal of the Audio Engineering Society**, vol. 39, no. 10 (October), pp. 768-774.

[20] Beranek, L. 1996: **Concert and Opera Halls – how they sound**, (Woodbury, NY: Acoustical Society of America).

[21] Griesinger, D. 1997: 'The Psychoacoustics of Apparent Source Width, Spaciousness and Envelopment in Performance Spaces', **Acta Acustica**, vol. 83, no. 4 (July/August), pp. 721-731.

[22] Potter, J. M., Bilsen, F. A. and Raatgever, J. 1995: 'Frequency Dependence of Spaciousness', **Acta Acustica**, vol. 3, pp. 417-427.

[23] Mason, R., Rumsey F. and de Bruyn, B. 2001: 'An investigation of interaural time difference fluctuations, part 3: the subjective effect of fluctuations in continuous stimuli delivered over loudspeakers', submitted to the **111th Audio Engineering Society Convention**, New York, 21st – 24th September.

[24] Mason, R., Rumsey F. and de Bruyn, B. 2001: 'An investigation of interaural time difference fluctuations, part 4: the subjective effect of fluctuations in decaying stimuli delivered over loudspeakers', submitted to the **111th Audio Engineering Society Convention**, New York, 21st – 24th September.

[25] Barron, M. and Marshall, A. H. 1981: 'Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure', **Journal of Sound and Vibration**, vol. 77, no. 2, pp.211-232.

[26] Hidaka, T., Beranek, L. L. and Okano, T. 1995: 'Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls', **Journal of the Acoustical Society of America**, vol. 98, no. 2 (August), pp. 988-1007.

[27] Bregman, A. S. 1990: **Auditory Scene Analysis: the perceptual organization of sound**, (Cambridge, USA: MIT Press).

[28] Griesinger, D. 2000: 'Surround: The Current Technological Situation', presented in Paris, November, available from http://www.world.std.com/~griesngr/.

[29] Theile, G. 2000: 'Multichannel Natural Recording based on Psychoacoustic Principles', **Audio Engineering Society Preprint**, 108th Convention, preprint no. 5156.

[30] Griesinger, D. 1998: 'Multichannel Sound Systems and Their Interaction with the Room', **Proceedings of the 15th International Audio Engineering Society Conference**, Copenhagen, Denmark, pp. 159-173.

[31] Harris, N. and Hawksford, M. O. J. 1999: 'Measurement and simulation results comparing the binaural acoustics of various direct radiators', **Audio Engineering Society Preprint**, 107th Convention, preprint no. 5015.

[32] Zacharov, N. 1998: 'Subjective Appraisal of Loudspeaker Directivity for Multichannel Reproduction', **Journal of the Audio Engineering Society**, vol. 46, no. 4 (April), pp. 288-303.

[33] Olive, S. E. and Toole, F. E. 1989: 'The Detection of Reflections in Typical Rooms', **Journal of the Audio Engineering Society**, vol. 37, no. 7/8 (July / August), pp. 539-553.

[34] Rumsey, F. 1998: 'Subjective Assessment of the Spatial Attributes of Reproduced Sound', **Proceedings of the 15th International Audio Engineering Society Conference**, Copenhagen, Denmark, pp. 122-135.

[35] Toole, F. E. 1985: 'Subjective Measurement of Loudspeaker Sound Quality and Listener Performance', **Journal of the Audio Engineering Society**, vol. 33, no. 1/2 (January / February), pp. 2-32.