

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
30 September 2004 (30.09.2004)

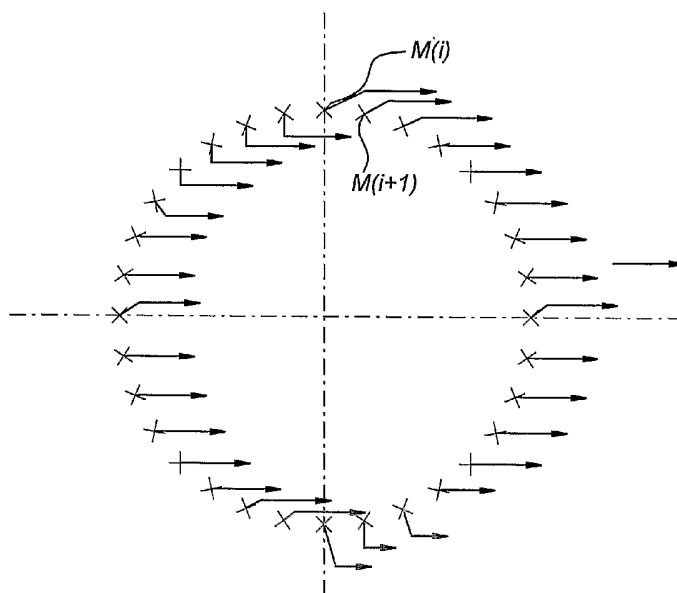
PCT

(10) International Publication Number
WO 2004/084577 A1

- (51) International Patent Classification⁷: H04R 1/40, 3/00
- (21) International Application Number: PCT/NL2004/000169
- (22) International Filing Date: 19 March 2004 (19.03.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 03075820.5 21 March 2003 (21.03.2003) EP
- (71) Applicant (for all designated States except US): TECHNISCHE UNIVERSITEIT DELFT [NL/NL]; Julianalaan 134, NL-2628 BL Delft (NL).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): HULSEBOS, Edo, M. [NL/NL]; Warande 26, NL-3705 ZC Zeist (NL). DE VRIES, Diemer [NL/NL]; Arendshof 12, NL-2651 WL Berkel en Rodenrijs (NL). BOONE, Marinus, Marias [NL/NL]; Voorweg 105A, NL-2715 NG Zoetermeer (NL). SCHUURMANS, Thomas, J., G. [NL/NL]; Lange Niezel 24, NL-1012 GT Amsterdam (NL).
- (74) Agent: VAN WESTENBRUGGE, Andries et al.; Nederlandsch Octrooibureau, Scheveningseweg 82, P.O. Box 29720, NL-2502 LS The Hague (NL).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: CIRCULAR MICROPHONE ARRAY FOR MULTI CHANNEL AUDIO RECORDING



(57) Abstract: An audio system has a circular microphone array with a number of microphones arranged on a circle for receiving a sound field. A digital signal processor is provided for processing output signals from these microphones. To establish well controlled and sharp directivity patterns the audio system performs a weighted integration of the sound field on predetermined portions of the circle. The weighted integration can be performed in a pre-processor reducing the number of channels by creating a smaller number of virtual microphones or by selecting specially designed microphones.

WO 2004/084577 A1



Published:

- *with international search report*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

CIRCULAR MICROPHONE ARRAY FOR MULTI CHANNEL AUDIO RECORDING

Field of the invention

5 The present invention relates to an audio system with a circular microphone array comprising a predetermined first number of microphones arranged on a circle, arranged for receiving a sound field and a digital signal processor for processing output signals from these microphones.

10 Prior art

Such a system is known from WO-01/58209-A1 that discloses an audio system with a circular microphone array for sound field recording. The system comprises a digital signal processor, frequency compensation filters and a sum and difference network. A
15 disadvantage of this prior art system is that a processor with a number of acquisition channels equal to the number of microphones is needed, whereas this number may be high, e.g., several hundreds. This necessitates a complex processor.

In previous AES conventions, papers were presented by E. Hulsebos e.a. on the subject of circular microphone arrays. In preprint 5337 of the 110th AES convention in
20 Amsterdam, which was published in a slightly modified form in the October 2002 issue of the AES journal, an elegant method was described to decompose the sound field measured on a circular array in terms of incoming and outgoing cylindrical harmonic solutions of the 2D wave equation [1,2]. These cylindrical harmonics were shown to be very closely related to the plane wave decomposition, a description of the sound field
25 that proved very useful for auralization purposes. In preprint 5579 presented at the 112th AES convention in Munich the theory for circular arrays was extended to focusing and spatial filtering to control aperture [3]. Furthermore, it was shown that it is possible to simulate virtual microphones with any desired directivity properties in post-processing which is not possible with currently available single microphones.

30

Summary of the invention

The object of the invention is to provide an audio system with a processor with a limited number of acquisition channels and yet providing a good frequency-
5 independent angular resolution over a large frequency range.

To that end, the audio system as defined at the outset is arranged to perform a weighted integration of the sound field on predetermined portions of the circle. With such a system the desired frequency-independent angular resolution (frequency independent
10 beam forming) can be obtained with only a relatively small number, e.g., 24, of acquisition channels. Moreover, well controlled and sharp directivity patterns can be obtained. There are no compromises anymore in terms of coverage, source localisation and channel separation.

15 In a first embodiment, the weighted integration of the sound field is performed by a preprocessor, arranged:

- (a) to receive the output signals from the first number of microphones,
- (b) to combine them to create a second number of virtual microphones to generate a second number of virtual microphone output signals and
- 20 (c) to transmit these virtual microphone output signals to the digital signal processor, the second number being less than the first number.

In such a system a very large number of microphones, e.g., 288, can be used whereas the number of virtual microphone output signals may be drastically reduced, e.g. to 24
25 signals, that can be easily captured using a multi-track computer interface with 24 acquisition channels and can be post-processed into up to 12 discrete reproduction audio channels.

In an implementation of this first embodiment, the pre-processor is arranged:

- 30 - to use a predetermined panning mechanism in action (b) to create the second number of virtual microphone output signals from the output signals, and the signal processor is arranged:
 - to calculate a 2D Fourier transform for the virtual microphone output signals,

- to calculate cylindrical harmonic components from the 2D Fourier transform,
- to calculate an inverse Fourier transform of the cylindrical harmonic components.

The panning mechanism may be implemented by using a resistor ladder network.

5

In order to map the second number of virtual microphone output signals to a desired number of loudspeakers, the processor may be arranged to process the inverse Fourier transform by a matrix operation to generate processor output signals for another set of virtual microphones.

10

In a second embodiment of the invention, the microphones have microphone capsules with different heights and are extended in space such that integration of a sound field over a predetermined area in space is done by the microphones. In this embodiment, the number of microphones can be kept as low as the number of acquisition channels.

15

In this second embodiment, the microphone capsules may be selected from the following group: microphone capsules with a triangular, Hanning, Hamming, and Chebyshev weighting window.

20

Also in this second embodiment, the number of microphone output signals may be mapped to a desired number of loudspeakers by performing a suitable matrix operation.

25

The invention also relates to a method of processing output signals from a circular microphone array, comprising a predetermined first number of microphones arranged on a circle for receiving a sound field, characterized by performing a weighted integration of the sound field on predetermined portions of the circle.

The method comprises similar embodiments as the system defined above.

30

The invention also relates to a computer program product, comprising data and instructions to allow processing of the actions (a), (b) and (c), as defined above.

Finally, the invention relates to a data carrier provided with such a computer program.

Brief description of the drawings

Below, the invention will be explained with reference to some drawings which are only
5 intended as examples and to clarify the invention but not to limit its scope. The
numbers in the examples to follow are also intended to clarify and not to limit the
invention. The invention is only limited by the annexed claims and its technical
equivalences.

Figures 1a, 1b and 1c show a 2D Fourier transform of a simulated circular plane wave
10 recording, using omni, figure-of-eight and cardioid microphones, respectively;

Figure 2 shows a cylindrical harmonics decomposition for a circular array with a 1
meter radius;

Figures 3a and 3b show a desired and actual directivity pattern, respectively of a virtual
microphone created using a 1 meter radius array;

15 Figures 4a-4f show a circular array processing for 12 virtual microphones;

Figures 5a-5e show a circular array processing for 12 virtual microphones according to
the invention;

Figures 6a, 6b show panning in the analog signal domain using small microphones with
weighting resistor ladders (a), and using an extended microphone membrane to apply
20 spatial weighting and integration of the sound field (b), respectively;

Figures 7a, 7b, 7c, 7d show four different panning windows and their frequency
responses;

Figures 8a-8h show aliasing performance of several different panning windows
recording a plane wave with 0 degree incidence, ideally being reproduced by 1 channel
25 only;

Figures 9a-9h show aliasing performance of several different panning windows
recording a plane wave with 15 degree incidence, ideally being panned equally between
2 channels only;

Figure 10 shows virtual microphone directivity patterns;

30 Figure 11 shows a matrix for 12 channel output;

Figure 12 shows a matrix for 5.1 output using a gradual panning between all 5
channels;

- Figure 13 shows desired and actual virtual microphone directivity patterns for 5.1 output using gradual panning between all 5 channels. The black lines are the desired patterns and the coloured areas are the actual patterns;
- Figure 14 shows a matrix for 5.1 output using sharp transition between front and rear channels;
- Figure 15 shows desired and actual virtual microphone directivity patterns for 5.1 output using sharp transition between front and rear channels. The black lines are the desired patterns and the coloured areas are the actual patterns;
- Figure 16 shows a matrix for 24 discrete output channels;
- Figure 17 shows desired and actual virtual microphone patterns for 24 discrete output channels;
- Figure 18 shows frequency response of the 12 discrete channels for different elevation angles;
- Figure 19 shows noise levels from a single microphone and the complete array system;
- Figure 20 shows an arrangement for carrying out the invention.

Description of embodiments

- One implementation of the system according to the invention is shown in figure 20.
- Figure 20 shows a circular array of microphones $M(i)$, $i = 1, 2, \dots, I$. Each of the microphones $M(i)$ is located on a circle, as shown, and is provided with an output to provide an output signal, as schematically indicated with an arrow connected to each of the microphones $M(i)$.
- The outputs of the microphones $M(i)$ are provided to a pre-processor 1. The pre-processor 1 is provided with J outputs $O_{pp}(j)$, $j = 1, 2, \dots, J$. In accordance with an implementation of the invention, $J < I$.
- The outputs $O_{pp}(j)$ are connected to a digital signal processor 3. The digital signal processor 3 has N outputs $O_p(n)$, $n = 1, 2, \dots, N$.
- The outputs $O_p(n)$ are connected to N loudspeakers $L(n)$, $n = 1, 2, 3, \dots, N$. In the embodiment of figure 1, the focus is on building a full array, which can also be used for live audio recording purposes. The pre-processor 1 receives all the output signals from the microphones $M(i)$. The pre-processor 1 is arranged to create a number of virtual microphones on the array having well controlled high order directivity patterns to

deliver a large number of discrete output channels $O_{pp}(j)$ without any unnecessary cross talk between them. In the embodiment of figure 20, the number J of virtual microphones is smaller than the number I of the microphones $M(i)$.

A prototype array was designed to deliver up to $N = 12$ discrete coincident channels of audio with good channel separation down to 100 Hz and a spatial aliasing frequency above 15 kHz. To achieve this, a circle with a radius of 1 m and 288 small cardioid microphones was used. The high number of microphones was not a big problems in this case: various cheap and reasonable quality capsules are available on the market. The number $I = 288$ microphones was reduced to $J = 24$ acquisition channels for the digital signal processor 3. The following sections show how this can be done.

Wave decomposition

If on the circular array both pressure and normal particle velocity are recorded, the approach is to decompose the recorded sound field into cylindrical harmonics:

$$\begin{aligned}
 M^{(1)}(k_\theta, \omega) &= \frac{H_{k_\theta}^{(2)}(kR)P(k_\theta, \omega) - H_{k_\theta}^{(2)}(kR)j\rho cV_n(k_\theta, \omega)}{H_{k_\theta}^{(1)}(kR)H_{k_\theta}^{(2)}(kR) - H_{k_\theta}^{(2)}(kR)H_{k_\theta}^{(1)}(kR)} \\
 M^{(2)}(k_\theta, \omega) &= \frac{H_{k_\theta}^{(1)}(kR)P(k_\theta, \omega) - H_{k_\theta}^{(1)}(kR)j\rho cV_n(k_\theta, \omega)}{H_{k_\theta}^{(2)}(kR)H_{k_\theta}^{(1)}(kR) - H_{k_\theta}^{(1)}(kR)H_{k_\theta}^{(2)}(kR)}
 \end{aligned} \tag{0.1}$$

In these equations $M^{(1)}$ and $M^{(2)}$ are the incoming and outgoing cylindrical harmonic decompositions of the sound field, $P(k_\theta, \omega)$ and $V_n(k_\theta, \omega)$ are the spatial and temporal Fourier transforms of the pressure $p(\theta, t)$ and normal velocity $v_n(\theta, t)$ measured on a circular array with radius R , $k = \omega/c$ is the temporal wave number, $k_\theta = \dots, -1, 0, 1, \dots$ is the angular wave number and $H_{k_\theta}^{(1,2)}$ are the Hankel functions of the first and second kind [4]. See [3] for more details on the cylindrical harmonic decomposition. Since in normal recording situations all sound waves are coming from the exterior of the circle, only the incoming part $M^{(1)}$ needs to be considered. The plane wave decomposition can easily be calculated from $M^{(1)}$ using

$$s_\infty(\theta, \omega) = \frac{1}{2\pi} \sum_{k_\theta} j^{(1-k_\theta)} M^{(1)}(k_\theta, \omega) e^{jk_\theta\theta} \tag{0.2}$$

which is up to a rotation factor $j^{(1-k_\theta)}$ equal to the inverse Fourier transform of $M^{(1)}$

[1,2]. Instead of using both pressure and normal particle velocity, it is sufficient to record the sound field on the array using outward pointing cardioid microphones. Using only omnidirectional or only figure-of-eight microphones is not sufficient though; not all cylindrical harmonics can be obtained from such a recording. In equation (1) it is found that the cylindrical harmonics are proportional to the 2-dimensional Fourier transform components of the recorded sound field. In figure 1 the amplitudes for the 2D Fourier transforms of a plane wave recorded on a circular array using omni, figure-of-eights and cardioid are shown. The omni and figure-of-eight recordings from figure 1(a) and (b) both have zeros in their Fourier transform components and therefore cannot be used separately for a calculation of the non-zero cylindrical harmonic components of the plane wave shown in figure 2. However, since the zero components for the omni and the figure-of-eight microphone arrays occur at different frequency components, they can be used together for a proper cylindrical harmonic decomposition as is done in equation (1). For the cardioid case the 2D Fourier transform components are already free from zeros and can easily be used for this purpose as well. This way the number of acquisition channels can be reduced by a factor 2. Note however that it is not possible to separate between incoming and outgoing sound waves when using only one set of outward pointing cardioids. Also equation (1) and (2) cannot be used directly in this case, but the appropriate plane wave decomposition operator can easily be calculated numerically by creating an inverse for the 2D Fourier transform of a plane Dirac wave simulated on a cardioid array.

Virtual microphones

In one embodiment, the aim is to deliver 12 discrete coincident channels of audio. For that purpose 12 virtual microphones with suitable directivity characteristics can be created by combining the appropriate cylindrical harmonic components with appropriate weighting factors. The only limitation in creating virtual microphones is that the directivity pattern of the virtual microphone cannot contain cylindrical harmonics with spatial frequencies k_θ beyond the angular resolution available for the given array size. Notice that the maximal available angular frequency component is proportional to the radius (aperture) of the circular array and the temporal frequency as can be seen in figure 1. In practice this usually means that for high temporal frequencies the created virtual microphones will match the desired directivity pattern

perfectly, but below a certain temporal frequency the highest angular frequency components in the desired directivity pattern are missing. The value of this temporal cutoff frequency depends on the radius of the circular array and the sharpness of the desired directivity pattern. The sharper it is, the higher the maximal required angular frequency is and the higher the temporal cutoff frequency will be.

Virtual microphones can be created easily using the plane wave decomposition. The plane wave decomposition can be considered as a set virtual microphone responses, which have the maximal directivity sharpness that can possibly be obtained for the given microphone array aperture. Suppose $s_{\infty}(\theta, \omega)$ is the plane wave decomposition of a sound field and $d(\theta, \omega)$ is the possibly frequency dependent (in this paper only frequency independent directivity patterns are used though) desired directivity pattern of a virtual microphone. Then the output of the virtual microphone can be obtained from the plane wave decomposition using

$$O(\omega) = \sum_{\theta} d(\theta, \omega) \cdot s_{\infty}(\theta, \omega) \quad (0.3)$$

The actual obtained directivity pattern $d_c(\theta, \omega)$ of the virtual microphone is given by

$$D_c(k_{\theta}, \omega) = D(k_{\theta}, \omega) \cdot S_{\text{plane}}(k_{\theta}, \omega) \quad (0.4)$$

where $D_c(k_{\theta}, \omega)$ and $D(k_{\theta}, \omega)$ are the angular Fourier transform of $d_c(\theta, \omega)$ and $d(\theta, \omega)$, respectively, and $S_{\text{plane}}(k_{\theta}, \omega)$ is the 2D Fourier transform of the plane wave decomposition of a plane Dirac wave with angle of incidence equal to zero and arriving at the center of the circle at $t = 0$, which is limited in angular resolution by the radius of the circular array. $S_{\text{plane}}(k_{\theta}, \omega)$ for a circle with a 1 meter radius is shown in figure 2. The actual directivity pattern is thus a band limited version of the desired directivity pattern in terms of the angle. See figure 3a for an example of one of the desired directivity patterns that can be used for obtaining 12 discrete audio channels. The realized directivity pattern using a 288 microphone circular array with a radius of 1 meter is shown in figure 3b. Notice that apart from the angular bandwidth limitation starts taking place at 250 Hz, also spatial aliasing artifacts occur above 15 kHz. These are determined by the distance between the microphones and can be avoided by using even more closely spaced microphones.

Analog Preprocessing

The number of microphones used in the circular recording array is not a big problem as mentioned before; various cheap and reasonable quality capsules are available on the market. The problem however, which is going to be solved next, is the number of acquisition channels. A wave decomposition has to be calculated from all the 288 microphone signals, which must be done in the digital domain using 2D Fast Fourier transform algorithms. This means that 288 microphones need to be digitally acquired and processed in real time, something which isn't feasible in practice or at least very expensive.

10

In order to understand how this problem can be solved easily by drastically reducing the number of acquisition channels, a closer look at the processing scheme is required. See figure 4. In figure 4a a circular recording of a plane wave is shown. The 2D Fourier transform for this recorded wave is then calculated and the amplitudes of these complex components are shown in figure 4b. From this 2D Fourier transform the cylindrical harmonic components in figure 4c are calculated. By calculating the 2D inverse Fourier transform of this field the 288 channel plane wave decomposition from figure 4d can be obtained using equation (2). The last step is to combine these 288 plane wave virtual microphone signals using equation (3) to obtain a lower number of virtual microphones with the desired directivity patterns for the multichannel sound format in figure 4e. As can be seen in figure 4f, which is the 2D Fourier transform of the data from figure 4e, these final outputs only contain low cylindrical harmonic components. This means that actually a large number of angular frequency components have been acquired and processed that are not necessary at all for the final output.

25

The invention is based on the insight that this throwing away of high order angular frequency components, which was done in the last panning step of the processing, can just as well be done as a first step, since the order in which linear operations (multiplications) in the k_θ, ω -domain take place, does not effect the end result. In figures 5a-5e the wave field from figure 5a is first combined into a lower number of channels in figure 5b and then a 2D Fourier transform is calculated in figure 5c, containing only the low order cylindrical harmonics in figure 5d needed for the final virtual microphones in figure 5e. This approach is much more efficient in terms of

30

processing power than the approach from figures 4a-4f since only a limited number of 2D Fourier components need to be considered. When done correctly, it delivers the same output as the approach from figures 4a-4f. Furthermore, the panning done between figure 5a and 5b does not necessarily need to be done in the digital domain after acquiring the full 288 input channels; since it is a simple operation that can easily be done in the analog domain, a much lower number of acquisition channels is required. Note that this second processing approach implies that the shapes of the directivity patterns are equal for the different virtual microphones; the virtual microphones are rotated copies of each other.

10

If discrete microphones are used in the microphone array, the panning can be done by combining the outputs for the individual microphones using resistor ladder networks, as shown in figure 6a. The resistor ladder network comprises a plurality of resistors, the outputs of which being connected together to a summing circuit, e.g., an operational amplifier with feedback resistor, as shown in figure 6a. Each microphone $M(i)$ is connected to a corresponding resistor. The value of each resistor determines the weight of the particular microphone $M(i)$ on the summed output.

15

However, in an alternative embodiment, this resistor ladder network is not necessary and the pre-processor can be obviated. Instead of using small discrete microphone capsules with resistor ladder networks, it is also possible to use microphone capsules that are more extended in space such that the integration of the sound field over a larger area in space is done by the microphone itself. The weighting is in this case done by varying the height of the microphone capsules. See figure 6b for a simple schematic example of such a microphone capsule with a triangular weighting window. Each microphone $M(i)$ is then connected to an acquisition channel of the processor 3. In the arrangement shown in figure 6b, each of the outputs of the microphones $M(i)$ is connected to these acquisition channels via an operational amplifier with a feedback resistor.

20

25

30

Thus, in general terms, the invention is based on the insight that less acquisition channels of the processor 3 are needed when the system used performs a weighted integration of a sound field on predetermined portions of the circle where the

microphones $M(i)$ are located.

The shape of the weighting window plays an important role in the performance of such an array. First of all, it determines the precise shape of the directivity patterns for the virtual microphones. Furthermore, it determines the angular frequency content of the virtual directivity patterns, which determines the number of required acquisition channels and the angular aliasing performance, as will be shown next. In figures 7a-7d a number of weighting windows are shown: triangular, Hanning, Hamming, and Chebyshev, respectively. The sizes of the windows are chosen to obtain 12 uniformly distributed audio channels with full 360 degree coverage. Only the directivity window for the 0 degrees channel is shown in this figure. The windows for the other angles can be obtained by shifting the integration window over angles that are multiples of 30 degrees.

In figures 7a-7d the angular Fourier transforms of these windows are shown. All these windows contain approximately 24 non-zero frequency components. This implies that at least 24 samples in the angle (acquisition channels) are required for a sufficient aliasing-free wave field decomposition and virtual microphone output creation, which is twice the amount of desired output channels. If only 12 acquisition channels were used, severe angular aliasing artifacts would occur in the wave decomposition processing. Even with 24 samples the field is not entirely free from aliasing due to some frequency-lobing as can be seen from figures 7a-7d.

The effect of undersampling and frequency-lobing described above can be easily demonstrated by processing a simulated plane wave on the circular array using the processing scheme from figures 5a-5e. The performance of the full 288 channel processing using the processing scheme from figures 4a-4f is shown in figure 8a and figure 9a for angles of incidence of 0 and 15 degrees, respectively. This should be considered the best performance possible with the given array size. Cross talk between the 12 channels is extremely low, except for frequencies below 250 Hz, which, as discussed before, is due to the limited radius of the circular array. The case of using the scheme from figures 5a-5e using 24 triangular angles shows a relatively poor performance in figure 8b and figure 9b. This is due to the fact that the triangular

window has quite strong side lobes in the frequency domain which cause aliasing artifacts. In figure 8c-8g and 9c-9g 24 uniformly distributed Hanning, Hamming, Gaussian and the product of a Gaussian and a cosine window are used, some of which are giving a much better performance due to smaller side lobes. Only the odd 12 output channels are shown in these figures since only those are used for a discrete 12 channel reproduction system. See figure 10 for a sketch of the directivity patterns of all 12 used virtual microphones. Note that the shapes are only valid between 150 and 15000 Hz and that they will be distorted by aliasing caused by the frequency lobing.

Using all 24 output channels directly for a discrete 24 channel reproduction system is probably not a good idea, since the even channels have a strong overlap in angle with the odd ones implying unnecessary crosstalk between them that results in comb filter effects between the channels during reproduction. However, as will be shown later, it is possible to create 24 better virtual microphones with less crosstalk from the original ones using an output matrix.

Real-time wave decomposition processing

Until now only examples where shown using simple, short and easy to process wave fields, like a simulated Dirac plane wave recording. In practice it is necessary to decompose a continuous stream of 24 channels of audio in real time. This can be done by splitting the audio input stream into time blocks, applying a FFT2-based convolution filter in the k_θ, ω -domain, transforming the data back θ, ω -domain and using the overlap-add approach to combine the blocks into an output audio stream. This type of processing is certainly feasible in real time on a single pc nowadays. The FFT2-based convolution filter can be designed to account for both wave decomposition and compensation for the non-ideal actual performance of the cardioid microphones in terms of frequency response and frequency dependent directivity pattern.

Output Matrix

Having a discrete 12 channel microphone array is nice, however it will not always be flexible and desirable to use the 12 fixed virtual microphones from figure 10 directly. This would imply a different array design for each different reproduction system. For example, the 12 discrete output channels on itself cannot be used directly for a 5.1 or

7.1 reproduction. However, by choosing the 24 panned acquisition channels one is not limited to the 12 virtual microphones from figure 10 but can in fact use any linear combination of all 24 virtual microphones to create new virtual microphones. This can be implemented by using a simple output matrix after the operation of figure 5e. The 24 original channels are routed through a matrix to obtain any desired number of output channels. Suppose that $i_1 \dots i_{24}$ are the panned and decomposed output channels of the array and thus the input channels for the matrix and suppose $o_1 \dots o_n$ are the N desired output channels. Then one can write

$$\begin{pmatrix} o_1 \\ \vdots \\ o_N \end{pmatrix} = \begin{pmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,24} \\ M_{2,1} & M_{2,2} & \dots & M_{2,24} \\ \vdots & \ddots & \ddots & \vdots \\ M_{N,1} & M_{N,2} & \dots & M_{N,24} \end{pmatrix} \begin{pmatrix} i_1 \\ \vdots \\ i_{24} \end{pmatrix} \quad (0.5)$$

10

The coefficients $M_{n,m}$ of the matrix determine the directivity patterns for the N virtual microphones. In this case a frequency independent virtual microphone is considered. If one wants to create a frequency dependent directivity pattern, the output matrix coefficients should be functions of the temporal frequency ω instead of being constants. As a first trivial example for an output matrix one could consider the full 12 channel discrete output system. The matrix for this example is shown as a coloured map in figure 11. In this case the matrix is very simple: only the 12 odd input channels are used and sent directly to the 12 outputs.

15

20 Suppose now as a second, more interesting example of 5 discrete virtual microphones compatible with 5.1 reproduction angles and covering the full 360 degrees angle. If smooth transition between the left, right and surround channels is desired, the matrix could look like figure 12. This matrix was obtained by creating a band-limited (24 samples long) inversion filter for the used Hanning pre-panning window from figure 7b in the angular frequency domain and convolving this with the 24 sample band-limited version of the desired 5.1 directivity patterns, using multiplication in the angular Fourier domain. The directivity patterns for the desired and actually created virtual microphones resulting from this matrix are shown in figure 13.

25

30 If one desires a much sharper transition between the channels resulting in less overlap

and correlation between the front and the rear channels, the matrix from figure 14 can be used. The directivity patterns of the resulting virtual microphones in this are shown in figure 15. Notice that the obtained virtual microphones are not perfectly equal to the desired ones, since the desired ones contain a small amount of high frequency components that cannot be resolved.

These were only a few simple examples. Notice that any directivity pattern can be created in this way as long as it is not sharper than the directivity patterns of the original 24 microphone output channels. Clearly it is not possible to achieve really sharper directivity patterns, since for that purpose higher order cylindrical harmonics, that have been filtered out by the panning step in figure 5 are required. It is possible however to create slightly more directive beams. For example, the matrix from figure 16 delivers 24 discrete audio channels, but at the price of more sidelobes (crosstalk) and more noise sensitivity. Their directivity patterns are shown in figure 17.

The output matrix can be generalized even further by using complex matrix coefficients instead of only real ones. A complex coefficient implies, apart from the amplitude gain and sign, also a phase rotation in the audio signal. This could particularly be useful as a decorrelation/diffusion filter for the surround channels in the previous 5.1 examples, in case they are only used for the recording of ambiance (for example room effect and applause).

The output matrixing is a process that can be applied and optimized to the taste of the sound engineer afterwards, if during the live recording the full 24 channels were stored. This approach is comparable B-format recording, in which case the 4 B-format signals can also be combined (matrixed) into the final outputs afterwards by using a B-format processor [5].

Elevation angles

Elevated sources cannot be expected to be properly imaged by a circular array. A spherical array, being quite unfeasible in practice currently, should ideally be used for that purpose. The use of such an array can however only be justified if the reproduction system is also capable of reproduction of elevated sources, which is not the case for most reproduction systems, including WFS. It will be shown that although the

recording performance for elevated sources is not ideal, this should not be a major problem in practice. In figure 18 the content of the 12 output channels of a plane Dirac wave with various elevation angles simulated on the array system is shown. As can be concluded from this figure, an elevated source will have contributions from two angles,
5 namely the angle of the real source and the opposite angle. An elevated source in the front will thus be reproduced as a source in front and a source in the back. Furthermore, as can also be seen from figure 18, the angular resolution for an elevated source is more limited, which affects the low temporal frequencies. The energy ratio between front and back is determined by the elevation angle. For extreme elevation angles within 1-2
10 degrees from the vertical axis, all output channels at all angles are affected and quite a strong high frequency amplification occurs. However, since this is only a very small fraction of the full 4π space angle where normally no direct sound sources or early reflection mirror images are present, it can be expected not to have a strong effect on the overall array performance in real-life recording situations.

15

Noise performance of a circular array

Next, noise sensitivity and the effect of variations between individual microphone capsules on the overall array performance are investigated. In figure 19 the noise
20 spectrum of an individual cardioid microphone is shown. The microphones used produce a pink noise signal. The total theoretical noise spectrum for the array including processing and 12 channel reproduction at the sweet spot is shown in figure 19. The noise level is calibrated using a 1 kHz plane wave in such a way that the sound pressure level of the wave in the center of the microphone array is equal to the sound pressure
25 level of reproduced 1 kHz tone at the sweet spot of the 12 channel reproduction system. Figure 19 shows that the spectrum of the noise after the processing has become white. The signal to noise ratio for the whole array system improves drastically for low frequencies compared to the single microphone performance, however, for frequencies above 3 kHz the noise performance deteriorates. This results in an overall noise level
30 increase of 1-2 dB(A).

Variations between individual microphone capsules, which will mainly be level variations caused by the build in microphone pre-amplifiers and weighting resistors,

were also simulated. Fluctuations of 1-2 dB in microphone sensitivities were assumed. Although, as expected, the actual cross talk separation between the 12 channels becomes a bit worse in the case of these variations, the overall array performance of the array is not too sensitive to such variations and 1-2 dB variation is acceptable.

5 Building a prototype array

Currently a prototype circular array is being build. The array consists of 4 parts of pre-bended aluminium strips that can be jointed together easily to form a circle. Holes are drilled in the strips in which the microphone capsules are mounted. At the backside of the strips shielded multi-layer print boards are attached that contain connections to the
10 microphones, 4×288 weighting resistors, 24 preamplifiers and some other required electronic components. The boards are constructed as thin as possible and have large holes at the positions of the microphones to avoid acoustic shielding at the backside of the cardioids as much as possible. The microphone capsules used are type EM-135 from Primomic. These relatively cheap microphones should not be expected to deliver hi-end
15 studio quality in terms of noise and high frequency performance. However, the spatial quality of the system should be convincing enough to demonstrate the advantages of the proposed circular microphone array technology. The development and design of a microphone array using only 24 microphone capsules which are more extended in space is currently not taking place but could be a very interesting and relatively cheap
20 high performance solution for the future.

Conclusions

A circular array has been proposed that can deliver discrete multichannel audio by simulation virtual microphones in post processing with well controlled and sharp directivity patterns that cannot be achieved with conventional microphone technology.
25 Although such an array requires a large number of microphone capsules to avoid spatial aliasing, the number of recording channels doesn't have to be large at all since part of the processing can already be done in the analog domain using simple resistor ladder networks. Furthermore, if well shaped, in space extended microphones could be designed, such ladder networks would become unnecessary and aliasing could be
30 completely avoided. By using a complex output matrix almost any desired virtual microphone set can be created and auditioned in real time from the original recorded microphone output channels. Since the array under consideration is circular and not

spherical, elevation angles cannot be recorded and reproduced properly. In practical applications however this should not cause a big problem, even in case of strong ceiling reflections. The array system also is not very sensitive to noise and fluctuations in the microphone capsules and weighting resistors.

5 References

- [1] E. Hulsebos, D. de Vries and Emmanuelle Bourdillat, *Improved microphone array configurations for auralization of sound fields by Wave Field Synthesis*, preprint 5337 of 110th AES Convention, Amsterdam, Netherlands (2001)
- 10 [2] E. Hulsebos, D. de Vries and Emmanuelle Bourdillat, *Improved microphone array configurations for auralization of sound fields by Wave Field Synthesis*, J. Audio Eng. Soc., Volume 50, number 10, pp. 779-790, New York, USA (2002)
- [3] E. Hulsebos and D. de Vries, *Parameterization and reproduction of concert*
15 *hall acoustics measured with a circular microphone array*, preprint 5579 of 112th AES Convention, Munich, Germany (2002)
- [4] M. Abramowitz and I. A. Stegun, *Handbook of mathematic functions*, Dover Publications, New York, United States (1965)
- [5] K. Farrar, *Soundfield microphone*, Wireless world, october edition pp. 48--50
20 (1979)

Claims

1. Audio system with a circular microphone array comprising a predetermined first number of microphones arranged on a circle for receiving a sound field and a digital
5 signal processor for processing output signals from these microphones characterized in that said audio system is arranged to perform a weighted integration of said sound field on predetermined portions of said circle.
- 10 2. Audio system according to claim 1, the audio system comprising a pre-processor arranged:
- (a) to receive said output signals from said first number of microphones,
 - (b) to combine them to create a second number of virtual microphones to generate a second number of virtual microphone output signals and
 - 15 (c) to transmit these virtual microphone output signals to said digital signal processor, said second number being less than said first number.
3. Audio system according to claim 2, wherein said pre-processor is arranged:
- to use a predetermined panning mechanism in action (b) to create said second
20 number of virtual microphone output signals from said output signals, and said signal processor is arranged:
 - to calculate a 2D Fourier transform for said virtual microphone output signals,
 - to calculate cylindrical harmonic components from said 2D Fourier transform,
 - to calculate an inverse Fourier transform of said cylindrical harmonic
25 components.
4. Audio system according to claim 3, wherein said pre-processor is arranged to perform said panning mechanism in the analog domain.
- 30 5. Audio system according to claim 4, wherein said pre-processor comprises a resistor ladder network to perform said panning mechanism.

6. Audio system according to any of the claims 3 through 5, said processor being arranged to process said inverse Fourier transform by a matrix operation to generate processor output signals for another set of virtual microphones.
- 5 7. Audio system according to claim 1, wherein said microphones have microphone capsules with different heights and are extended in space such that integration of a sound field over a predetermined area in space is done by the microphones.
8. Audio system according to claim 7, wherein said microphone capsules are
10 selected from the following group: microphone capsules with a triangular, Hanning, Hamming, and Chebyshev weighting window.
9. Audio system according to any of the claims 7 and 8, said processor being arranged to process said output signals from said microphones by a matrix operation to
15 generate output signals for a set of virtual microphones.
10. Audio system according to any of the preceding claims, the system being arranged to:
- splitting said output signals into time blocks,
 - 20 • applying a FFT2-based convolution filter in k_o, ω domain to render filtered data,
 - transferring said filtered data back to θ, ω domain and
 - using an overlap-add approach to provide an output audio stream.
11. Audio system according to any of the preceding claims, the system comprising
25 microphones on a spherical array.
12. Method of processing output signals from a circular microphone array comprising a predetermined first number of microphones arranged on a circle for receiving a sound field characterized by performing a weighted integration of said sound field on
30 predetermined portions of said circle.
13. Method according to claim 12 characterized by:
- (a) receiving said output signals from said first number of microphones,

- (b) combining them to create a second number of virtual microphones to generate a second number of virtual microphone output signals and
- (c) transmitting these virtual microphone output signals to said digital signal processor, said second number being less than said first number.

5

14. Method according to claim 13 comprising:

- using a predetermined panning mechanism in said action (b) to create said second number of virtual microphone output signals from said output signals,
- to calculate a 2D Fourier transform for said virtual microphone output signals,
- 10 - to calculate cylindrical harmonic components from said 2D Fourier transform,
- to calculate an inverse Fourier transform of said cylindrical harmonic components.

15. Method according to claim 14, comprising processing said inverse Fourier
15 transform by a matrix operation to generate output signals for another set of virtual microphones.

16. Method according claim 12, comprising using microphones with microphone
20 capsules with different heights and being extended in space such that integration of a sound field over a predetermined area in space is done by the microphones.

17. method according to claim 16, comprising processing said output signals from
said microphones by a matrix operation to generate output signals for a set of virtual
microphones.

25

18. Computer program product comprising data and instructions to allow processing
actions (a), (b) and (c) as defined in claim 13.

19. Data carrier provided with a computer program product as claimed in claim 18.

Fig 1

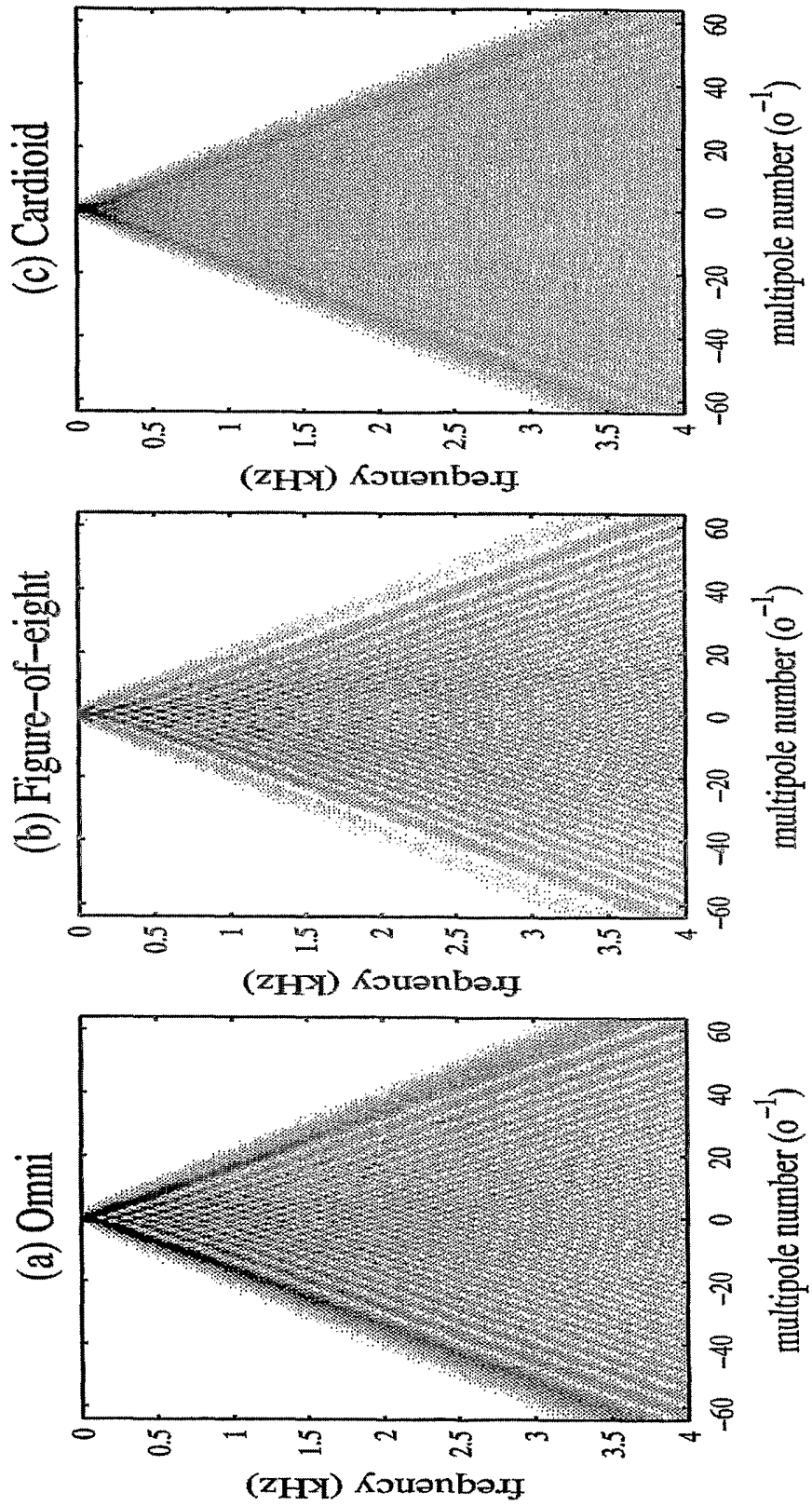


Fig 2

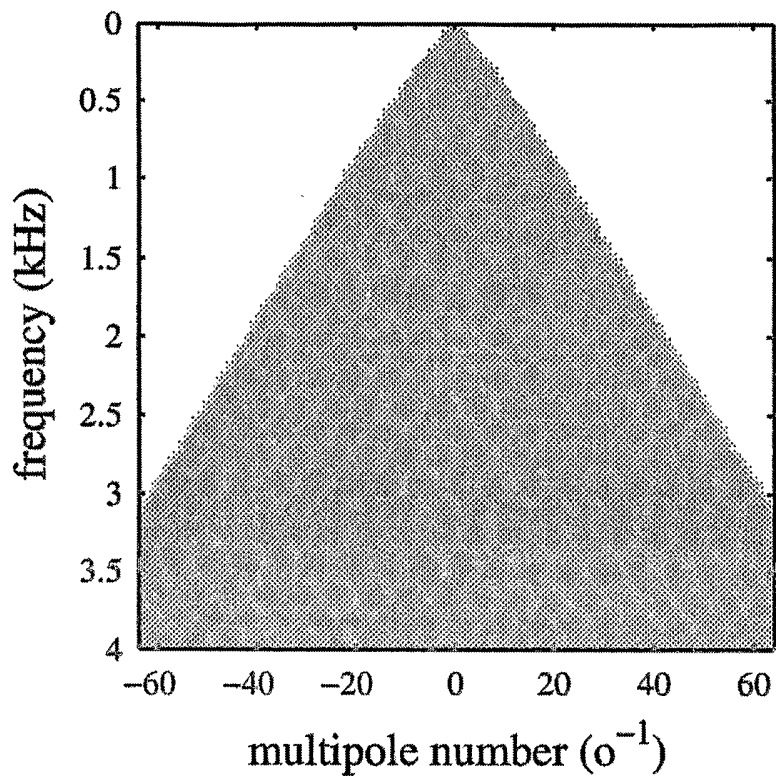


Fig 3

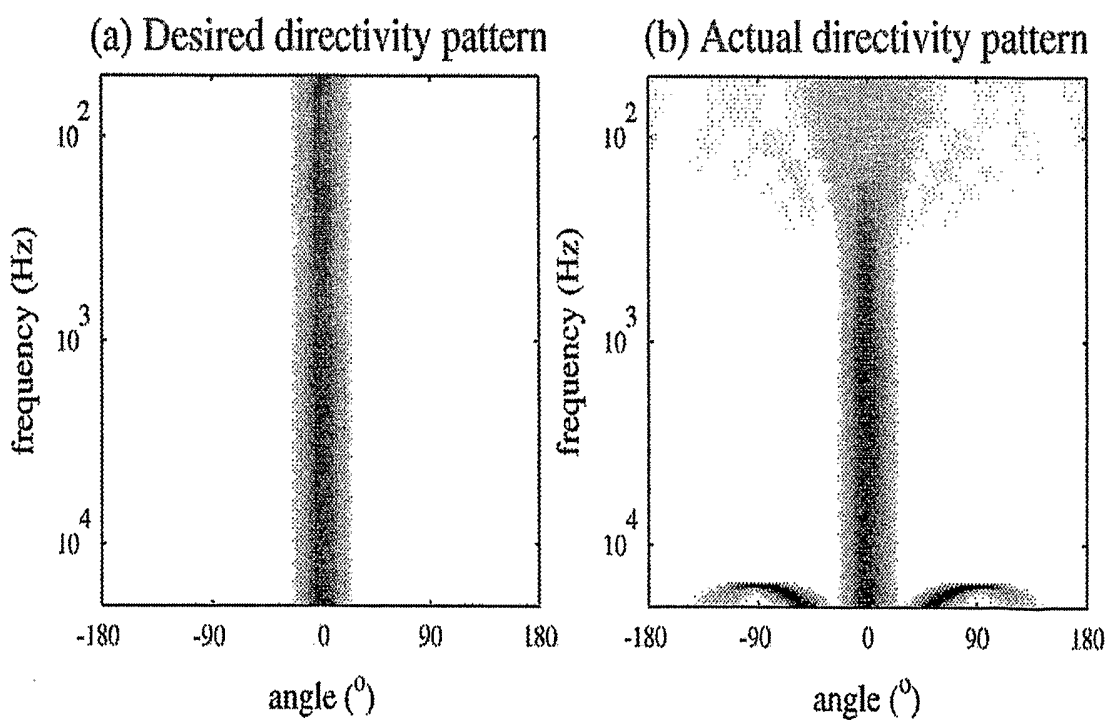


Fig 4

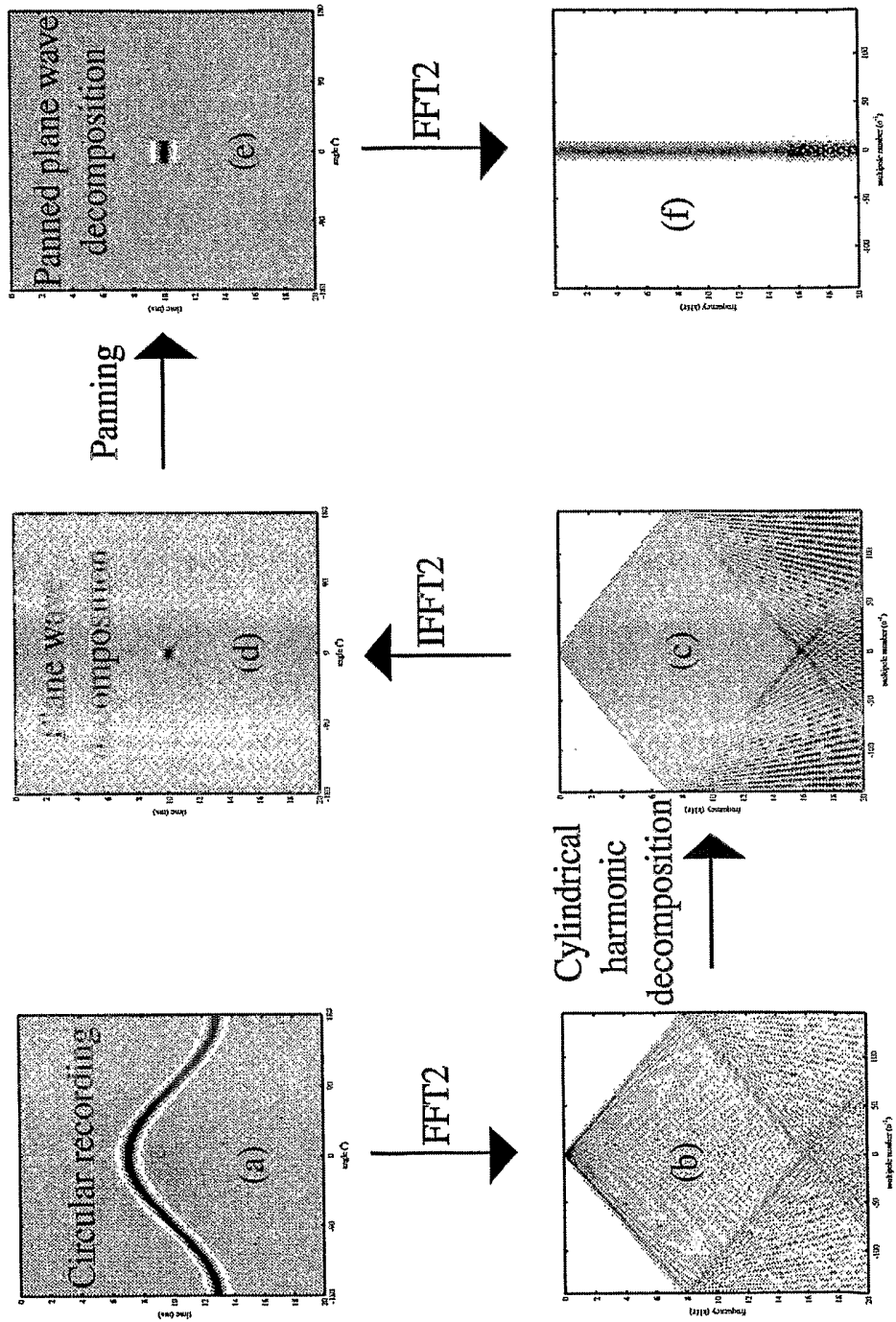


Fig 5

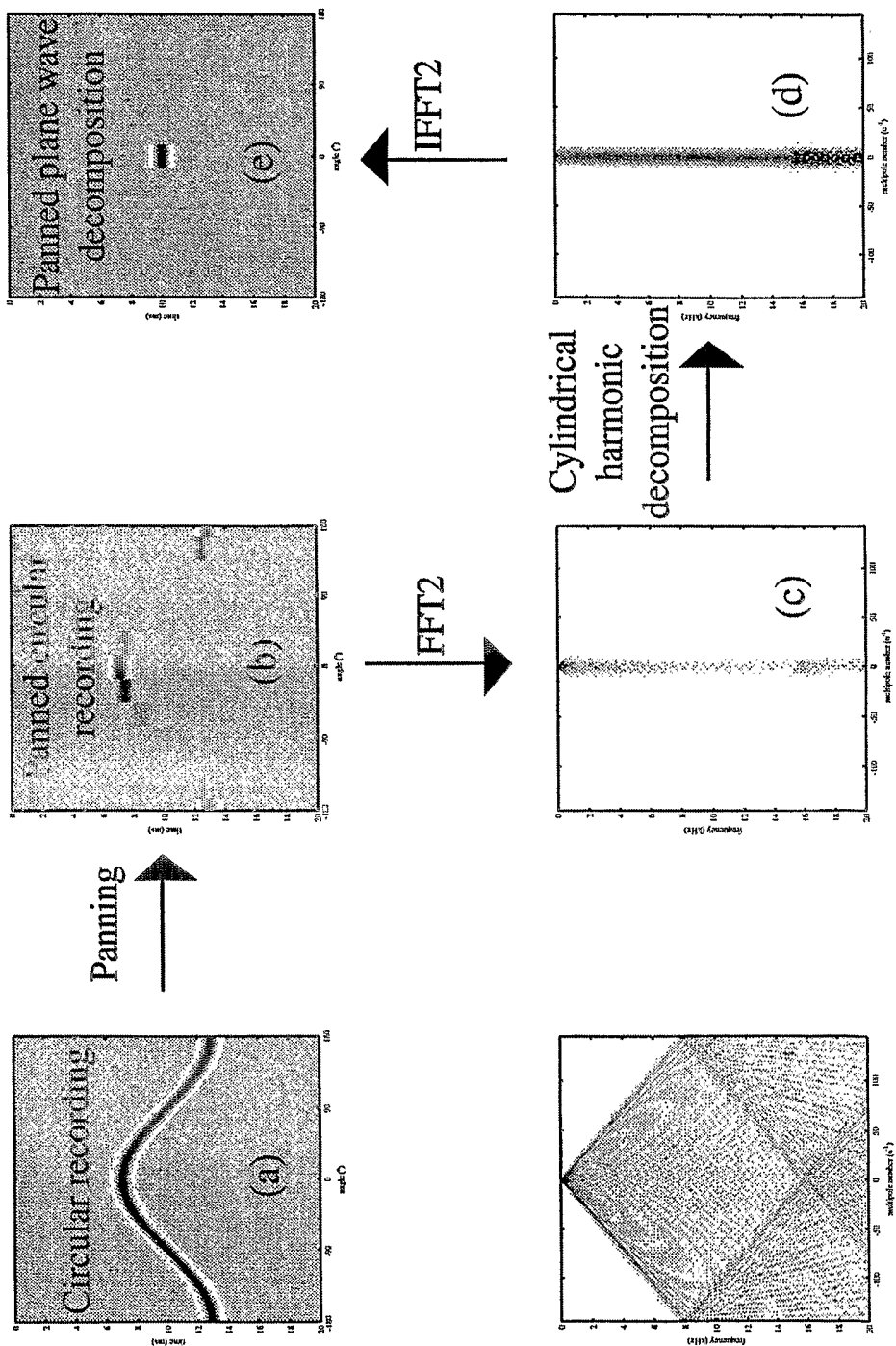


Fig 6a

(a) Small microphone capsules with weighting resistors

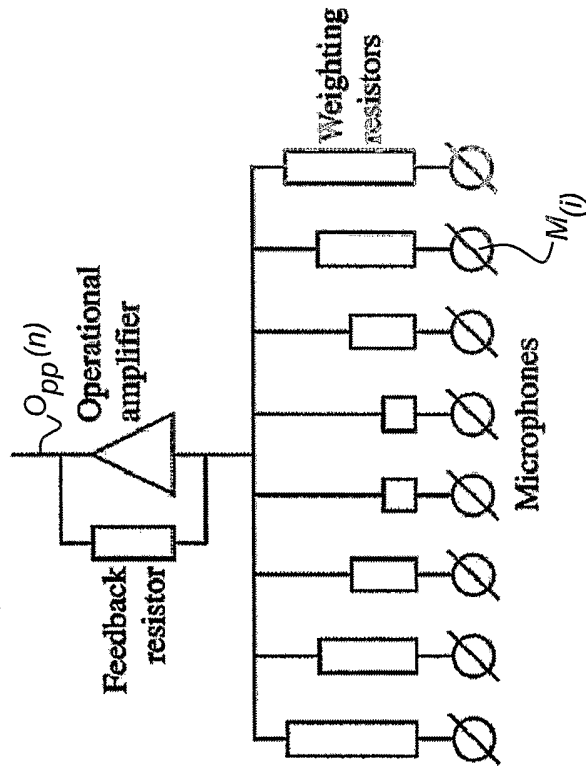


Fig 6b

(b) Large triangular microphone capsule

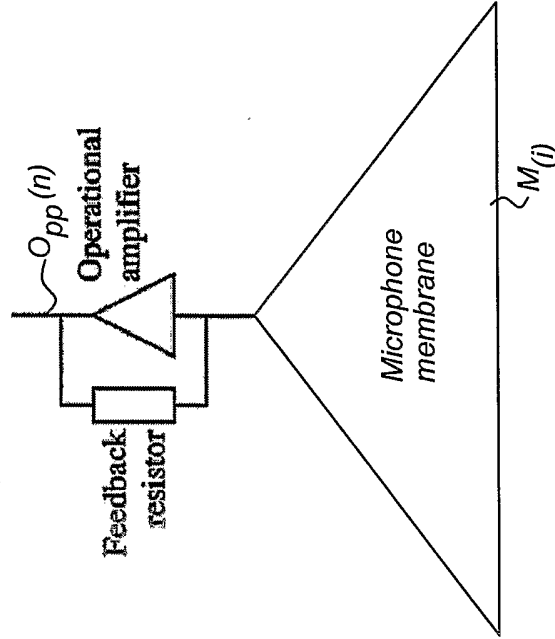


Fig 7(1)

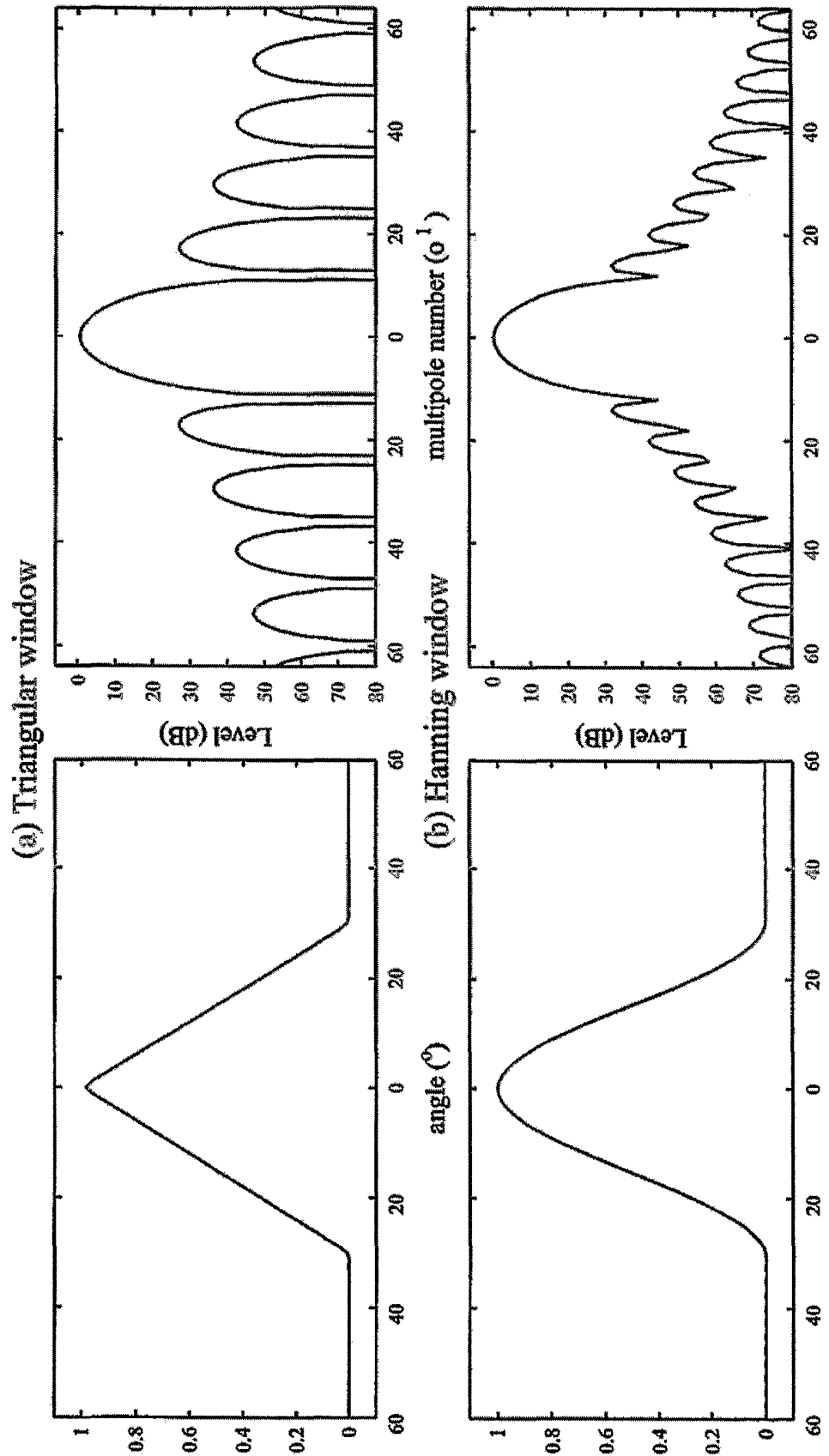


Fig 7(2)

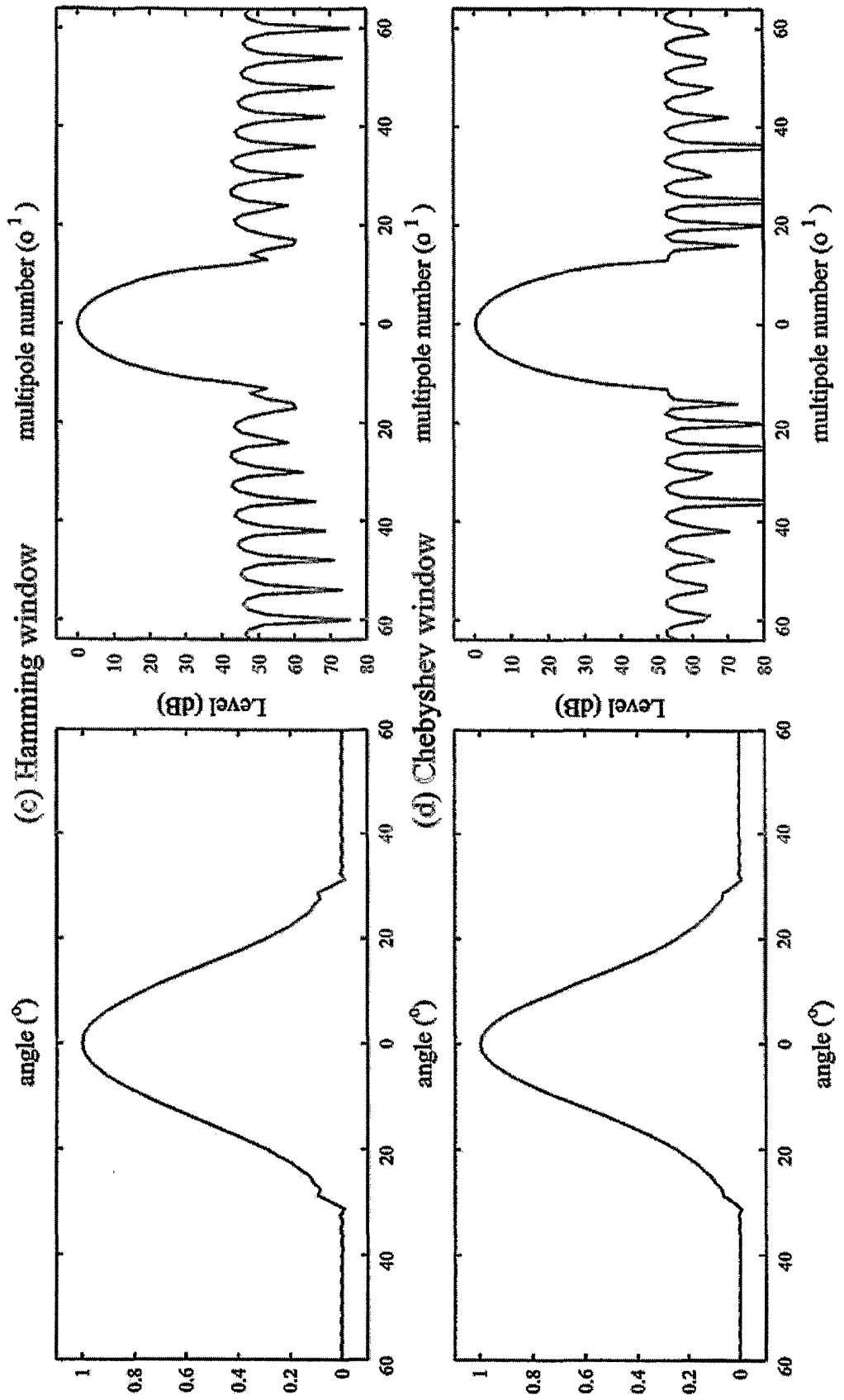


Fig 8(1)

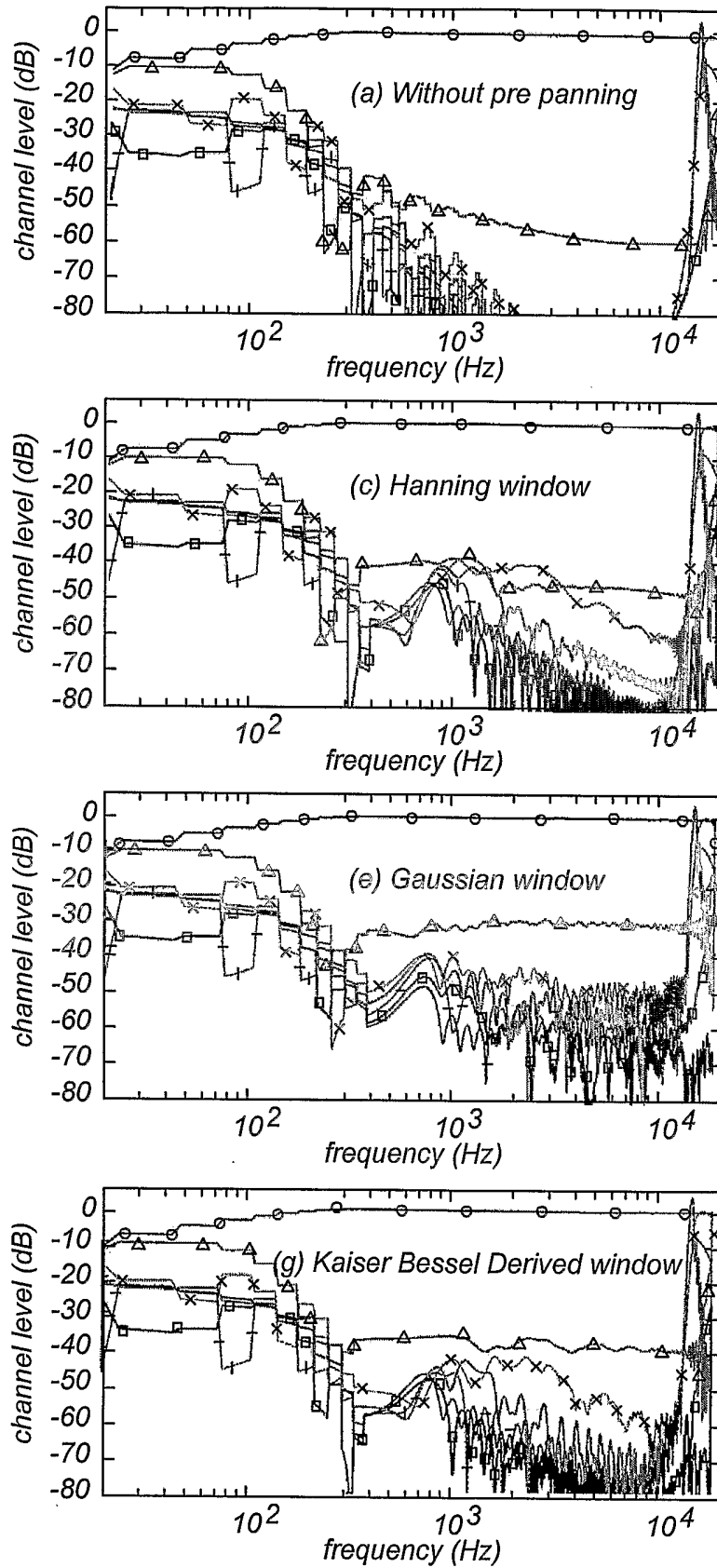


Fig 8(2)

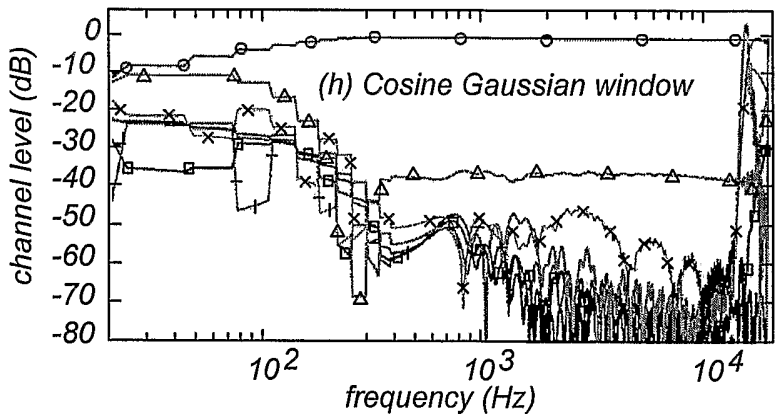
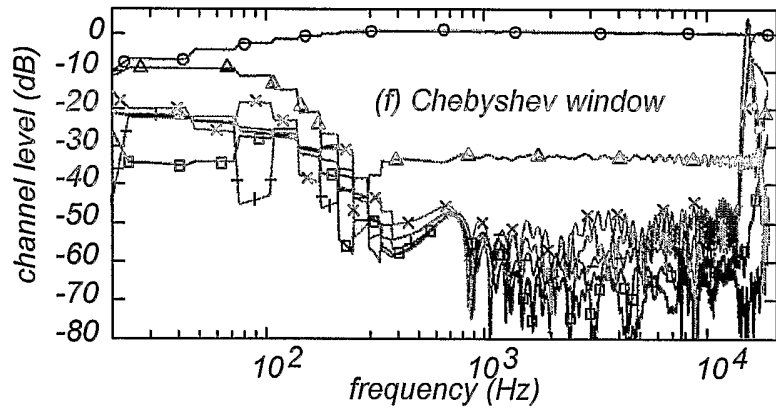
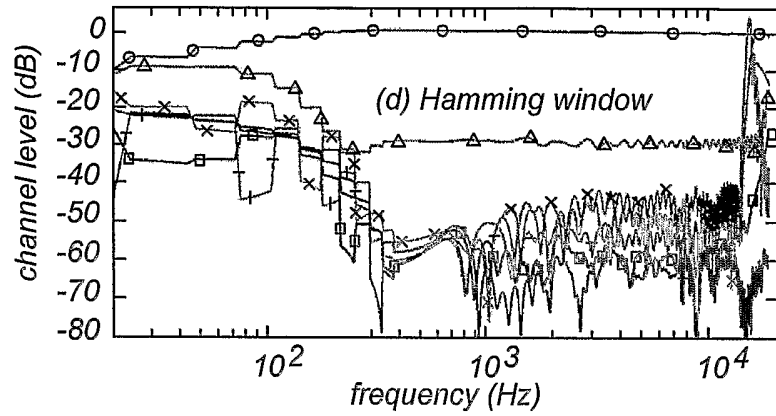
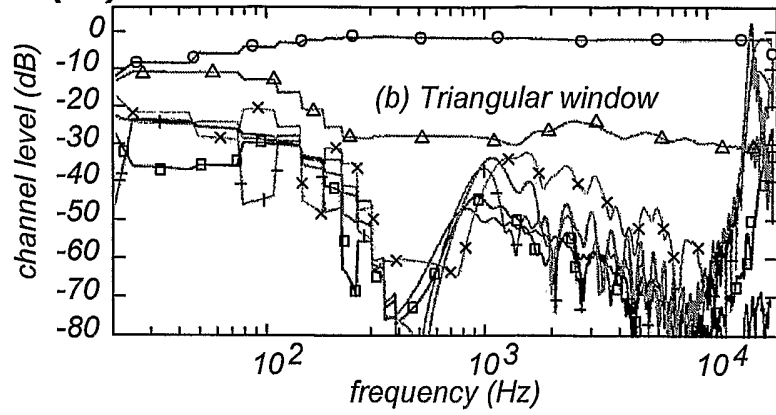


Fig 9(1)

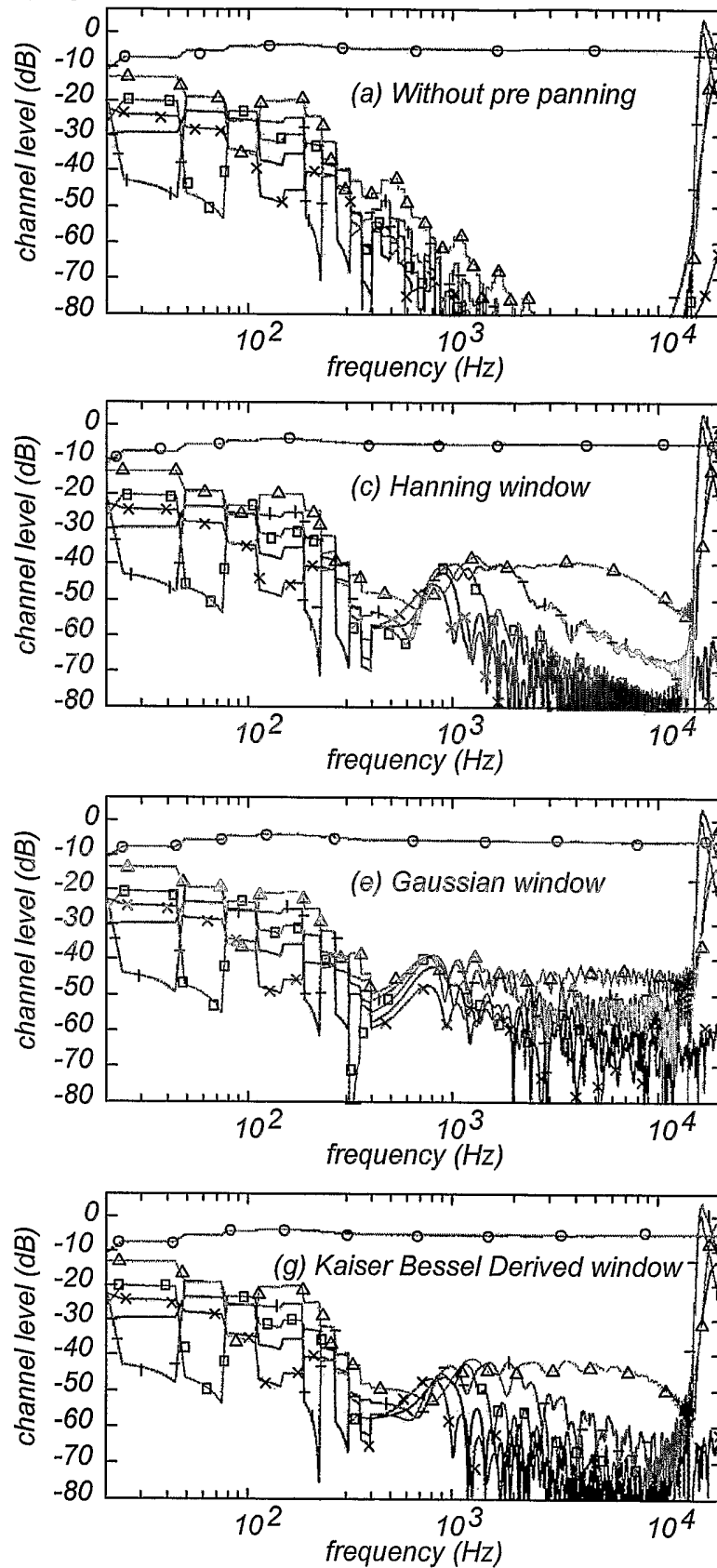


Fig 9(2)

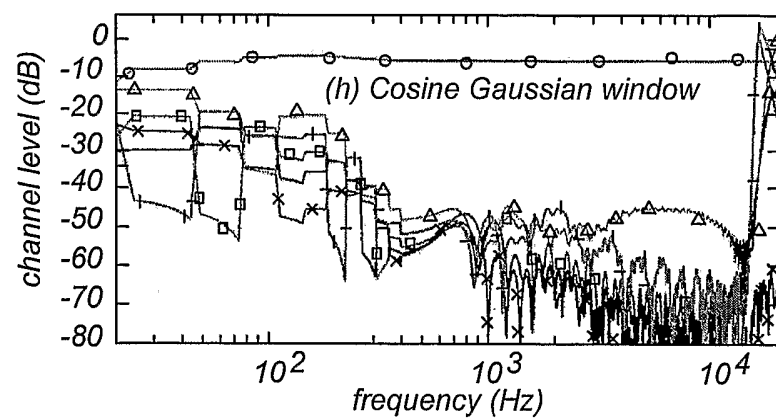
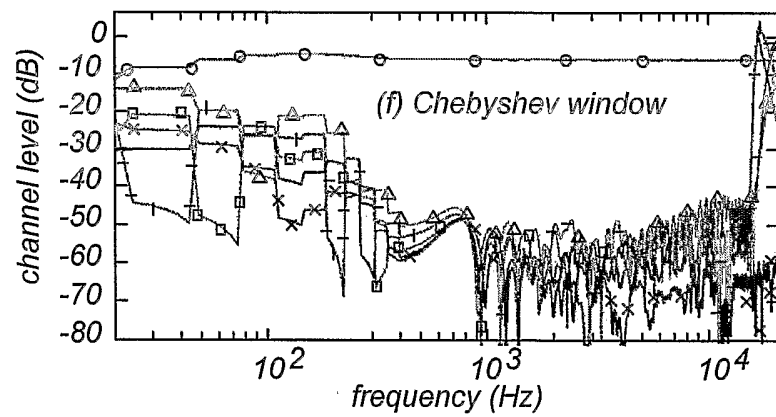
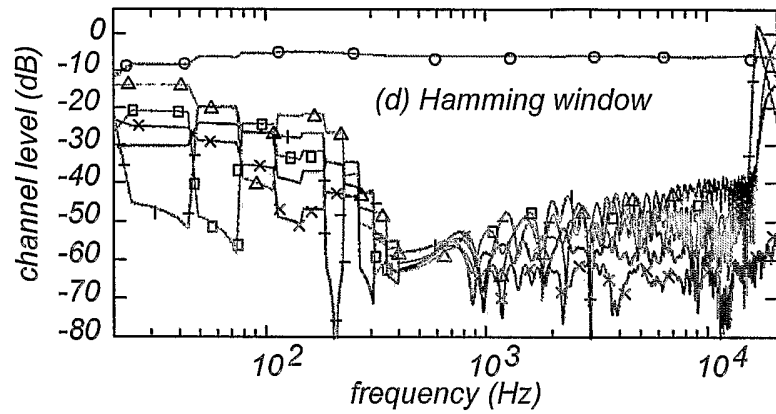
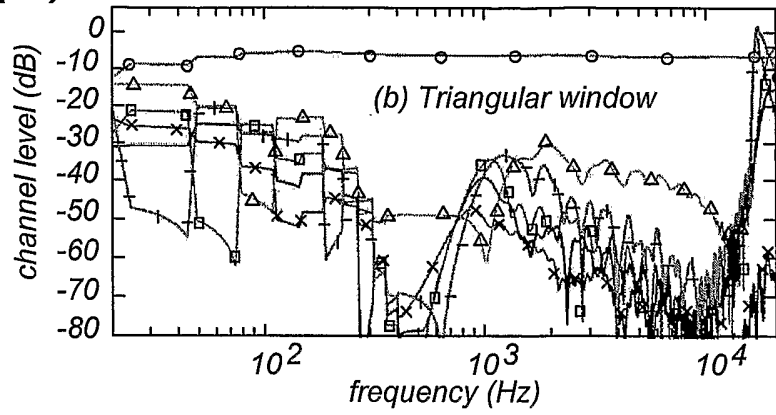


Fig 10

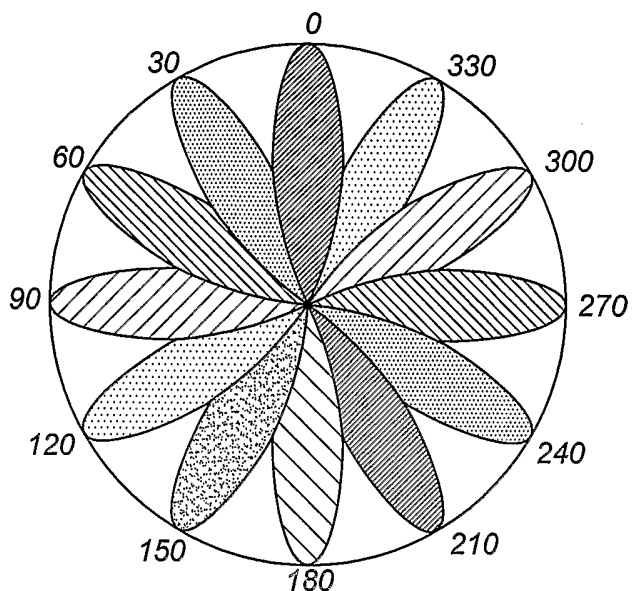


Fig 11

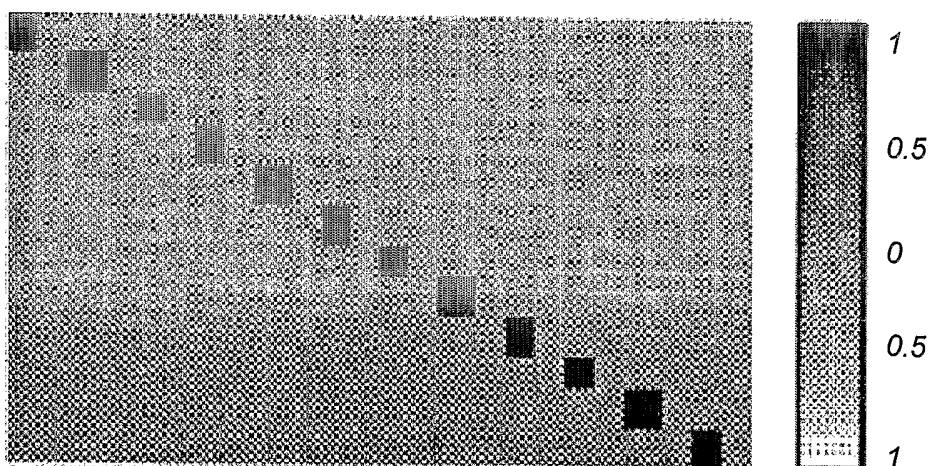


Fig 12

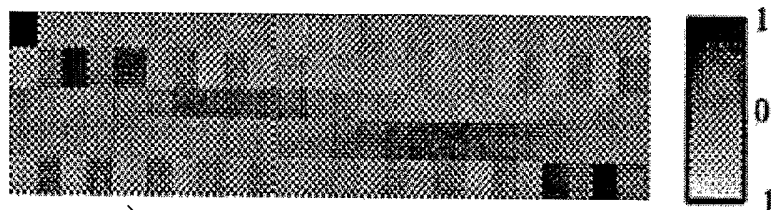


Fig 13

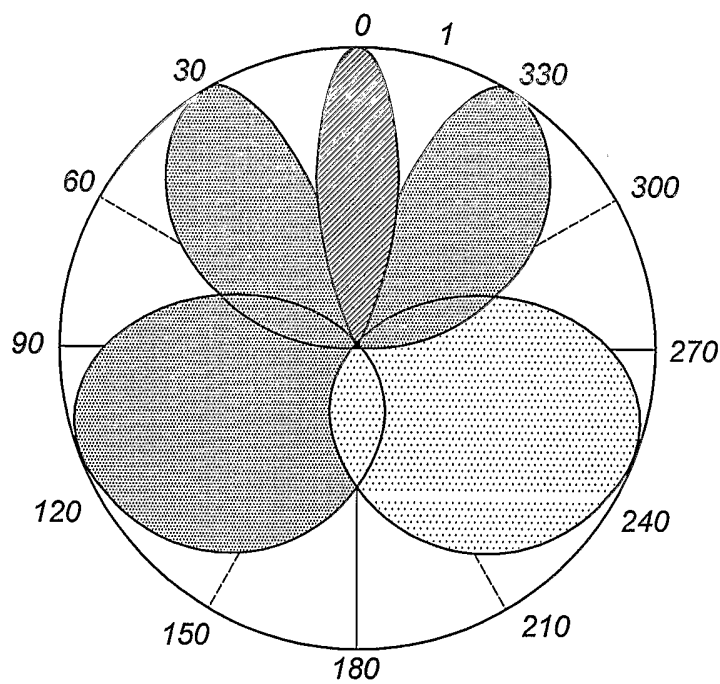


Fig 14

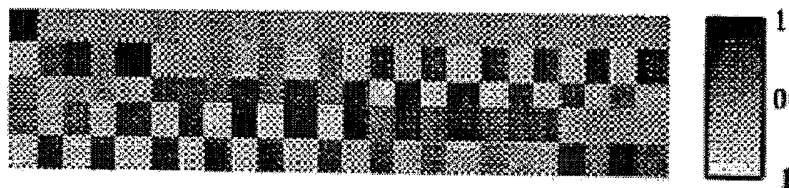


Fig 15

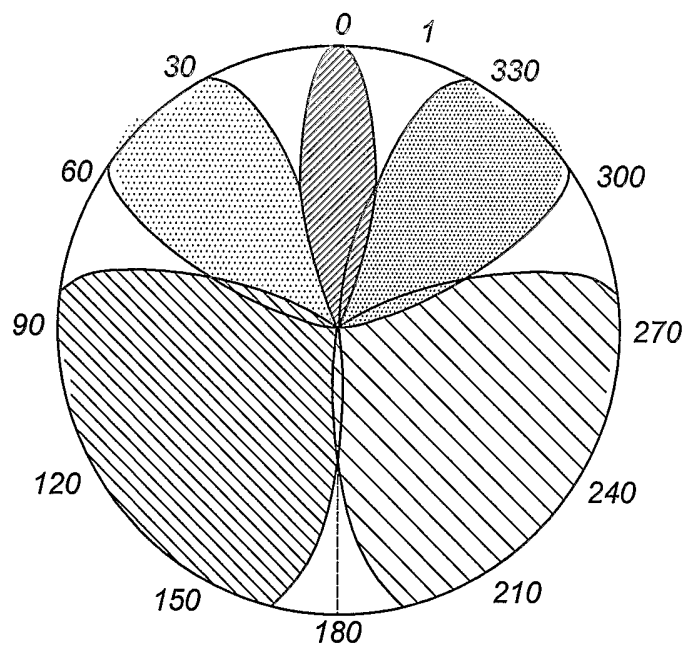


Fig 16

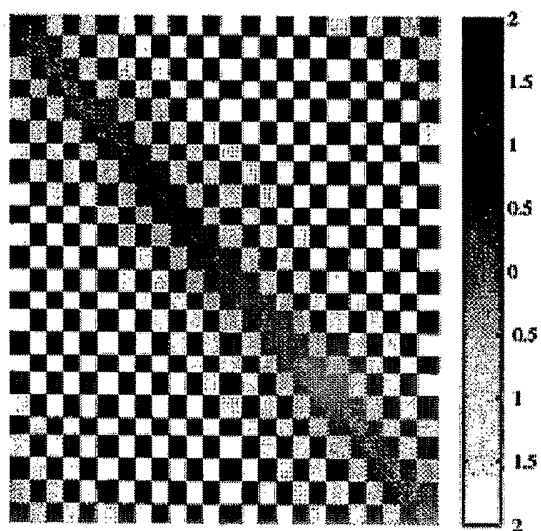


Fig 17

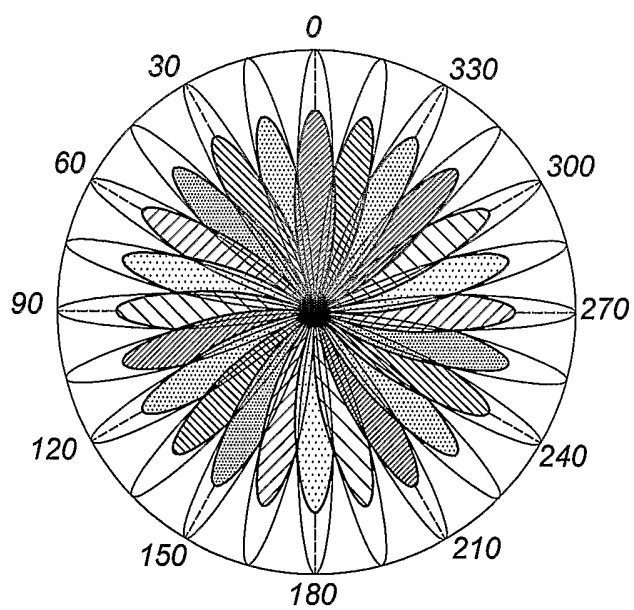


Fig 18

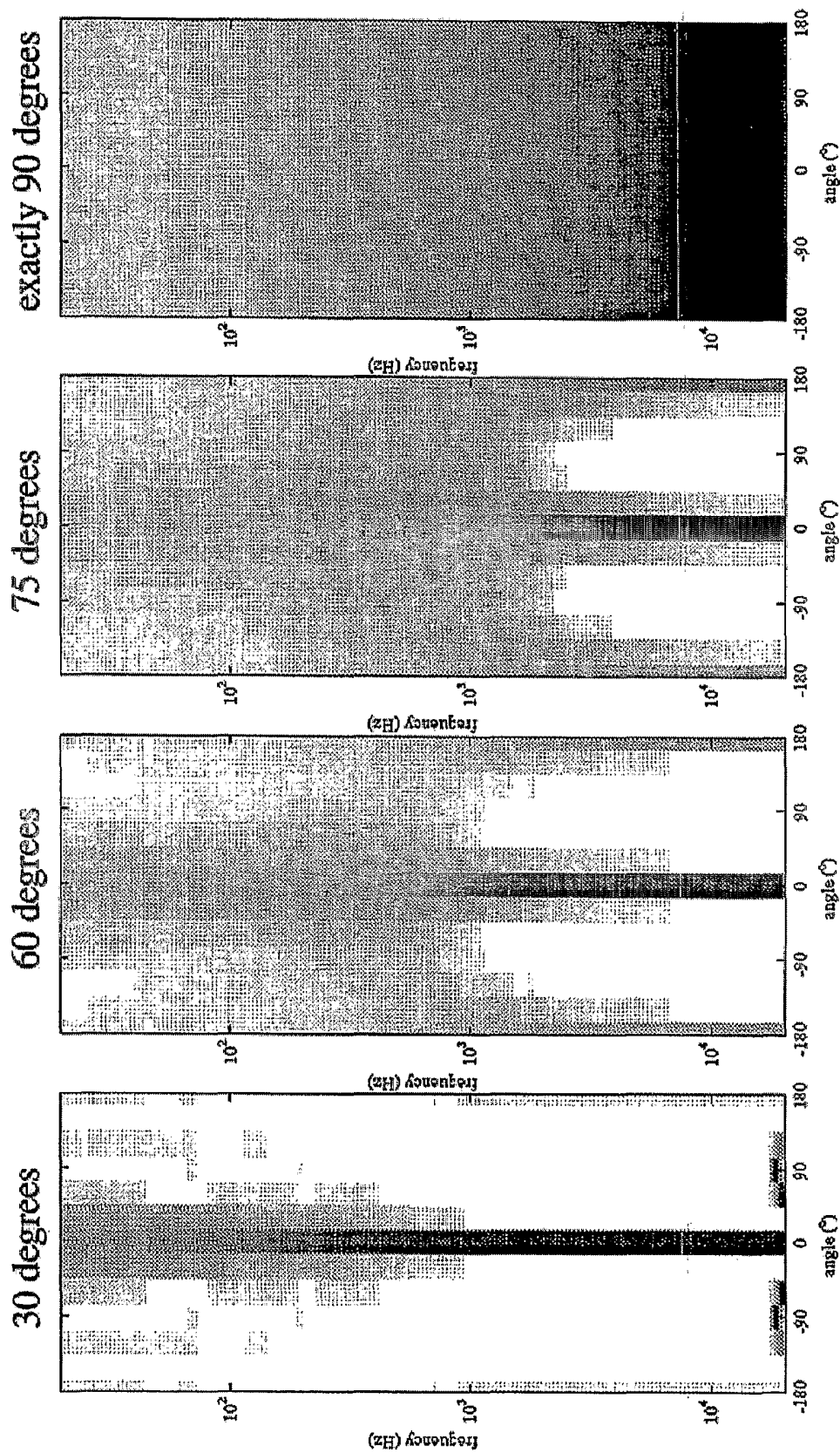


Fig 19

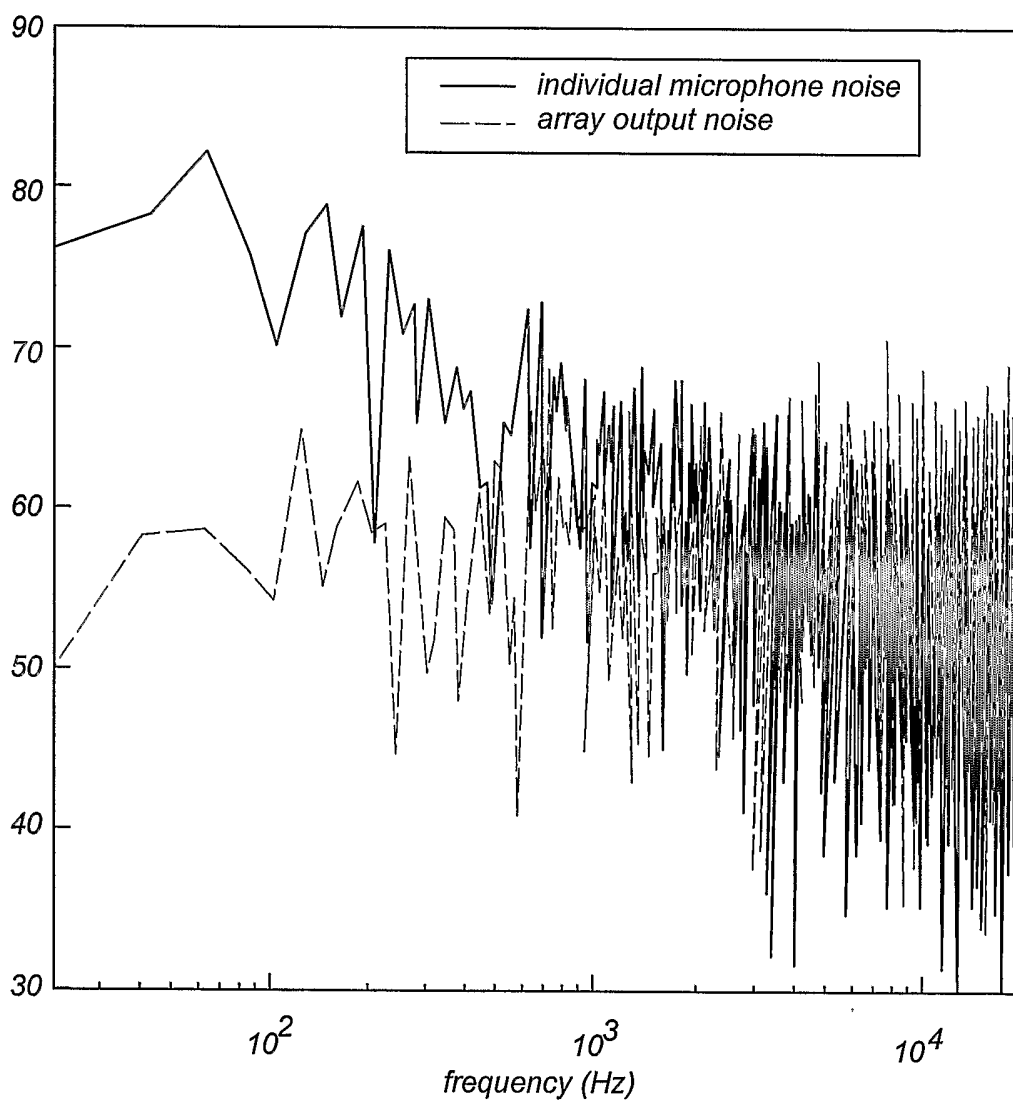


Fig 20

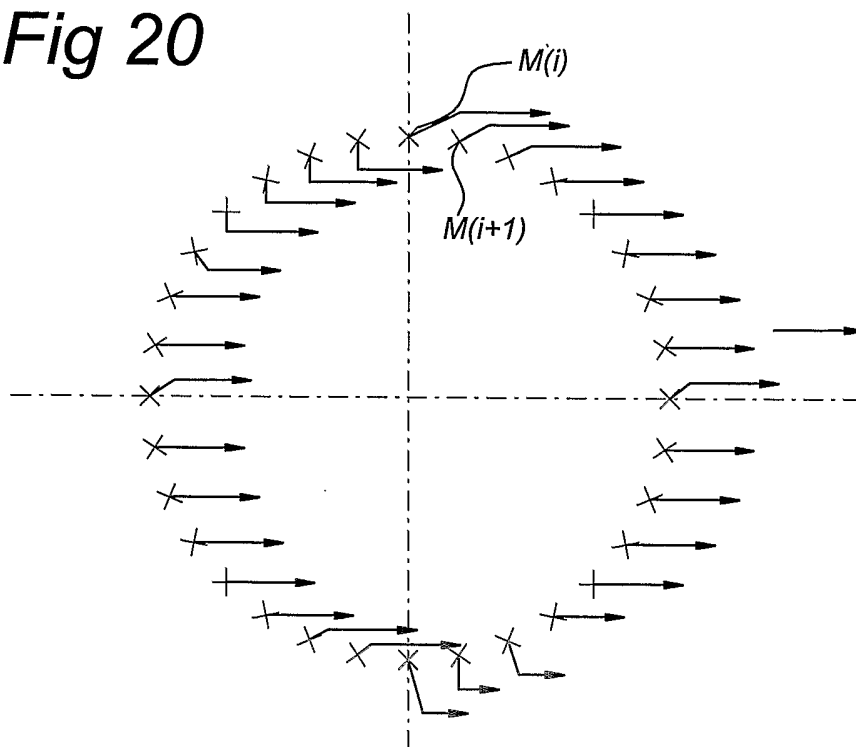
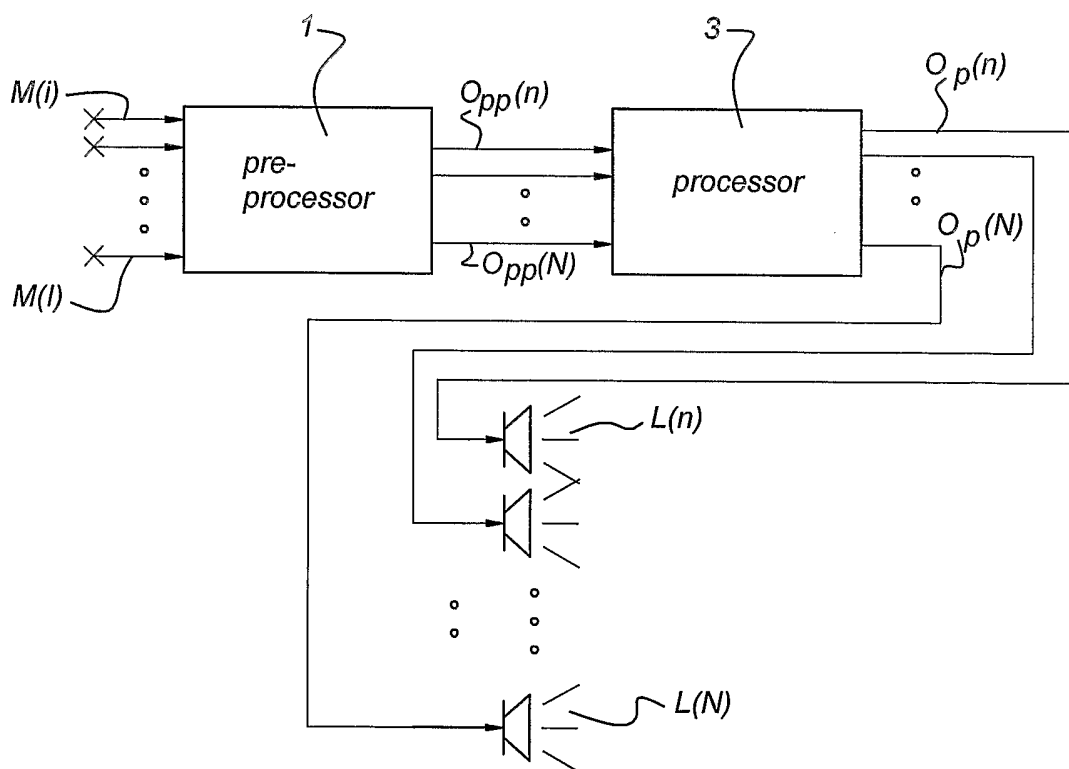


Fig 21



INTERNATIONAL SEARCH REPORT

PCT/NL2004/000169

A. CLASSIFICATION OF SUBJECT MATTER
 IPC 7 H04R1/40 H04R3/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 IPC 7 H04R

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, PAJ, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	US 5 715 319 A (CHU PETER L) 3 February 1998 (1998-02-03) the whole document	1-9, 11-19 10
A	US 6 192 134 B1 (ANDREWS JR WARNER B ET AL) 20 February 2001 (2001-02-20) column 5, line 33 - column 6, line 43	1-19
A	EP 0 414 264 A (SONY CORP) 27 February 1991 (1991-02-27) abstract; figure 3	3, 14
A	PATENT ABSTRACTS OF JAPAN vol. 008, no. 149 (E-255), 12 July 1984 (1984-07-12) & JP 59 052996 A (NIPPON DENSHIN DENWA KOSHA), 27 March 1984 (1984-03-27) abstract	5

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

° Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- * & * document member of the same patent family

Date of the actual completion of the international search

5 August 2004

Date of mailing of the international search report

17/08/2004

Name and mailing address of the ISA
 European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Fülöp, I

INTERNATIONAL SEARCH REPORT

PCT/NL2004/000169

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 03/009639 A (CARLILE SIMON ;JIN CRAIG (AU); LEUNG JOHANN (AU); VAN SCHAIK ANDRE) 30 January 2003 (2003-01-30) abstract; figure 1 -----	7,11,16

INTERNATIONAL SEARCH REPORT

Information on patent family members

PCT/NL2004/000169

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 5715319	A	03-02-1998	CA 2256485 A1	04-12-1997
			EP 0903056 A1	24-03-1999
			JP 3521914 B2	26-04-2004
			JP 2001502126 T	13-02-2001
			WO 9746048 A1	04-12-1997
US 6192134	B1	20-02-2001	WO 9927754 A1	03-06-1999
EP 0414264	A	27-02-1991	JP 2687613 B2	08-12-1997
			JP 3083490 A	09-04-1991
			DE 69022926 D1	16-11-1995
			DE 69022926 T2	30-05-1996
			EP 0414264 A2	27-02-1991
			US 5051964 A	24-09-1991
JP 59052996	A	27-03-1984	NONE	
WO 03009639	A	30-01-2003	WO 03009639 A1	30-01-2003
			EP 1433355 A1	30-06-2004