Audio Engineering Society

Convention Paper

Presented at the 128th Convention
2010 May 22–25 London, UK

Design of a Circular Microphone Array for Panoramic Audio Recording and Reproduction: Microphone Directivity

Hüseyin Hacıhabiboğlu, Enzo de Sena, and Zoran Cvetković

Centre for Digital Signal Processing Research (CDSPR), Division of Engineering, King's College London, Strand, London, WC2R 2LS, United Kingdom

Correspondence should be addressed to Hüseyin Hacıhabiboğlu (huseyin.hacihabiboglu@kcl.ac.uk)

ABSTRACT

Design of a circularly symmetric multichannel recording and reproduction system is discussed in this paper. The system consists of an array of directional microphones evenly distributed on a circle and a matching array of loudspeakers. The relation between the microphone directivity and the radius of the circular array is established within the context of time-intensity stereophony. The microphone directivity design is identified as a constrained linear least squares optimisation problem. Results of a preliminary informal subjective evaluation are presented which indicate the usefulness of the proposed microphone array design technique.

1. INTRODUCTION

We are surrounded by sound. Sound sources can be situated at any direction on the horizontal plane. A good surround sound system should therefore reproduce sources situated at different directions equally accurately. Commercially available multichannel systems usually employ uneven loudspeaker positions [1] favouring the front direction, and the audio material to be played back over such systems is typically engineered heavily at the post-processing stages so as to provide a good localization and ambience perception. While satisfactory listener experience can be achieved most of the time, the perceptual consistency of the reproduced audio with the actual recording environment cannot be guaranteed and the reproduced sound field reflects the choices of the audio engineer rather than the properties of the actual recording venue.

There exist different audio reproduction systems based on the concept of reconstructing the sound field exactly. Ambisonics [2, 3], and wave-field synthesis (WFS) [4] are two such systems. The former achieves perfect reconstruction only at a narrow listening area. The latter requires significant computational resources and a high number of channels and is thus not feasible in a domestic setting. A multichannel recording and reproduction system was proposed by Johnston and Lam [5] that overcomes these limitations in order to provide a panoramic listening experience to the listener in a wider listening area. The recording system consists of a sparse circular array of five near-coincident hypercardioid microphones evenly distributed on the horizontal plane. The original array radius was selected as 15.5 cm and it was suggested that this radius allowed capturing spatial hearing cues on the horizontal plane[1]. In addition, two superdirectional microphones were positioned at the apexes of the open sphere defined by the centre and radius of the microphones on the horizontal plane. The horizontal reproduction setup consisted of five loudspeakers evenly distributed on a circle.

A pilot study by Hall and Cvetković [7] indicated that better subjective results could be obtained if the microphone array radius is larger using the original microphone directivity proposed by Johnston and Lam. Hacihabiboglu and Cvetković [8] suggested that Johnston-Lam array could be considered as a conjoined set of near-coincident stereophonic pairs and can also be generalized to more than five channels. A mathematical analysis of the system's response to monochromatic plane waves revealed that better performance and a wider listening area could be achieved by minimizing the crosstalk between non-adjacent microphones. A microphone directivity design method based on stereophonic intensity panning laws was also proposed.

This paper is concerned with the microphone array directivity design problem from the perspective of time-intensity stereophony. It is well known that both the time delay and the level differences between channels play a role in subjective localization [9]. The circular array of directional microphones imposes both level differences (due to microphone directivity) and time delays (due to non-zero array radius) on the recorded sound source. In turn, time delays between each microphone are related to both the number of channels and the array radius. Therefore, microphone directivity design problem cannot be separated from the selection of array radius.

[1] See also our companion paper [6]

This paper is concerned with the design of microphone directivity pattern for a given microphone array radius which results in a good subjective localisation performance. The circular microphone and loudspeaker arrays as originally proposed by Johnston and Lam are briefly reviewed in Sec. 2. An analytical formulation of the multichannel system based on active intensity will be presented in Sec. 3. The cross-talk terms will be identified and the maximum allowable inter-channel crosstalk level will be also be discussed. Franssen's time-intensity panning curves will be briefly reviewed and their relevance in the context of the proposed multichannel system will be discussed. Microphone directivity design problem will then be identified as a constrained linear least squares optimization problem in Sec. 4. Several design examples will be provided. Results of a listening experiment which compares the subjective localisation performance of different microphone directivity patterns under free-field conditions will be reported in Sec. 5. The results will be discussed and conclusions drawn in Sec. 6.

2. JOHNSTON-LAM ARRAY

Johnston and Lam proposed [5] a circularly symmetric microphone array composed of five first-order microphones on the horizontal plane facing outwards and two superdirectional microphones facing up and down. The stated aim of Johnston-Lam array was to accurately capture interaural cues of binaural hearing. The recorded audio could be played back with a corresponding loudspeaker array consisting of five equispaced loudspeakers on a circle to provide panoramic audio to the listeners. The signals recorded using up and down facing microphones were mixed to signals obtained with the horizontal microphones. The system was reported to provide very realistic spatial perception. In a later patent [10], the setup was generalised to having odd number of microphones on the horizontal plane. It was also suggested in the patent that the vertical microphones can be omitted from the system without much subjective degeneration in the reproduced sound field.

In the original proposal [5] and also in the subsequent patent by Johnston and Wagner [10] the directivity pattern of the individual array elements

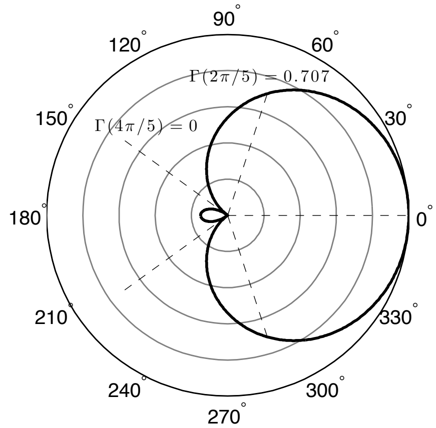

Fig. 1: The directivity function proposed by Johnston and Wagner [10].

were selected so as to have a gain of 3 dB below the front direction gain at the look direction of the neighbouring microphone and a zero at the next non-consecutive channel. For the original proposal which considered five channels on the horizontal plane this requirement corresponded to having a 3 dB decrease at 72° and a zero at 144°. The second-order microphone directivity which satisfies these design considerations is given in Fig. 1.

A first-order microphone directivity which roughly approximated this condition was evaluated by Hall and Cvetković [7] and it was observed that the subjective performance of the recording/playback system increased with increasing array radius. In a recent study by Hacıhabiboğlu and Cvetković [8], it was suggested that Johnston-Lam array can be considered as a conjoined set of near-coincident stereophonic pairs, and thus microphone directivity can be indirectly related to earlier work on stereophonic recording.

In the following section, the analysis presented in that study will be briefly reviewed and the effects of using conjoined near-coincident stereophonic microphone pairs on directional reproduction of audio are outlined. As we are only concerned with recording and reproduction of spatial audio on the horizontal plane, only the horizontal components of the circular microphone array are considered in this paper.

3. ANALYTIC EVALUATION OF CIRCULARLY SYMMETRIC ARRAYS

A stationary sound field can be represented as a sum of monochromatic plane waves with different amplitudes, frequencies, and propagation directions. An objective analysis of the directional reproduction capabilities of multichannel audio systems is thus possible by analysing the response of the system for a single monochromatic plane wave [11]. The analysis presented in this section follows this approach.

The microphone array studied in this paper consists of an array of N directional microphones with the same directivity function, $\Gamma(f, \theta)$, positioned on a circle of radius r_m at equal angular intervals with their acoustical axes pointing out (see Fig. 2). Directivity functions of real microphones are functions of both the angle of incidence and of frequency. However, we assume in this paper that the 'ideal' microphones in the hypothetical array are frequency independent (i.e. $\Gamma(f, \theta) = \Gamma(\theta)$ for all frequencies).

Let us consider a complex monochromatic plane wave of frequency f_0, incident from the horizontal direction, θ_s. The signal recorded by the m^{th} microphone is:

$$P_m(t) = A\Gamma_m(\theta_s)\, e^{jk_0\left[ct - r_a \cos\left(\theta_s - \frac{2\pi m}{N}\right)\right]}, \quad (1)$$

where A is the peak amplitude, $\Gamma_m(\theta_s) = \Gamma(\theta_s - 2\pi m/N)$ is the sensitivity (i.e. directivity) of the microphone, $k_0 = 2\pi f_0/c$ is the wave number, and c is the sound speed.

The reproduction setup consists of N angularly equispaced loudspeakers on a circle (See Fig. 3). Each loudspeaker plays back the audio signal recorded by the microphone with the corresponding angle without any additional processing. Let us assume that the loudspeakers are positioned in the acoustic far-field and thus effectively behave as plane-wave sources. We can express the pressure component of the sound field at an arbitrary listening position, $\mathbf{x}_e = r_e [\cos\psi_e \ \sin\psi_e]$, within the listening area due to the loudspeaker m as:

$$p_m(\mathbf{x}_e) = A\Gamma_m(\theta_s)\, e^{jk_0\left[ct - r_a \cos\left(\theta_s - \frac{2\pi k}{N}\right) - r_e \cos\left(\frac{2\pi k}{N} - \psi_e\right)\right]}. \quad (2)$$

Here, $r_e = |\mathbf{x}_e|$ is the radial distance from the centre of the circle defining the loudspeaker array, and

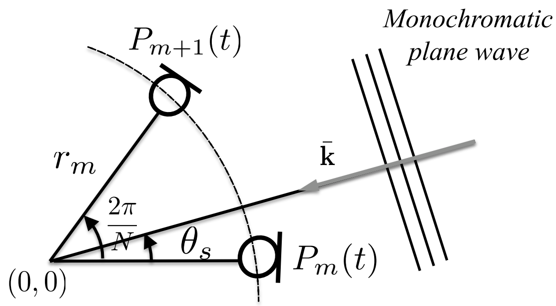

Fig. 2: Two elements of the N-channel microphone array. Adapted from [8].

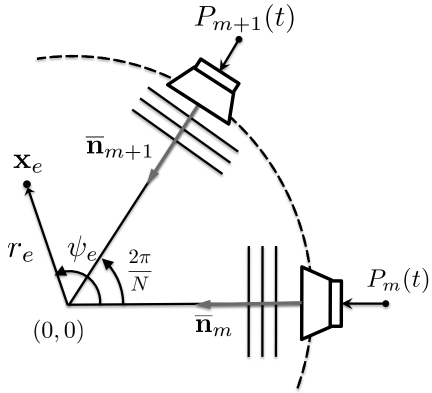

Fig. 3: Two elements of the N-channel loudspeaker array. Adapted from [8].

ψ_e denotes the angular positioning of the listening position.

The complex pressure and velocity components of the acoustical field in the listening area due to the loudspeaker array will be a sum of individual components due to these N loudspeakers:

$$p(\mathbf{x}_e) = \sum_{m=1}^{N} p_m(\mathbf{x}_e), \tag{3}$$

$$\overline{\mathbf{v}}(\mathbf{x}_e) = \frac{1}{\rho c} \sum_{k=1}^{N} p_k(\mathbf{r}_l)\overline{\mathbf{n}}_m. \tag{4}$$

where $\overline{\mathbf{n}}_m$ is the unit vector co-directional with the acoustic axis of the loudspeaker m.

The product of pressure and (complex conjugate) velocity components is known as the complex intensity. Complex intensity is not time-dependent for a complex monochromatic plane wave as opposed to instantaneous intensity. The complex intensity, $\mathbf{I}_c(\mathbf{x}_e)$, can be expressed using the pressure and velocity components as:

$$\begin{aligned}\mathbf{I}_c(\mathbf{x}_e) &= \frac{1}{2} p(\mathbf{x}_e)\overline{\mathbf{v}}^*(\mathbf{x}_e), \\ &= \frac{1}{2\rho c}\sum_{k=1}^{N}\sum_{m=1}^{N} p_k(\mathbf{x}_e)p_m^*(\mathbf{x}_e)\overline{\mathbf{n}}_m\end{aligned} \tag{5}$$

The summand can be expressed as

$$\mathbf{I}_{c,km}(\mathbf{x}_e) = A^2\gamma_{km}(\theta)e^{j2k_0 d_{km}\sin\xi_{km}}\overline{\mathbf{n}}_m \tag{6}$$

where $\gamma_{km}(\theta) = \Gamma_m(\theta)\Gamma_k(\theta)$ and,

$$d_{km} = \sin\left[\frac{(k-m)\pi}{N}\right]\sqrt{r_a^2 + r_e^2 + 2r_a r_e\cos(\psi_e - \theta_s)},$$

$$\xi_{km} = \theta - \frac{(k+m)\pi}{N} + \tan^{-1}\left[\frac{r_e\sin(\psi_e-\theta_s)}{r_a + r_e\cos(\psi_e - \theta_s)}\right].$$

The real part of complex intensity, also known as *active intensity* [12], can be used to investigate the directional properties of the reproduced sound field. Active intensity is co-directional with the propagation direction of a plane wave at a given location. The active intensity due to the combination of recording and reproduction systems is then given by:

$$\mathbf{I}_{a,km}(\mathbf{x}_e) = A^2\gamma_{km}(\theta_s)\cos\left(2k_0 d_{km}\sin\xi_{km}\right)\overline{\mathbf{n}}_m. \tag{7}$$

and, the total active intensity is:

$$\mathbf{I}_a(\mathbf{x}_e) = \sum_{k=1}^{N}\sum_{m=1}^{N}\mathbf{I}_{a,km}(\mathbf{x}_e) \tag{8}$$

It may be observed that the active intensity is related not only to the active intensities of individual loudspeakers, $\mathbf{I}_{a,mm}(\mathbf{x}_e)$, but also the *cross-talk* terms $\mathbf{I}_{a,km}(\mathbf{x}_e)$, $m \neq k$, occurring due to their interaction.

Correct reconstruction of the plane wave requires the reproduced active intensity $\mathbf{I}_a(\mathbf{x}_e)$ to be codirectional with the direction of wave propagation. The magnitude of active intensity determines the strength of the directional property of the reproduced sound field. Therefore, in order to reproduce the plane wave correctly active intensity should have a large magnitude and also be in the same direction as the propagation direction of the recorded plane wave.

4. MICROPHONE DIRECTIVITY DESIGN

4.1. Cross-talk Components

At least two loudspeakers are required to reproduce the direction of a plane wave correctly around the optimal listening area also know as the 'sweet spot' (i.e. $\mathbf{x}_e = 0$). Therefore, the aim of the proposed multichannel system is to have only two loudspeakers active for a *single plane wave*. For example, if the plane wave is incident from the direction, θ_s, such that $\frac{2\pi m}{N} \leq \theta \leq \frac{2\pi(m+1)}{N}$, only the loudspeakers m and $m+1$ should be active. This allows using stereophonic panning laws for designing the common microphone directivity pattern. In order to achieve this, the cross-terms, $\gamma_{km}(\theta)$, for non-consecutive microphones, m and k, should be minimised. This requires designing superdirectional pressure-gradient microphones [13] of the form:

$$\Gamma(\theta) = \sum_{i=0}^{M} a_k \cos^i(\theta), \tag{9}$$

for which

(i) $\Gamma(0) = 1$, and

(ii) $\Gamma(2m\pi/N) = 0$ for $m = 2...N-2$.

In order to satisfy the second condition, each additional zero in the directivity function will correspond to an increase in the order of directivity by one. Although there exists no comprehensive study of the audibility thresholds of reflections incident from behind the listener, cross-talk may be considered to be effectively zero if its level is at least 15 dB below the front direction sensitivity of the microphone. If this condition is satisfied, only two loudspeakers will be effectively active for any given source direction. In other words, the levels of the remaining loudspeakers will be too low to be audible.

4.2. Time-Intensity Panning

There exist a great body of literature investigating different stereophonic recording techniques. The virtues and vices of coincident, near-coincident, and noncoincident stereophonic recording have been studied thoroughly [14]. The specific microphone array which is the topic of this study behaves like conjoined near-coincident stereophonic pairs if the cross-talk terms are eliminated. In other words both time and intensity differences will be present at each recorded channel.

Time delay between two channels causes the precedence effect [9, 15] to influence the perceived direction of the sound source. Unless, summing localisation threshold [15] is exceeded in the time delay between the channels, time delay will be an important contributing factor in the formation of the perceived direction of the auditory event. From audio engineering perspective, a practical (if slightly heuristic) approach to mapping time and intensity differences to the perceived direction of the auditory image was given by Franssen [16].

Fig. 4 shows the stereophonic time-intensity panning curves adapted from Franssen [16]. The curves represent the level difference between right and left channels. The upper curve, $R(\tau)$, represents the limit at which the auditory image is perceived to be located at the right loudspeaker for a given interchannel delay of τ. The lower curve, $L(\tau)$, represents the limit at which the auditory image is perceived at the left loudspeaker. Operating lines are defined in order to pan the stereophonic image between two loudspeakers with a given maximum interchannel delay. In the figure, two such operating lines are shown. The lines from A_R to A_L, and B_R to B_L are the operating curves for time-intensity panning for a maximum interchannel delay of 1 and 2 ms, respectively.

Let us consider levels of left and right channels (g_L and g_R) of a stereophonic setup. The time-intensity curves given in the figure represent the ratio:

$$\rho(\tau) = 10 \log \left[g_R(\tau)/g_L(\tau) \right] \tag{10}$$

where $\tau = \tau_r - \tau_l$ represents the interchannel delay. If $\rho(\tau) \geq R(\tau)$, the auditory image is perceived at the right loudspeaker. If $\rho(\tau) \leq L(\tau)$, the auditory image is perceived at the left loudspeaker. The operating lines then represent the interchannel delay and loudspeaker level ratios that will cause the auditory image to be panned between the loudspeakers. Additionally, total sound power should be constant:

$$|g_R(\tau)|^2 + |g_L(\tau)|^2 = 1. \tag{11}$$

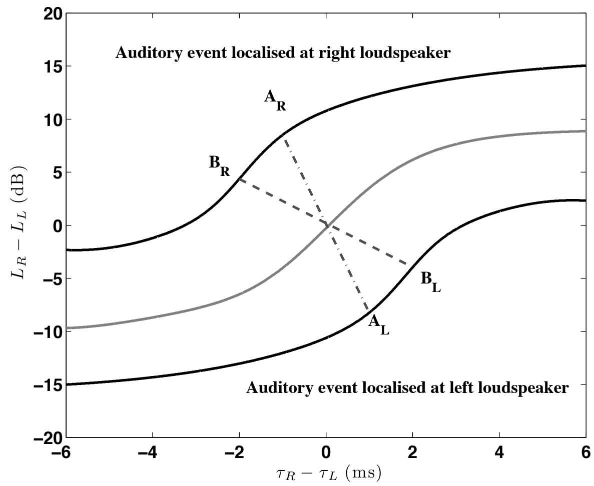

Fig. 4: Time-intensity panning curves, adapted from Franssen [16].

This way, total sound level at the listening position will be constant independent of the direction of the sound source.

Let us also define a maximum delay $\tau_{\max}$ between two channels. The operating line has a slope of:

$$\kappa_o = \frac{L(-\tau_{\max}) - R(-\tau_{\max})}{2\tau_{\max}} \qquad (12)$$

The gain of left (or right) channel can therefore be obtained simply as:

$$g(\tau) = \sqrt{\frac{K(\tau)^2}{K(\tau)^2 - 1}} \qquad (13)$$

where $K(\tau) = 10^{\kappa_o \tau / 10}$.

4.3. Microphone Directivity Design as an Constrained Optimisation Problem

Three conditions are taken into account while designing the directivity function using time-intensity curves:

1. The designed directivity function when paired with the consecutive microphone channels of the recording array should emulate time-intensity panning for angles of incidence between two consecutive channels,
2. The directivity function, $\Gamma(\theta)$, should be at least 15 dB below its value for frontal direction for $\theta > 2\pi/N$ and $\theta < -2\pi/N$, and
3. The directivity function should be effectively zero for non-consecutive channels.

Let us consider a sound source in the acoustical far-field incident from a direction, $2m\pi/N \leq \theta_s \leq 2(m+1)\pi/N$, between two consecutive channels of the circular microphone array. Let us also assume that the cross-talk terms are zero, so the source is effectively recorded by two microphones, m and $m+1$ only. The interchannel delay between these two channels depends on the direction of the source, θ_s, (see Fig. 2) and can be calculated as:

$$\tau(\theta_s) = -2\frac{r_m}{c}\sin\left(\frac{\pi}{N}\right)\sin\left(\theta_s + \frac{\pi}{N}\right) \qquad (14)$$

The maximum delay between channels (when signals at both channels are non-zero) is:

$$\tau_{\max} = -2\frac{r_m}{c}\sin^2\left(\frac{\pi}{N}\right). \qquad (15)$$

A time-intensity panning operating line can be obtained from this maximum interchannel delay value from (12). This operating line can be used to obtain the corresponding gain which essentially is the sensitivity of the microphone for the given source direction.

The conditions stated above can be imposed analytically as a constrained linear least-squares optimisation problem:

$$\min_{\mathbf{a}} \|\mathbf{G}_m\mathbf{a} - \psi\|_2^2 \quad \text{such that} \begin{cases} \mathbf{G}_t\mathbf{a} & \leq & \beta \\ \mathbf{G}_z\mathbf{a} & = & 0 \end{cases} \qquad (16)$$

where

$$\begin{aligned} \mathbf{G}_m &= [\cos^p \theta_{m,q}] \quad q = 0...Q_m \quad p = 0...M, \\ \mathbf{G}_t &= [\cos^p \theta_{t,q}] \quad q = 0...Q_t \quad p = 0...M, \\ \mathbf{G}_z &= [\cos^p \theta_{z,q}] \quad q = 0...Q_z \quad p = 0...M, \\ \mathbf{a} &= [a_0\ a_1\ ...\ a_M]^T, \\ \psi &= [g(\tau(\theta_{m,0}))\ ...\ g(\tau(\theta_{m,Q_m}))]^T, \end{aligned}$$

β is the maximum allowable crosstalk level between non-consecutive channels, $0 \leq \theta_{m,q} \leq 2\pi/N$, $2\pi/N \leq \theta_{t,q} \leq \pi$, and $\theta_{z,q} = 2\pi i/N$. Here, $\theta_{m,q}$

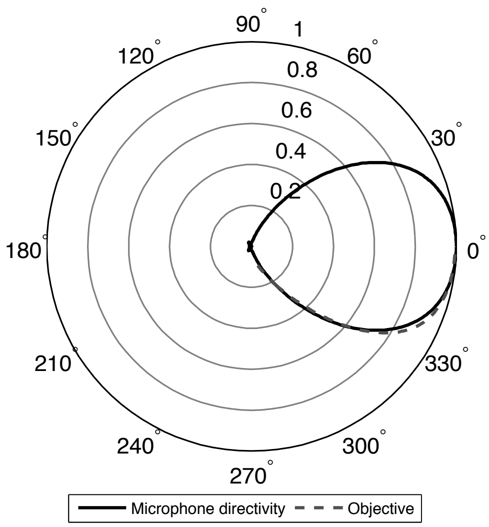

Fig. 5: New directivity function based on time-intensity panning for $N = 5$, $M = 6$, and $r_m = 15.5$ cm.

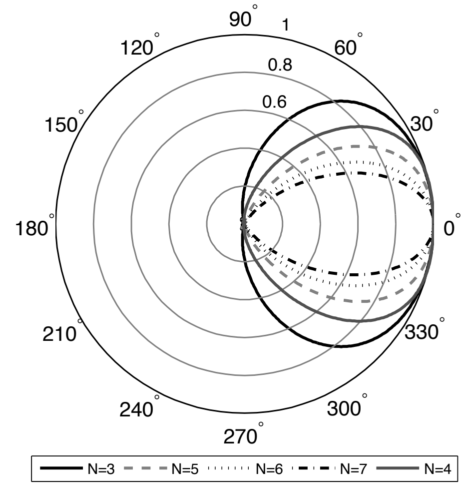

Fig. 6: The new directivity function based on time-intensity panning.

are the angles at which the difference between the directivity function and time-intensity panning gain is minimised, $\theta_{t,i}$ are the angles at which the cross-talk constraint is applied, and $\theta_{z,q}$ are the angles at which the directivity function is zero.

Fig. 5 shows directivity pattern of the microphone designed for $r_m = 15$ cc=m, $M = 6$, and $N = 5$. The design objective due to time-intensity panning law is also overlaid on the directivity plot. It may be observed that a very good approximation to the design criteria can be obtained with a sixth-order design.

Fig. 6 shows the proposed directivities for different number of channels, $N = 3$ to $N = 7$ for $r_m = 15.5$ cm for $M = 6$. It may be observed that as the number of channels increases, a narrower beamwidth is required.

5. SUBJECTIVE EVALUATION

5.1. Subjects

Six subjects (5 male; 1 female), with no reported hearing impairment participated in the experiment. Three of the subjects were the authors of this paper. The subjects reported no difficulties in rating the stimuli during the test.

5.2. Test setup

The subjects were seated in an acoustically isolated sound booth ($T_{60} \approx 200$ ms). The walls and the ceiling of the booth are almost completely absorbent and the only major reflection is from the floor.

The test setup consisted of five MACKIE HR824 active monitor loudspeakers positioned regularly on a circle with a radius of 2 m. Eight Genelec 6010 loudspeakers were positioned on the same circle at equal intervals with $8°$ separation between two consecutive channels of the five-channel system. Both MACKIE and Genelec speakers were calibrated to a nominal level of 78 ± 0.25 dBA using pink noise as measured at the centre of the circle [17]. The MACKIE system was used to play back the simulated five-channel system, while Genelec loudspeakers were used as acoustic pointers. The subjects were seated at the centre of the circle defining the positions of the multichannel system. All the loudspeakers were positioned at the ear level facing the subject. A computer monitor positioned in front of the listener was connected to the audio workstation over which the test routine was executed.

The listeners registered their responses by clicking buttons on a graphical user interface displayed on

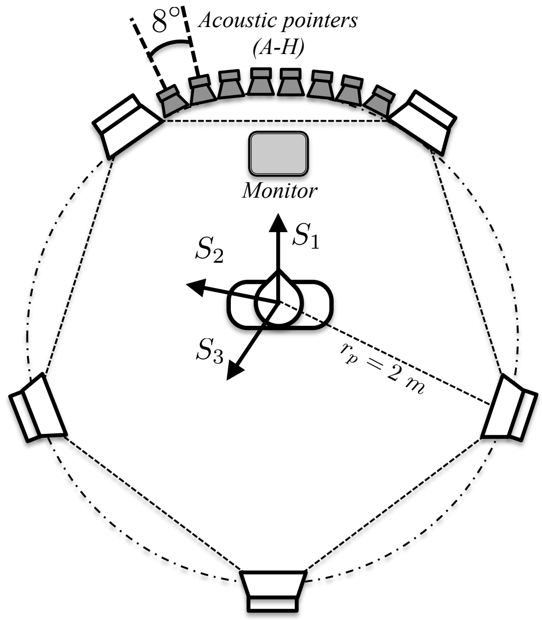

Fig. 7: The test setup. The white loudspeakers constitute the five channel reproduction system. The gray loudspeakers are the acoustic pointers. Three listener seating directions, S_1, S_2, and S_3 are denoted as arrows.

the monitor (see Sec. 5.3 for details). Fig. 7 shows the test setup.

5.3. Methodology and stimuli

The listening test aimed to compare the localisation performance of the original microphone directivity proposed by Johnston and Wagner [10] with the directivity pattern proposed in this paper and the *tanpan* directivity proposed earlier [8]. Three sets of stimuli were synthesised. First set used the directivity function proposed by Johston and Wagner [10], the second set used the *tanpan* directivity, and the third set used the *TI pan* directivity function proposed in this paper. The array radius was set to 15.5 cm for both cases as originally suggested in Johnston and Wagner's patent [10]. The first microphone directivity used a second order design with the coefficients $a_0 = 0.5405$, $a_2 = 0.5748$, and $a_2 = -0.1153$. The second directivity function was of a fifth-order microphone designed according to the tangent panning law. The coefficients were $a_0 = -.0402$, $a_1 = -.0697$, $a_2 = .6771$, $a_3 = 1.2247$, $a_4 = -.1314$, and $a_5 = -.6622$. The third directivity function was of a sixth-order microphone designed using the method proposed in this paper for $r_m = 15.5$ cm, and $N = 5$. The coefficients were $a_0 = 0.0413$, $a_1 = 0.3993$, $a_2 = 0.9683$, $a_3 = -0.0084$, $a_4 = -1.1963$, $a_5 = 0.0889$, and $a_6 = 0.6666$.

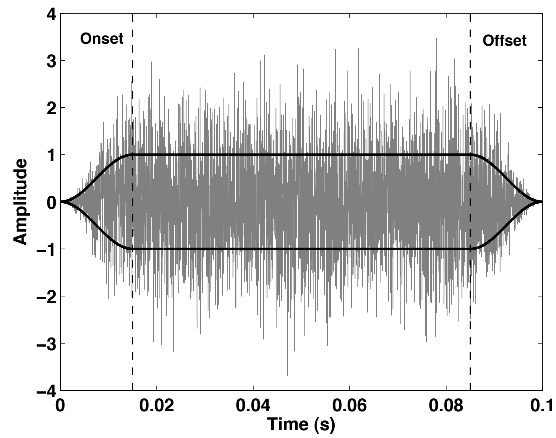

Fig. 8: An example stimulus used in the test.

For all three directivities, the gains and delays at each microphone were calculated due to a simulated source in acoustical free field positioned 10 m away from the centre of the microphone array for eight different directions corresponding to the directions of the acoustic pointers. In order to obtain more accurate stimuli, the delays to each microphone were simulated using allpass fractional delay filters [18].

Windowed white Gaussian noise of 0.1 s duration ($F_s = 44.1$ kHz) was used as a stimulus. The stimulus was resynthesised randomly at each trial in order to eliminate bias due to fixed stimulus spectrum. A cosine tapered window with a 30% taper ratio was used to obtain a relatively smooth stimulus onset and offset in order to reduce transient response of the loudspeakers. Fig. 8 shows an example stimulus used in the test. The dark lines represent the cosine tapered window. The onset and offset portions of the stimulus are also denoted.

The subjects' task was to listen to the simulated free-field recording over the five-channel system and

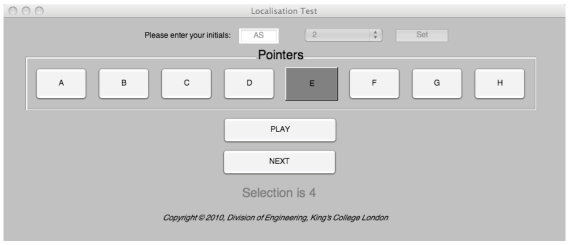

Fig. 9: The GUI used in the test to collect responses.

respond by listening to and selecting the acoustic pointer which is closest to the perceived direction of the auditory image. The responses were elicited using the MATLAB graphical user interface (GUI) shown in Fig. 9.

Three different listener seating directions, S_1, S_2, and S_3 were used in order to test localisation performances at different source directions (see Fig. 7). Each subject took the test for all different directions in a total of three test blocks. At each seating direction, each directivity-direction pair was repeated 15 times and the presentation order was fully randomised. Therefore, 120 responses each for Johnston-Lam microphone directivity, *tanpan* directivity, and the *TI pan* directivity were obtained. Each test block took around 40 mins to complete for each subject. Breaks were given between each block to prevent fatigue.

5.4. Results

For the purpose of quantifying the localisation performance of different systems under test, we define the localisation error as the difference between the simulated angle and the angle corresponding to the acoustic pointer response given by the subjects.

The mean localisation errors and standard deviations for the tested directivities are given in Table 1. It may be observed from these statistics that both tanpan and TI pan directivities perform better than the Johnston/Lam directivity under the given experimental conditions. These statistics may be misleading however as the performance of each directivity changes significantly with the direction of seating.

Figs. 10-12 shows the localisation errors for different angles and 95% confidence intervals. From Fig. 10, it may be observed that the performance of *TI pan* directivity is better than both *Johnston/Lam* and *tanpan* directivitie for front listening direction. Please also note the symmetry of the results which is due to the symmetry of the hearing system. Figs. 11 and 12 shows that there are angular intervals (between 44° and 68° for S_2 and between 116° and 140° for S_3) where the performance of all systems are equally bad. This is possibly due to cone of confusion which reduces the localisation accuracy and not due to the tested systems.

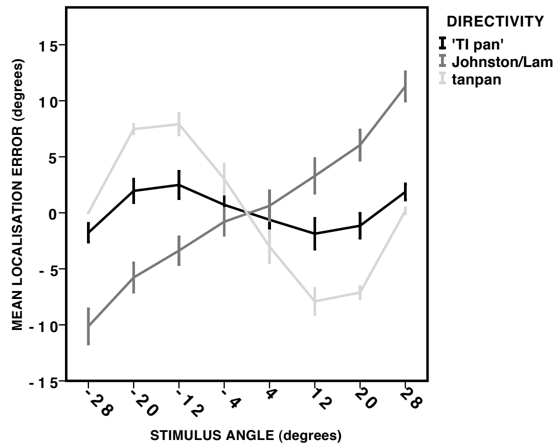

Fig. 10: Mean localization errors for the first listening position averaged across all subjects, S_1 (front). The error bars show the 95% confidence intervals of the mean errors.

An important degradation in multichannel audio systems is due to the widening of the auditory image which reduces the locatedness [9] of the auditory image. We hypothesise that the spread of localisation error reflects the level of difficulty at which the

Table 1: Means and standard deviations of the tested directivities.

Directivity	Mean error	Std. deviation
Johnston/Lam	6.64°	13.74°
Tanpan	2.26°	10.10°
TI pan	4.44°	10.80°

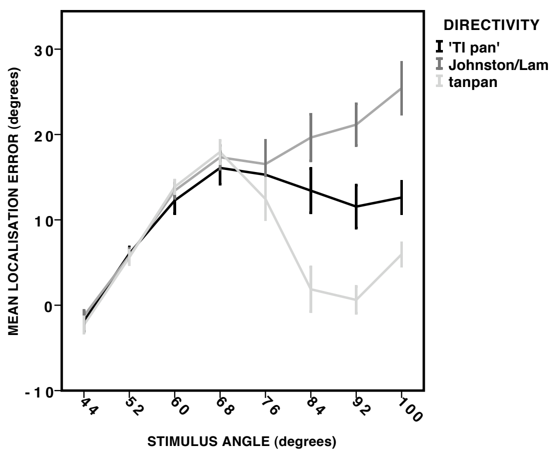

Fig. 11: Mean localization errors for the first listening position averaged across all subjects, S_2 (side). The error bars show the 95% confidence intervals of the mean errors.

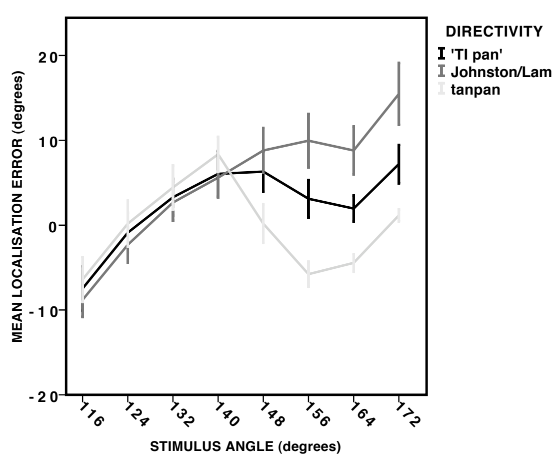

Fig. 12: Mean localization errors for the first listening position averaged across all subjects, S_3 (back). The error bars show the 95% confidence intervals of the mean errors.

subjects gave their reponses. Therefore, pairwise comparisons of standard deviations of the subjective localisation errors were carried out via F-tests. Both *tanpan* and *TI pan* directivities have similar error standard deviations while standard deviation of *Johnston/Lam* directivity was higher. Therefore, we applied a right-tailed test to see if this difference is statistically significant. The null hypothesis that the error variances are equal for *Johnston/Lam* and *tanpan* directivities can be rejected at $\alpha = 0.01$ level ($F(2195, 2195) = 1.6191$; $p < 0.01$; right-tailed). The null hypothesis that the error variances are equal for *Johnston/Lam* and *TI pan* can be rejected at $\alpha = 0.01$ level ($F(2195, 2195) = 1.8515$; $p < 0.01$; right-tailed). The null hypothesis that tanpan and "TI pan" directivities have the same error variance cannot be rejected at $\alpha = 0.01$ level. This indicates that those two directivities provide similar performance in terms of perceived source width.

It should be noted that, in a typical listening condition the localisation accuracy in the front direction will be more important than localisation accuracy at the sides and at the back, directions at which human auditory system is not very accurate. Therefore, we argue that *TI pan* directivity proposed in this paper provides a better choice than both Johnston/Lam directivity and *tanpan* directivity which we have proposed previously.

6. CONCLUSIONS

A new microphone directivity design method based on the concept of stereophonic time-intensity panning was proposed in this paper. The proposed method is based on the minimisation of the error between the microphone sensitivity and time-intensity panning curves, as well as the interchannel crosstalk. The designed microphone directivities are useful in near-coincident multichannel recording as proposed by Johnston and Lam [5]. The design procedure involved the solution of a constrained linear least squares optimisation problem.

Results of a subjective listening test comparing the localisation performances of the proposed microphone array with the original Johnston/Lam directivity and the *tanpan* directivity were also presented. The subjective localisation test involved matching acoustic pointers with the simulated source directions under well-controlled experimental conditions. It was shown that the new design method allows better subjective localisation of simulated sound sources. A more comprehensive evaluation to compare the proposed system with second-order Ambisonics is under way [20].

ACKNOWLEDGMENTS

The work reported in this paper was funded by the Engineering and Physical Sciences Research Council (EPSRC) Research Grant EP/F001142/1.

7. REFERENCES

[1] International Telecommunication Union, "ITU-R BS.1116-1: Methods for the subjective assessment of small impairments in audio systems including multichannel audio systems," 1997.

[2] M. A. Gerzon, "Ambisonics in multichannel broadcasting and video," *J. Audio Eng. Soc.*, vol. 33, no. 11, pp. 859–871, November 1985.

[3] R. Furness, "Ambisonics - an overview," in *Proc. AES* 8^{th} *Int. Conf., "The Sound of Audio"*, 1990, pp. 181–190.

[4] M. M. Boone, U. Horbach, and W. P. J. Bruijn, "Spatial sound-field reproduction by wave-field synthesis," *J. Audio Eng. Soc.*, vol. 43, no. 12, pp. 1003–1012, December 1995.

[5] J. D. Johnston and Y. H. Lam, "Perceptual soundfield reconstruction," Presented at the AES 109th Convention, Los Angeles, USA, Preprint #2399, 22-25 September 2000.

[6] E. de Sena, H. Hacıhabiboğlu, and Z. Cvetković, "Design of a circular microphone array for panoramic audio recording and reproduction: Naturalness," Accepted for presentation at the AES 128th Convention, London, UK, May 2010.

[7] J. Hall and Z. Cvetković, "Coherent multichannel emulation of acoustic spaces," in *Proc. AES* 28^{th} *Conference, Piteå, Sweden*, 30 June - 2 July 2006.

[8] H. Hacıhabiboğlu and Z. Cvetković, "Panoramic recording and reproduction of multichannel audio using a circular microphone array," in *Proc. 2009 IEEE Workshop on Appl. of Signal Process. to Audio and Acoust. (WASPAA'09)*, oct. 2009, pp. 117 –120.

[9] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization.* Cambridge, MA, USA: MIT Press, 1997.

[10] J. D. Johnston and E. R. Wagner, "Microphone array for preserving soundfield perceptual cues," United States Patent, US 6,845,163 B1, 18 January 2005.

[11] M. A. Poletti, "A unified theory of horizontal holographic sound systems," *J. Audio Eng. Soc.*, vol. 48, no. 12, pp. 1155–1182, December 2000.

[12] R. D. Heyser, "Instantaneous intensity," in *Presented at the AES 81st Convention, Los Angeles, USA, Preprint #2399*, 12-16 November 1986.

[13] G. W. Elko, "Differential microphone arrays," in *Audio signal processing for next-generation multimedia communication systems*, Y. Huang and J. Benesty, Eds. Boston, USA: Kluwer Academic Publishers, 2004.

[14] S. P. Lipshitz, "Stereo microphone techniques," *J. Audio Eng. Soc.*, vol. 34, no. 9, pp. 716–744, Sep. 1986.

[15] R. Y. Litovsky, H. S. Colburn, W. A. Yost, and S. J. Guzman, "The precedence effect," *J. Acoust. Soc. Am.*, vol. 106, no. 4, pp. 1633–1654, Oct. 1999.

[16] N. V. Franssen, *Stereophony.* Eindhoven, the Netherlands: Philips Research Laboratories, 1964.

[17] S. Bech and N. Zacharov, *Perceptual Audio Evaluation: Theory, Method and Application.* London, UK: John Wiley & Sons, 2006.

[18] H. Hacıhabiboğlu, B. Günel, and A. M. Kondoz, "Analysis of root displacement interpolation method for tunable allpass fractional-delay filters," *IEEE Tran. Signal Process.*, vol. 55, no. 10, pp. 4896–4906, Oct. 2007.

[19] E. de Sena, H. Hacıhabiboğlu, and Z. Cvetković, "Perceptual evaluation of a circularly symmetric microphone array for panoramic recording of audio," in *Accepted for publication in Proc. 2nd Int. Symp. on Ambisonics, and Spherical Acoustics*, Paris, France, 2010.