Abstract
Auditory width measurements based on the interaural cross-correlation coefficient (IACC) are often used in the field of concert hall acoustics. However, there are a number of problems with such measurements, including large variations around the centre of a room and a limited range of values at low frequencies. This paper explores how some of these problems can be solved by applying the IACC in a more perceptually valid manner and using it as part of a more complete hearing model. It is proposed that measurements based on the IACC may match the perceived width of stimuli more accurately if a source signal is measured rather than an impulse response, and when factors such as frequency and loudness are taken into account. Further developments are considered, including methods to integrate the results calculated in different frequency bands, and the temporal response of spatial perception.

1. Introduction
In the field of auditorium acoustics, measurements based on the interaural cross-correlation coefficient (IACC) have been used for a number of years to predict aspects of spatial impression, including auditory width\(^1\), with varied success. Based on this research, such measurements have also been applied in the field of sound recording and reproduction.

Although a great deal of research has been undertaken into the IACC as a predictor of auditory width, it is apparent that there are still a number of limitations to its application. For example, it has been shown that the results of IACC-based measurements can vary greatly across the width of a concert hall, as shown by de Vries et al [1], without a corresponding change in ASW. Also, it has been shown that IACC measurements of low frequency stimuli rarely differ greatly from a value of 1 [2], even though low frequency stimuli can vary greatly in perceived width [3]. However, it is possible that these limitations are not due to inherent problems with the cross-correlation calculation itself, but that they are caused by the manner in which this is applied.

The most common IACC-based measurement technique that relates to aspects of spatial impression is that derived by Hidaka et al [4], which is included in an appendix of ISO 3382 [5]. This technique is based on measuring segments of an impulse response that has been recorded from one or more source positions to one or more receiver positions within an auditorium. An IACC-based analysis of the impulse response up to 80 ms is used as a predictor of the ASW, and an IACC-based analysis of the impulse response beyond 80 ms is used as a predictor of the perceived envelopment\(^2\). The calculations are made in octave bands, and an average taken across the bands centred on 500, 1000, and 2000 Hz. This choice of frequency range is based on the assumption that the IACC of high frequency stimuli does not affect spatial impression, and to avoid potential problems at low frequencies.

In a series of related research papers, Morimoto and colleagues have proposed alternatives to this measurement method, including the use of more representative musical stimuli [6], the division of source and environment related aspects based on the law of the first wavefront [7], the effect of loudness on ASW [8], and the inclusion of the IACC calculation in a more complete model of binaural hearing [8].

---

\(^{1}\) ‘Auditory width’ may be interpreted to include the more common term in auditorium acoustics: ‘apparent source width’ (ASW).

\(^{2}\) The authors consider that the subjective attribute that relates to the IACC of a reverberant signal can be more accurately described as the ‘environment width’, and that the term ‘envelopment’ used in auditorium acoustics relates to a combination of the spatial impression and relative loudness of the reverberation, as discussed in [9].

---

Figure 1: Block diagram of the main processing stages of the binaural auditory width model
possible, using a similar approach to Morimoto and colleagues. A simplistic diagram of the components of this model is shown in Figure 1. The measurement should ideally be applicable to a wide range of stimulus types, for which results can be directly compared regardless of the stimulus type, and that can be related to a physical parameter, such as a subtended angle. This paper describes the research that has been undertaken so far, and outlines the work still required in order to create a practical measurement tool.

2. Binaural input signal

2.1. Measurement of representative source signals

As mentioned above, IACC-based measurements that relate to the perceived spatial impression of concert hall acoustics are commonly conducted by analysing the properties of recorded impulse responses. It is logical to use this approach, as a large number of factors can be derived from an impulse response. However, whilst this is an established method, it may not necessarily be the most accurate for relating objective measurements to a perceivable effect, as was discussed in [10]. This is due to the dissimilarity between an impulsive signal and the majority of the sounds that are produced in a concert hall during performances.

There are two main differences between impulses and tonal musical signals that affect the measured IACC. Firstly, as an impulse has a wide bandwidth, a measurement based on this is likely to give a different result to a measurement based on a tonal signal. This is because the characteristics of the interaction between the direct sound and the reflections of the acoustical environment are dependent on the precise frequency content of the signal as well as the reflection pattern of the acoustical environment [11]. Secondly, irrespective of the previous case, the short duration of the impulse means that little, if any, interaction between the direct sound and the reflections will occur [12]. This means that it will be difficult to predict the effect of the interactions that will occur with longer musical tones.

These differences between transient and musical signals result in differing IACC measurements in concert halls, as found by Griesinger [13]. He observed that the IACC that was measured by using a musical signal usually resulted in lower values than that of the corresponding impulse response. If it is assumed that it is the IACC of the musical stimuli that causes the perceived spatial impression, then measurements made of impulsive stimuli may be somewhat misleading.

2.2. Division of source and environment components

In order to generate objective measurement results that match the subjective attributes of the source (e.g. ASW) and the environment or reverberation (e.g. envelopment or environment width), a method is required for separating these factors. As mentioned above, this is commonly achieved by dividing the impulse response at 80 ms after the direct sound. From this, the characteristics of the impulse response before 80 ms are related to the perceived attributes of the sound source, and the characteristics of the impulse response after 80 ms are related to the perceived attributes of the reverberant environment.

The use of the specific value of 80 ms after the direct sound for the division between the source-related and environment-related segments of the impulse response appears to have been based on this being the threshold of the perception of disturbing echoes with musical stimuli. However, the echo threshold is dependent on a wide range of factors including the source signal, the level of the reflection, and the pattern of the preceding reflections, as discussed in [10]. In view of this, it seems somewhat arbitrary to attempt to apply results based on a single time constant to a wide range of source signals in a wide range of acoustical environments. In addition to this, research has indicated that the two time segments do not affect the perceived source width and envelopment independently [14]. Therefore it seems that a single time-based division of an impulse response at 80 ms after the arrival of the direct sound may not accurately separate the source-related and environment-related aspects of a sound.

An alternative basis for dividing the source and environment related aspects of a signal is perceptual grouping, as suggested by Griesinger [13]. This involves the division of a musical signal into segments that contain physical cues that are perceived to be an attribute of the source, and segments that contain physical cues that are perceived to be an attribute of the acoustical environment. This is a complex area of psychoacoustics which requires further research; however a simplistic approximation can be made by dividing the signal according to whether or not a direct sound is arriving at the receiver. In this case, it is expected that all components of the signal that arrive whilst the direct sound is reaching the receiver will be perceived to be an attribute of the sound source. Conversely, it is expected that all components of the signal that arrive at the receiver after the direct sound has ended will be perceived to be an attribute of the reverberation or the acoustical environment.

Evidence that supports this approach was reported in [15, 16], where it was found that the physical parameters contained in constant sounds affected the perceived width of the source, and the physical parameters contained in decaying sounds affected the perceived width of the reverberation. Further research is planned to confirm these results. The simplistic separation of source and environment components described above should provide reasonable results for simple single source signals (e.g. a single note from a single instrument) in most reverberant environments, however this may require further modification for more complex source signals.

3. Perception of variations in IACC over time

It has been shown in previous research that the IACC of natural signals can vary a great deal over time [17]. It is also known that variations in IACC over time can be perceived; though as with most perceptual processes, this perception has a limited temporal resolution [18]. Therefore, in order to accurately measure the width of a signal with a time-varying IACC, these variations need to be quantified. To do this, the input signal can be divided into a number of time windows, with IACC measurements made of each of these. A number of studies have been conducted that investigated the maximum rate of variations in IACC that can be perceived, and it appears that a suitable window length to use in IACC
measurements is between approximately 35 to 243 ms, where the results are dependent on both the subject and the experimental task [19].

Based on this research, the measurement model currently uses a rectangular integration window of between 35 and 80 ms, depending on the intended application of the results (i.e. whether the results should reflect the judgements of the most critical or an average listener). The integration window is moved through the input signal in steps of approximately 1.5 ms to allow for the possibility that rapid variations in IACC are perceivable. It must be noted that further work is needed to verify the most appropriate window shape, duration and step size to accurately reflect the perceived variations in width over time.

4. Frequency dependence of perceived width

A number of previous studies have indicated that the relationship between a narrow-band stimulus with a given IACC and its perceived width is dependent on its frequency [20]. However, until recently the causes of this effect and methods to compensate in a measurement had not been derived. The authors undertook a number of experiments to investigate certain related factors, including the perceived width of narrow-band stimuli with an IACC of 1, the effect of the breakdown of phase locking in the ear, and the effect of the limited range of IACC values encountered at low frequencies, as mentioned earlier [21].

The main findings of these experiments were as follows. Firstly, it was found that the IACC affects the perceived width of narrow-band stimuli with centre frequencies up to 15 kHz, not solely below a few kilohertz as has been asserted previously. However, it is necessary to simulate the breakdown of phase-locking in the ear, as the perceived width of high frequency stimuli is dependent on the IACC of the signal envelope, not the IACC of the fine temporal detail of the signal. Secondly, it was found that low frequency stimuli are perceived to be wider than stimuli at mid or high frequencies having the same IACC, most likely due to the auditory system compensating for the limited range of IACC values created in natural sound fields at low frequencies. Thirdly, it was found that even for stimuli that were identical at both ears, the width was dependent on frequency, possibly due to inaccuracies in the neural phase locking at each end of the audio spectrum. The experiments undertaken and the methods that were derived to take into account the frequency dependency are described in detail in [21].

5. Loudness dependence of perceived width

Previous studies have also indicated that the loudness or level of a sound affects the perceived width [8]. However, as these studies used wideband stimuli, it is possible that increasing the reproduction level caused the low frequency components to be more clearly audible due to the frequency response of the human ear. In order to determine whether this affected previous results, the authors undertook an experiment to investigate the effect of loudness on perceived width using similar narrow-band stimuli to those used in the experiments that investigated the frequency dependency of perceived width [22].

The experiment included stimuli with loudness levels of 50 to 70 phons, with a range of centre frequencies to investigate whether the effect is dependent on frequency. The results indicated that the effect of loudness on perceived width is not frequency dependent, and that it can be modelled by multiplying the measured IACC by a single loudness dependent linear function. Validation experiments have indicated that this gives a reasonable match between the objective and subjective results across a reasonably wide range of loudness values. The experiments undertaken and the methods that were derived to take into account the frequency dependency are described in detail in [22].

6. Further developments

The research conducted so far has resulted in a measurement model that can quantify the relative width of narrow-band stimuli with a wide range of centre frequencies and a wide range of loudness values. However, there are still a number of important factors that need to be determined in order to develop the model into a practical measurement tool. Some of these are summarised below.

As mentioned above, further research is required to determine the way in which temporal variations in IACC affect the perceived width of a sound, and the optimum method to model this. Factors that need to be determined are: the shape and duration of the integration window, the size of step in time between each window, and whether any additional temporal smoothing is required.

The results of the measurement are currently expressed as a scale that allows comparison between stimuli, but is not related to any physical measurement of width. Further research is required to allow prediction of the perceived width of stimuli in terms of a physical distance across for sounds perceived to be within the head, or in terms of a subtended angle for sounds perceived to be outside the head.

In order to measure the width of stimuli with a wide bandwidth and complex spectral content, the results of the calculations made in individual frequency bands need to be combined. The results of Ueda and Morimoto [23] indicated that differing signal levels in different frequency bands affected the perceived overall width of a stimulus. It is possible that this can be predicted from data on simultaneous across-frequency masking; however this needs to be investigated further. In addition, it is possible that the perceived location of the signal in each frequency band will be different (i.e. one frequency band is perceived to the left and another band is perceived to be central). This will need to be taken into account in the measurement in order to derive an accurate measurement of the total width of the stimulus.

Finally, once the measurement model is complete, the most practical and efficient method for applying it to concert hall acoustics needs to be determined. This will require the selection of a set of representative sound source signals that cover as many signal types as possible, without consuming an impractical amount of time.
7. Conclusions
This paper has outlined a number of ways in which the extant IACC-based auditory width measurements may be modified. This includes:

- measurement of representative stimuli instead of impulse responses
- division of the source and environment related aspects using perceptual grouping as opposed to a single time after the arrival of the direct sound
- measurement of the variation in width over time by calculating the IACC in a number of overlapping time windows
- measurement of the IACC over the whole audio bandwidth as opposed to three mid-frequency octave bands
- inclusion of a simulation of the breakdown of auditory neural phase locking
- inclusion of compensation for the dependence of width on frequency and loudness
- production of the measured results in terms of a physically relevant scale
- integration of results calculated in individual frequency bands into a single measurement that takes into account masking, across frequency interaction and differing perceived lateral locations.

Validation experiments for those factors that have been investigated have indicated that the modifications improve the match between the measured result and the subjective effect. A comparison of the complete measurement model with the subjective effect of fluctuations in continuous stimuli delivered over loudspeakers, "Audio Eng. Soc. Preprint, no. 5457, 2001.

8. Acknowledgements
This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), UK – grant GR/R55528/01.

9. References