nederlands akoestisch genootschap
NAG
journaal
nr. 153
nov. 2000

dr. J. A. Ródenas, dr. R.M. Aarts
Acoustics and Sound Reproduction
DSP Group, Philips Research
Prof. Holstlaan 4 (WY82)
5656 AA Eindhoven, The Netherlands
e-mail: {Josep. Rodenas, Ronald. M. Aarts}@philips.com

POSITION INDEPENDENT STEREOPHONIC SOUND REPRODUCTION

Summary: In this paper the correction of the degradation of stereophonic illusion due to *off-centre listening* is investigated. The main idea here is that the directivity pattern of a loudspeaker array should have a well defined shape such that a good stereo reproduction is achieved in a *large listening area*. Digital filters are designed and applied to individual drivers of linear loudspeaker arrays in order to obtain a directivity pattern of a specific shape. This shape is adapted to the time/intensity trading mechanism of the human auditory system via psychoacoustic experiments within a wide listening area. Therefore, the drivers are fed by digital filters with coefficients which are calculated by means of a numerical optimisation procedure. This was done in such a way that the loudspeaker arrays together radiated sound in a broad range of listening positions in accordance with the time/intensity trading results. This application is referred to as *"Position Independent (PI) Stereo"*.

INTRODUCTION

The basis of stereophony is the ability to create "*phantom*" images. What do we mean by a phantom image? It is known that the brain locates a mono signal originated from a single source by comparing the differences in the arrival time and intensity of that signal at each ear. What if that same mono signal is played through two loudspeakers on either side of the listener? If this happens, then the sound seems to appear from midway between the two loudspeakers since the travelling time of the signal arriving at each ear is the same. This is called a "*phantom*" image (see Figure 1) [1].

Generally it is considered as a serious artefact of the traditional stereo system that the listener is aware of the stereophonic illusion only in a limited region. Optimum stereo perception only occurs if the listener is placed exactly in the median plane between the two loudspeakers. If the head is moved away from the median plane then the stereo effect deteriorates. This *off-centre listening* problem becomes even more serious when the distance between the loudspeakers is not large in comparison with the deviations from the centre position as it occurs in multimedia PC monitor and TV applications; the latter normally has a wider stereo base but in some cases a smaller stereo base is desired. Thus, if the head is moved laterally, the sound rapidly seems to come from the nearest loudspeaker only. This is mainly because of two additional effects: the intensity of the nearest loudspeaker at the listener's head is highest, and its wavefront arrives earlier (law of the first wavefront or precedence effect) (Figure 1) [1, 2].

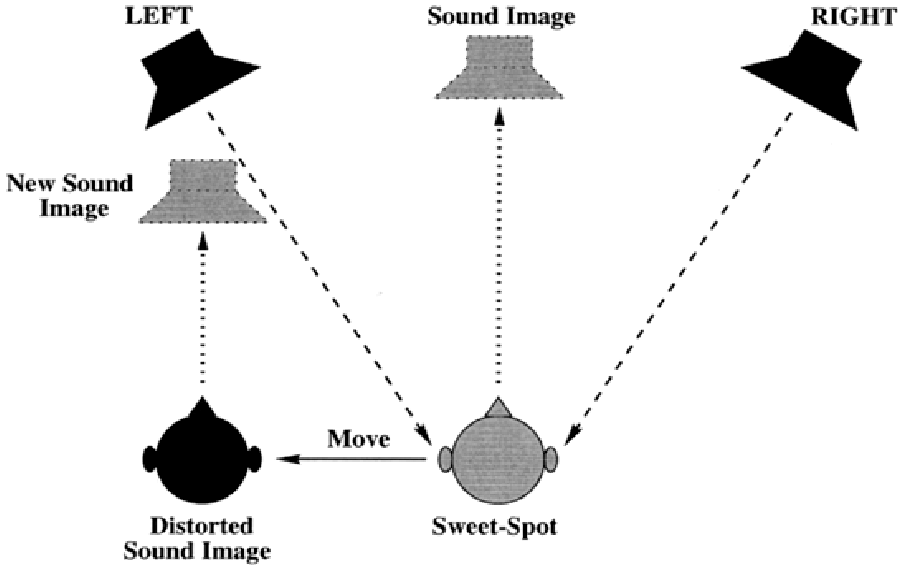

Figure 1 Stereo setup and off-centre listening problem when the listener is not in the sweet spot position.

In general, it can be stated that correct localisation within a wide listening area is beneficial for all applications where a good stereophonic sound image is required: audio, video and car stereo. The idea of achieving an increase in the listening area (*sweet spot widening*) in a stereo setup has been introduced and studied at the beginning of the 90's and was called "*Position Independent (PI) Stereo*". In these works, the idea of "*time/ intensity trading*", realised by an appropriate directivity pattern for the loudspeakers for a two channel stereo arrangement, was presented. The main idea was the following: if the listener moves to the left, the sound intensity from the right loudspeaker must increase, while that of the left loudspeaker must decrease in such a way that the virtual sound image remains in the middle. This can also be seen as a sort of *automatic balance control* depending on the position of the listener. Here, we will continue this idea and we will describe robust digital implementations for the PI-stereo system so as to achieve a better performance of the system in a large listening area. In summary, we will present the current state-of-the-art on PI-stereo sound reproduction to clarify some unclear parts that appeared in previous works, and we will also describe a digital filtering technique which will be applied to the individual drivers of the loudspeaker arrays to obtain a directivity pattern of a specific shape.

BASIS OF POSITION INDEPENDENT STEREO

The PI-stereo system is basically composed of two loudspeaker arrays, each fitted into a single cabinet, with an optimal directivity pattern which has been designed such that a good stereo reproduction is achieved in a large listening area [6, 7, 8, 9]. A "standard" listening setup for PI-stereo for Hi-Fi audio and TV setups is shown in Figure 2.

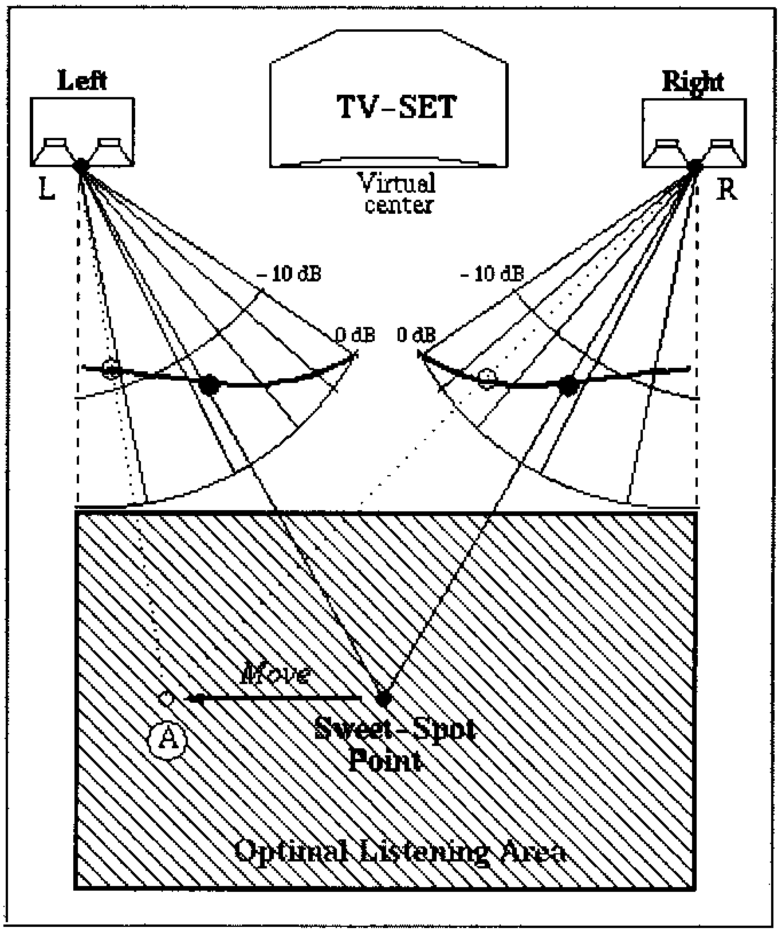

Figure 2 Optimal listening area for PI-stereo reproduction.

Loudspeaker array design and frequency range splitting

It has been proved that an adjustable directivity pattern for a loudspeaker can be realised by using an array of drivers positioned at a small distance from each other [9, 10]. In [3] a clever design of loudspeaker cabinets using several drivers to achieve PI-stereo sound reproduction was described. In that work, a pair of loudspeaker cabinets was equipped with a pair of drivers for each frequency range (high and mid) (see Figure 3). It was also decided to split the audible frequency band into 3 ranges (Figure 4):

1. Lower than frequency f_L = 200 Hz: the left and right channels signals are added (becomes a mono signal). No optimal directivity pattern is required since low frequencies are hard to localise.
2. Between frequencies f_L = 200 Hz and f_H = 2 kHz (mid range): Left and right channels are processed separately.
3. Greater than frequency f_H = 2 kHz (high range): The same as in the mid range.

A pair of loudspeaker cabinets (left and right) were equipped with a pair of drivers, which were separated with a given distance to achieve the two frequency ranges (high

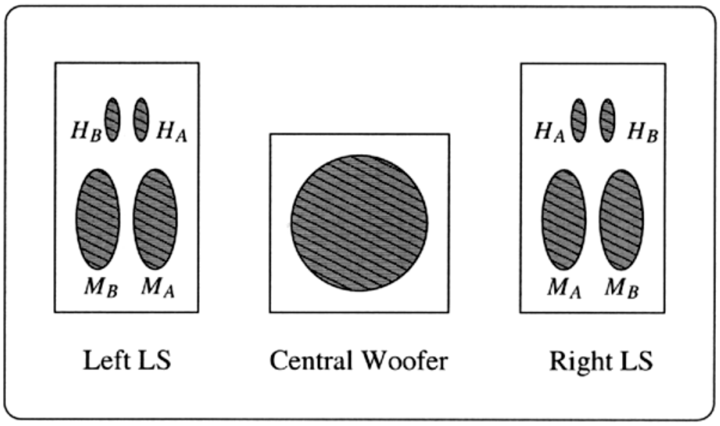

Figure 3 Description of the loudspeaker cabinets for the PI-stereo system. Optionally, a central woofer can be added to the PI-stereo system to reproduce the low frequencies.

and mid), so as to achieve the desired directivity pattern. The low frequencies can be optionally reproduced by means of a woofer (Figure 3). The drivers were finally fed by the calculated digital filters in such a way that both loudspeakers together radiate sound in a broad range of listening positions in accordance with the time/intensity trading experiments.

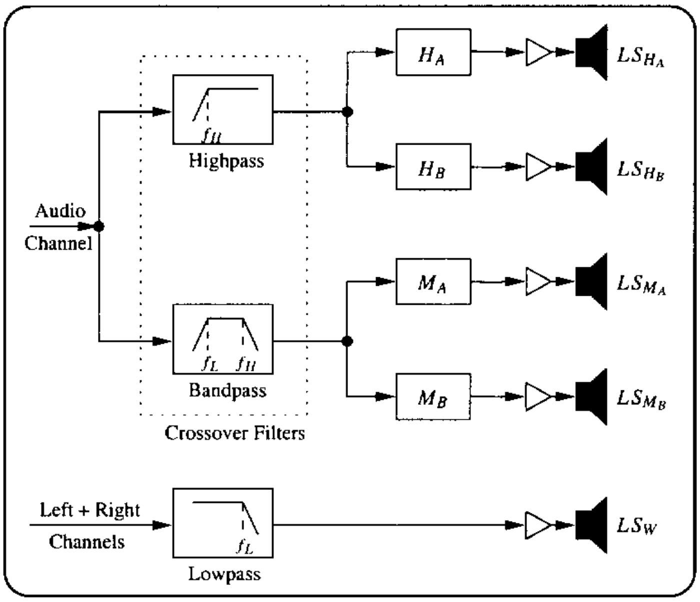

Figure 4 A schematic block diagram for the processing of the audio channels (left and right). The frequency range splitting is achieved by means of crossover filters. Filters H_A and H_B for the high frequency range and filters M_A and M_B for the mid frequency range are obtained by the proposed method. These filters drive the 2 different arrays of loudspeakers, LS_{H_A}, LS_{H_B} and LS_{M_A}, LS_{M_B}. Low frequencies of the audio channels (left and right) are reproduced by a woofer.

OPTIMAL DIRECTIVITY PATTERN

In order to calculate the optimal directivity pattern for the loudspeakers, listening tests in an anechoic room were conducted [3, 5]. During these tests, the differences in intensity levels between right *R* and left *L* loudspeakers for different listening positions to obtain a central sound image were measured (Figure 5). From these experiments, an optimal directivity pattern for the sound sources of the loudspeakers can be determined. We introduce next the approach of finding the optimal digital filters to drive the arrays of loudspeakers so as to achieve the required directivity pattern (*target function*).

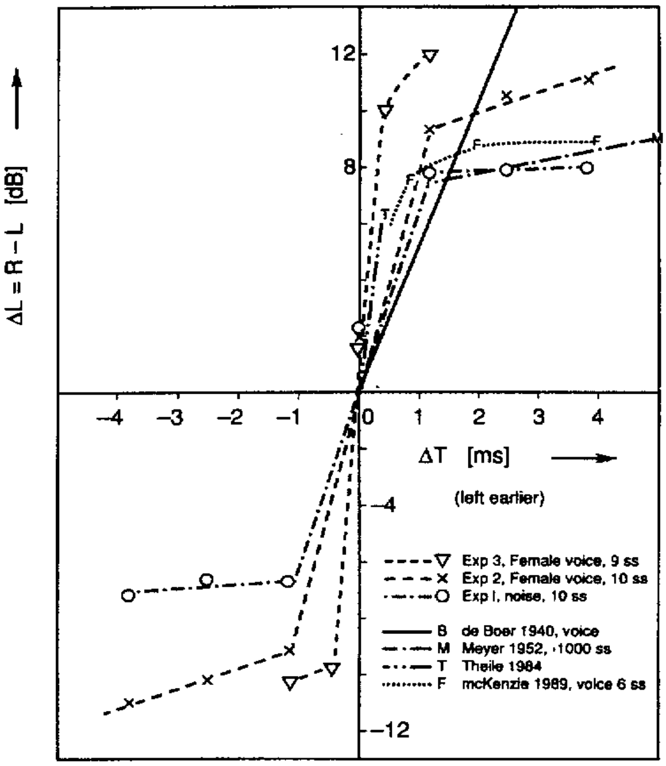

Figure 5 Time/intensity trading results from the experiments carried out in [5] in comparison with other experiments from the literature.

Calculation of the target function

Goossens [3] defined an *ad hoc* mathematical expression (*target function* F^t) for the sound pressure levels for the directivity patterns of the loudspeakers. This expression, which depends on two linear parameters (*A* and *B*) and one non-linear parameter (*C*), was defined as follows

$$F^t = 20\lg\left[A + B\sin^C\theta\right] \tag{1}$$

where θ is a set of angles. In order to fit time/intensity trading data to this target function, we can use the well known *non-linear least squares* method and the *linear optimisation* problem, which give the values for *A*, *B* and *C* [11, 12]. Making the following substitution $L^t = 10^{[F^t/20]}$, the system we have is then:

$$\begin{bmatrix} 1 & \sin^C \theta_1 \\ 1 & \sin^C \theta_2 \\ \vdots & \vdots \\ 1 & \sin^C \theta_n \end{bmatrix} \cdot \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} L_1^t \\ L_2^t \\ \vdots \\ L_n^t \end{bmatrix} \tag{2}$$

Thus, in matrix form this reduces to:

$$\mathbf{T} \cdot \underline{\mathbf{x}} = \underline{\mathbf{L}}^{\mathbf{t}} \tag{3}$$

where $\mathbf{T}$ is the $[N\times 2]$ matrix of the target function to optimise, $\underline{\mathbf{x}}$ is the $[2\times 1]$ vector of the linear parameters to find and $\underline{\mathbf{L}}^{\mathbf{t}}$ is the $[N\times 1]$ vector of the experimental time/intensity trading data to fit. This problem can be solved by first finding out the nonlinear parameter C and then the linear parameters A and B. This results in the best estimation for the system which gives the *minimum error* in the fit to $\mathbf{T}$ by

$$\min_{x\in R^2} \left\| \mathbf{T} \cdot \underline{\mathbf{x}} - \mathbf{L}^{\mathbf{t}} \right\|_2^2 \tag{4}$$

We then obtain the following target function which approximates the correct directivity pattern for the loudspeakers for a given time/intensity trading data:

$$F^t = 20 \log\left[0.11 + 1.51 \sin^3 \theta\right]. \tag{5}$$

Figure 6 shows the directivity polar plot for the optimised directivity pattern for the calculated time/intensity trading data.

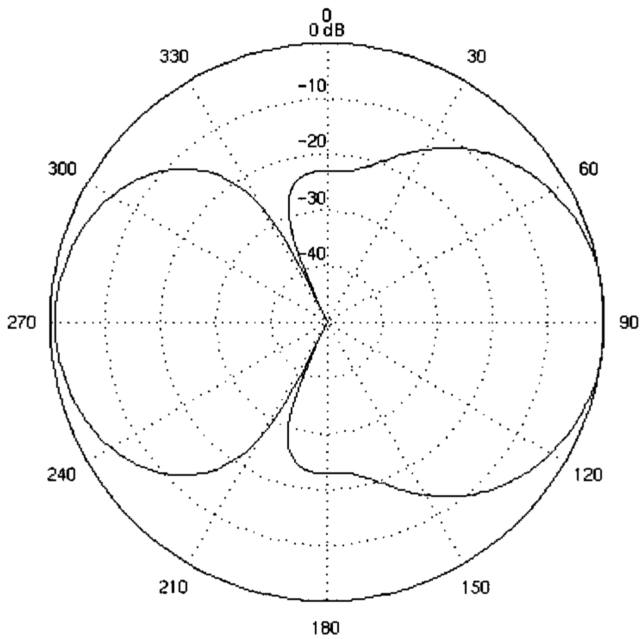

Figure 6 Polar plot of the calculated optimal directivity pattern of Equation (5).

FILTERING TECHNIQUE FOR DRIVING THE LOUDSPEAKER ARRAYS

In the next section, we focus on the optimisation problem of estimating the required FIR filter coefficients which achieve the calculated optimal target function in a wide listening area.

Optimisation problem: Digital FIR filters

We consider here the general case of having a *linear array* of N equal and equidistant omnidirectional sound sources separated by a distance d. It is well known that a FIR filter has a transfer function $H(z)$ of the form:

$$H(z) = \sum_{n=1}^{N} h_n z^{-n} \tag{6}$$

Using the acoustic pressure equation for a simple source and given that each sound source is driven by a FIR filter $h_{n,m}$ of M coefficients with N number of drivers, the *total sound pressure level* $P_{k,\ell}$ for a frequency ω_k and an angle θ_ℓ is given by:

$$P(\omega_k,\theta_\ell) = \sum_{n=1}^{N}\sum_{m=1}^{M} h_{n,m} e^{j\omega_k\left[\left(n-\frac{N+1}{2}\right)\frac{d}{c}\sin\theta_\ell - \frac{m-1}{f_s}\right]} \tag{7}$$

where f_s is the sampling frequency. The last expression can also be written as:

$$\begin{aligned} P(\omega_k,\theta_\ell) = \quad & \sum_{n=1}^{N}\sum_{m=1}^{M} h_{n,m}\left[\cos\left(\omega_k\left[\left(n-\frac{N+1}{2}\right)\frac{d}{c}\sin\theta_\ell - \frac{m-1}{f_s}\right]\right)\right. \\ & \left. + j\sin\left(\omega_k\left[\left(n-\frac{N+1}{2}\right)\frac{d}{c}\sin\theta_\ell - \frac{m-1}{f_s}\right]\right)\right] \end{aligned} \tag{8}$$

Since we want to obtain a sound pressure level equivalent to that given by the optimal target function in Equation (5), we can define the *required sound pressure level* $S_{k,\ell}$, for the target function as:

$$\begin{aligned} S_{k,\ell} \quad & = F_\ell^t e^{\left[-j\omega_k \frac{M-1}{2f_s}\right]} \\ & = F_\ell^t\left[\cos\left(\omega_k \frac{M-1}{2f_s}\right) \pm j\sin\left(\omega_k \frac{M-1}{2f_s}\right)\right] \end{aligned} \tag{9}$$

where F_ℓ^t are the sound pressure levels at a angles ℓ given by the target function previously described, and the phase term corresponds to a *constant group delay* of $T(M–1)/2$. We can now formulate the *least squares optimisation* problem, so as to find the $h_{n,m}$ FIR coefficients for the different array drivers, as follows:

$$\mathbf{T} \cdot \underline{\mathbf{H}} = \underline{\mathbf{S}} \tag{10}$$

where the vector $\underline{\mathbf{H}}$ of dimensions $[NM \times 1]$ contains the filter coefficients and the matrix

T of dimensions $[2KL \times NM]$ is defined by Equation (8) as:

$$T_{(k-1)L+\ell,(n-1)M+m} = \cos\left(\omega_k\left[\left(n-\frac{N+1}{2}\right)\frac{d}{c}\sin\theta_\ell - \frac{m-1}{f_s}\right]\right) \tag{11}$$

$$T_{(K+k-1)L+\ell,(n-1)M+m} = \sin\left(\omega_k\left[\left(n-\frac{N+1}{2}\right)\frac{d}{c}\sin\theta_\ell - \frac{m-1}{f_s}\right]\right) \tag{12}$$

and the vector $\underline{\mathbf{S}}$ of dimensions $[2KL \times 1]$ is defined by Equation (9) as:

$$S_{(k-1)L+\ell,1} = F_\ell^t \cos\left(\omega_k \frac{M-1}{2 f_s}\right) \tag{13}$$

$$S_{(K+k-1)L+\ell,1} = -F_\ell^t \sin\left(\omega_k \frac{M-1}{2 f_s}\right) \tag{14}$$

This gives the following matrix form system:

$$\begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1,NM} \\ T_{21} & T_{22} & \cdots & T_{2,NM} \\ \vdots & \vdots & \ddots & \vdots \\ T_{2KL,1} & T_{2KL,2} & \cdots & T_{2KL,NM} \end{bmatrix} \cdot \begin{bmatrix} h_{11} \\ \vdots \\ h_{1M} \\ h_{21} \\ \vdots \\ h_{NM} \end{bmatrix} = \begin{bmatrix} S_{11} \\ \vdots \\ S_{1L} \\ S_{21} \\ \vdots \\ S_{2KL} \end{bmatrix} \tag{15}$$

The resulting FIR coefficients $h_{n,m}$ from the optimisation method are real for all frequencies. Hence from Equation (7) it follows that the directivity pattern is symmetric.

$$P_{k,\ell} = \sum_{n=1}^{N} h_n e^{\left[j\pi f_s\left(n-\frac{N+1}{2}\right)\frac{d}{c}\sin\theta_\ell\right]} \Rightarrow \left|P_{k,\ell}(\theta_\ell)\right| = \left|P_{k,\ell}(-\theta_\ell)\right| \tag{16}$$

This can cause problems if the optimal target function is strongly asymmetric. The most practical solution to overcome this difficulty in the optimisation problem is turning the whole array by an angle θ_0, which simply corresponds to a rotation of the target function by an angle of $-\theta_0$, that is replacing all θ's by $\theta - \theta_0$.

Therefore, this proposed method results in the calculation of the FIR filter coefficients $h_{n,m}$ (n = 2 and m = 20) for the two sound sources or drivers (we will refer to them from now on as being *A* and *B*). We performed some simulations and it appeared that they converged to the optimum choice of filter coefficients for each of the two frequency bands (mid and high).

Figure 7 shows the frequency responses of the calculated digital filters. Figure 8 illustrates the phase difference response between the two *A* and *B* filters for both drivers in the mid and high frequency ranges. We can see that for the mid range filters, the *phase difference* $\gamma = \varphi_A - \varphi_B$ ranges from 172.5˚ (200 Hz) to 134˚ (2 kHz), and for the high range filters it goes from 166˚ (2 kHz) to 126.5˚ (12 kHz).

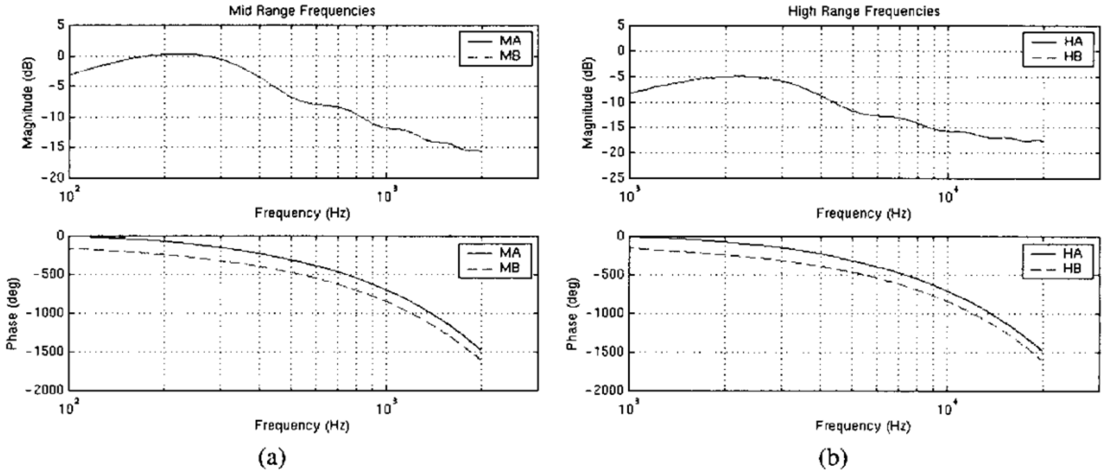

Figure 7 Frequency responses of the digital FIR filters for the two drivers (*A* and *B*) with N = 20. (a) Mid range frequencies and (b) High range frequencies.

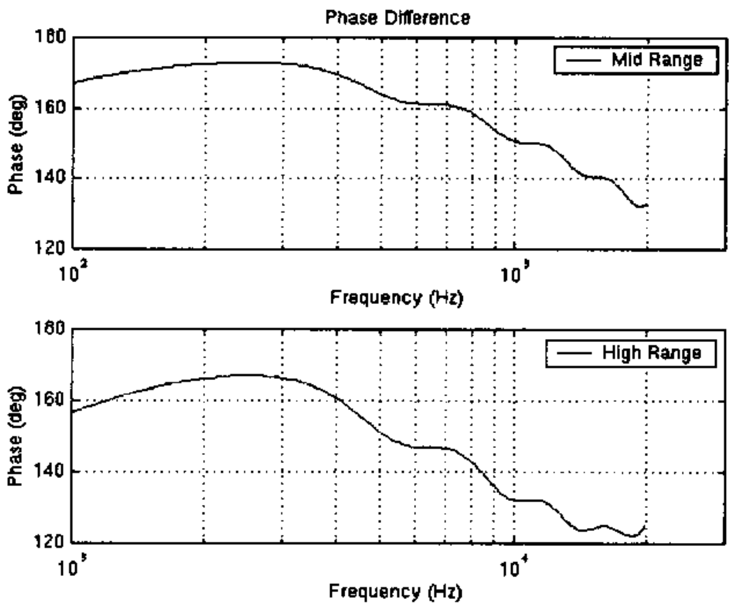

Figure 8 Phase differences between the two drivers (*A* and *B*) for the digital FIR filters.

MEASURED DIRECTIVITY PATTERNS AND LISTENING EXPERIMENTS

For the study of the directivity polar plots the right loudspeaker response was considered, that is the main lobe is on the right-hand side and the minor lobe is on the left-hand side. Another consideration here is that when listening to the PI-stereo system the loudspeaker boxes are face on, not at 30˚ pointing inwards to the listener as in normal stereo. To compensate for this, the frequency responses of the loudspeaker boxes were taken at 30˚ clockwise for the left and 30˚ anti-clockwise for the right box so that the frequency response in the middle of the working region was considered as opposed to the response at the edge.

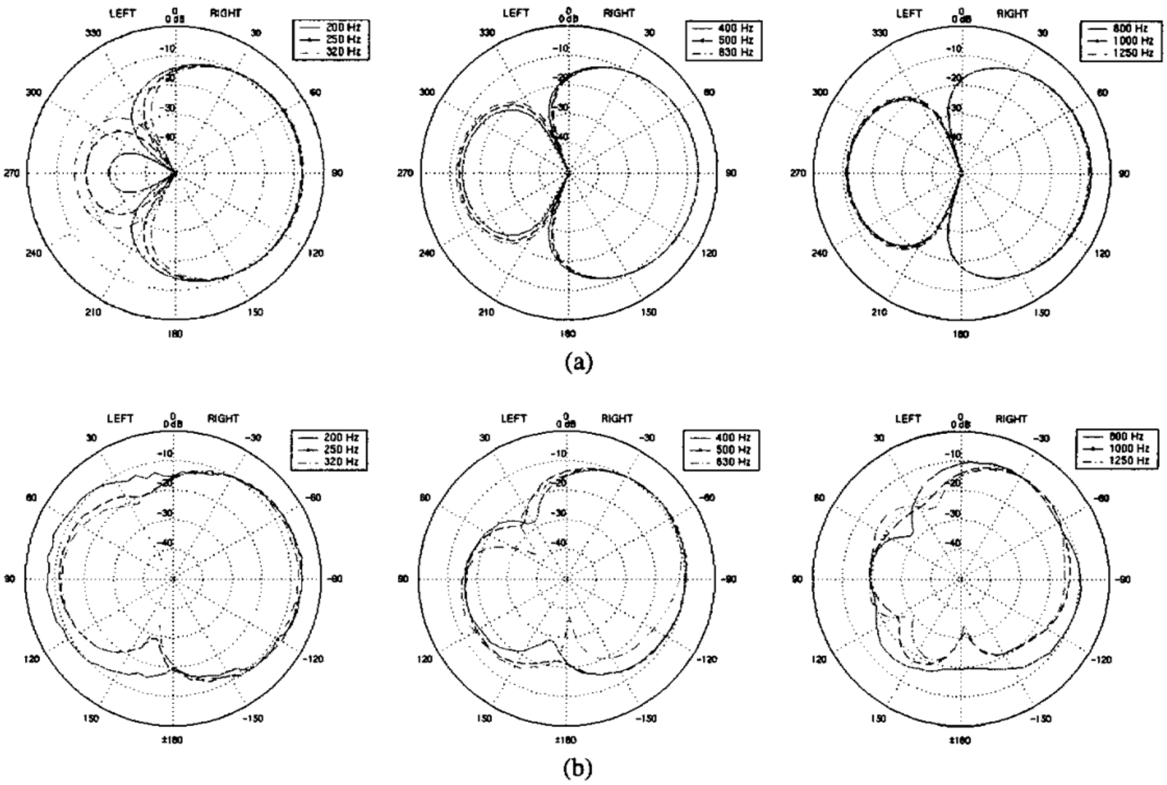

Figure 9 Theoretical (a) and measured (b) directivity pattern plots for the PI-stereo system for frequencies ranging from 200-1250 Hz. Polar plots for higher frequency ranges appeared to be very close to these ones.

Figure 9(a) shows simulations of theoretical directivity polar plots for the PI-stereo system for frequencies ranging between 200-1250 Hz. Figure 9(b) shows the measured directivity polar plots in an anechoic room using the right loudspeaker box. Note here that all polar plots have been normalised at 30˚ which is the centre of the considered working region. Theoretical and measured directivity polar plots for higher frequencies appeared to be very close to the ones in this figure and due to space limitations are not reproduced in this paper.

Preliminary listening experiments have shown that the PI-stereo system worked as predicted by the theoretical models. Correct central sound localisation for voices and other effects have also been demonstrated for a number of listening positions. Both normal stereo and PI-stereo were reproduced by the same loudspeaker cabinets, so there was no shift in the stereo image. The degradation of the position independent stereo image for lateral positions was acceptable and was observed to be of better quality compared to that for normal stereo. The difference between normal stereo and PI-stereo for central listening was almost unperceivable. The stereo sound sensation, in particular the placement of central voices, was independent of the listening position within a large area.

CONCLUSIONS

We have described and developed a new stereo sound reproduction system that offers a natural high quality stereophonic sound image in a large listening area. A digital filtering technique has been applied to individual drivers of linear loudspeaker arrays in order to obtain a directivity pattern having an optimal shape. This optimal shape was adapted to time/intensity trading experiments for enlarging the sweet spot area.

The outcome of this work showed that optimal directivity patterns for loudspeaker arrays in stereophonic applications can be very useful for sweet spot widening. PI-stereo can be applied to any systems where a good stereo sound reproduction in a large listening area is required, such as: TV-sets, Hi-Fi's, multimedia, home theatre, car stereo and portable audio.

REFERENCES

1. J. Blauert, 1983. Spatial Hearing, The MIT Press, Cambridge Massachusetts, 3rd edition.
2. N.V. Franssen, July 1960. *Some Considerations on the Mechanism of Directional Hearing*, PhD thesis, Technical University of Delft, The Netherlands.
3. P. Goossens, 1990. Unpublished Internal Technical Report, Philips Research Labs.
4. R.A. Overton, Andrew I. Ruxton, and R.M. Aarts, 1991. Unpublished Internal Technical Report, Philips Research Labs.
5. R.M. Aarts, 1993. Enlarging the sweet-spot for stereophony for (HD)TV and audio applications, In *94th AES Convention*, Berlin, Germany, Preprint 3473.
6. B.B. Bauer, April 1960. Broadening the area of stereophonic perception, J. Audio Eng. Soc., **8**(2):91-94.
7. E.J. Jordan, February 1971. Loudspeaker stereo techniques: How to combine left and right signals and get the message from the medium, Wireless World, pages 67-70.
8. J. Kates, November 1980. Optimum loudspeaker directional patterns, J. Audio Eng. Soc., **28** (11).
9. M.F. Davis, November 1987. Loudspeaker systems with optimized wide-listening area imaging, J. Audio Eng. Soc., **35**(11).
10. M.F. Davis, March 1985. Loudspeaker system, U.S. Patent, (4,503,553).
11. D. Marquardt, 1963. An algorithm for least-squares estimation of nonlinear parameters, SIAM J. Appl. Math., **11**:431-441.
12. J.E. Dennis, 1977. *Nonlinear Least Squares: State of the Art in Numerical Analysis*, Academic Press.