



(51) International Patent Classification:

H04R 3/00 (2006.01) *H04R 1/40* (2006.01)
H04R 3/02 (2006.01)

(21) International Application Number:

PCT/GB2010/000730

(22) International Filing Date:

9 April 2010 (09.04.2010)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

0906269.6 9 April 2009 (09.04.2009) GB

(71) Applicant (for all designated States except US): **NTNU TECHNOLOGY TRANSFER AS** [NO/NO]; Klæbuveien 153, N-7491 Trondheim (NO).

(71) Applicant (for MG only): **GARDINER, Stephen** [GB/GB]; Dehns, St Bride's House, 10 Salisbury Square, London, EC4Y 8JD (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **SUN, Haohai** [CN/NO]; Maristuveien 12, N-7030 Trondheim (NO). **YAN, Shefeng** [CN/CN]; No. 21 Beisihuan XiFu, 100190 Beijing (CN). **SVENSSON, U. Peter** [SE/NO]; Lassvebergvegen 13B, N-7562 Hundhamaren (NO).

(74) Agent: **DEHNS**; St. Bride's House, 10 Salisbury Square, London EC4Y 8JD (GB).

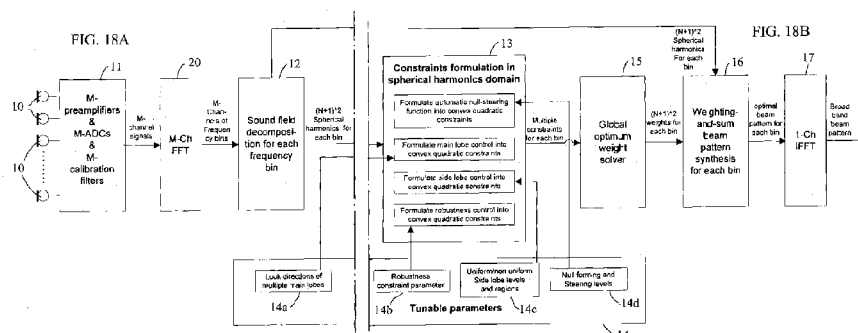
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, 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, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

(54) Title: OPTIMAL MODAL BEAMFORMER FOR SENSOR ARRAYS



(57) Abstract: A method of forming a beam pattern in a beamformer of the type in which the beamformer receives input signals from a sensor array, decomposes the input signals into the spherical harmonics domain, applies weighting coefficients to the spherical harmonics and combines them to form an output signal, wherein the weighting coefficients are optimized for a given set of input parameters by convex optimization. Formulations are provided for forming second order cone programming constraints for multiple main lobe generation, uniform and non-uniform side lobe control, automatic null steering, robustness and white noise gain.

- 1 -

Optimal Modal Beamformer for Sensor Arrays

The present invention relates to beamforming.

Beamforming is a technique for combining the inputs from several sensors in an array. Each sensor in the array generates a different signal depending on its location, these signals being representative of the overall scene. By combining these signals in different ways, e.g. by applying a different weighting factor or a different filter to each received signal, different aspects of the scene can be highlighted and/or suppressed. In particular, the directivity of the array can be changed by increasing the weights corresponding to a particular direction, thus making the array more sensitive in a chosen direction.

Beamforming can be applied to both electromagnetic waves and sound waves and has been used, for example, in radar and sonar. The sensor arrays can take on virtually any size or shape, depending on the application and the wavelengths involved. In simple applications, a one-dimensional linear array may suffice. For more complex applications, arrays in two or three dimensions may be required. Recently, beamforming has been used in the fields of 3-dimensional (3-D) sound reception, sound field analysis for room acoustics, voice pick up in video and teleconferencing, direction of arrival estimation and noise control applications. For these applications, arrays of microphones in three dimensions are required to allow a full 3-D acoustic analysis.

Of the possible three dimensional array arrangements, spherical arrays are of particular interest as more flexible three dimensional beam pattern synthesis can be realized than with other standard array geometries, and array processing can be performed using the mathematical framework of the spherical harmonics domain. A spherical array typically takes the form of a sphere with sensors distributed over its surface. The most common implementations include the "rigid sphere" in which the sensors are arranged on a physical sphere surface, and the "open sphere" in which the surface is only notional, but the sensors are held in position on this notional surface by other means. Other configurations such as dual open spheres (sensors arranged on two concentric notional spherical surfaces, one inside the other), spherical shell arrays (sensors arranged in between two concentric notional spherical surfaces, i.e. within the shell defined by them), single open spheres with Cardioid Microphones, and hemispheres are also suitable

- 2 -

implementations. All of these can be used for decomposition of the sound field into spherical harmonics.

For a given array (of e.g. microphones or hydrophones for acoustic applications or antennas for radio applications), the weights applied to each of the sensors in the array define a "beam pattern" for the array. However, typically, when one or more parts of the array are weighted more heavily than others, the beam pattern develops "lobes" which indicate areas of strong reception and good signal gain and "nulls" which indicate areas of weak reception where incident waves will be highly attenuated. The arrangement of lobes and nulls depends both on the weights applied to the sensors and to the physical arrangement of the sensors. However, typically, the beam pattern will include a "main" lobe for the strongest signal receiving direction (i.e. the principle maximum of the pattern) and one or more "side" lobes for the secondary (and other order) maxima of the pattern. Nulls are formed between the lobes.

In acoustic applications, considering the analysis of an auditory scene, the problem can be likened to the cocktail party problem in which it is desired to listen to a particular source (e.g. a friend who is talking to you), while ignoring or blocking out sounds from particular interfering sources (e.g. another conversation going on next to you). At the same time, it is also desirable to ignore or block out the background noise of the party in general. Similarly, the beamforming problem in a microphone array is to focus the receiving power of the array onto the desired source(s) while minimising the influence of the interfering sources and the background noise.

These problems can be of particular importance in applications such as teleconferencing in which two rooms are communicatively linked via microphone arrays and loudspeakers, i.e. each room has a microphone array to pick up sounds for transmission as audio signals to the other room and loudspeakers to convert signals received from the other room into sound. At any given time in one of the rooms (the near end), there may be one or more speaking persons whose voices must be captured, interference sources which should ideally be blocked, such as the loudspeakers which generate the sound from the other side of the call (the far end) and background noise e.g. air conditioning noises or echoes and reverberation due to the speaking persons and/or the loudspeakers.

This problem is generally addressed by the process known as "beamsteering" in which the main lobe of the beam pattern is aimed in the direction of the signal of interest, while nulls in the beam

- 3 -

pattern (also known as notches) are steered towards the direction(s) of interference signal(s) ("null steering").

The side lobes generally represent regions of the beampattern which receive a stronger than desired signal, i.e. they are unwanted local maxima of the beampattern. Side lobes are unavoidable, but by suitable choice of the weighting coefficients, the size of the side lobes can be controlled.

It is also possible to create multiple main lobes in the beampattern when there is more than one signal direction of interest. Other aspects of the beampattern which it is desirable to control are the beamwidth of the main lobe(s), robustness, i.e. the ability of the system to stand up to abnormal or unexpected inputs, and array signal gain (i.e. the gain in signal-to-noise ratio (SNR)).

In most environments, the auditory scene is constantly changing. Signals of interest come and go, signals from interference sources come and go, signals can change direction and amplitude noise levels can increase. In these situations, the sensor array ideally needs to be able to adapt to the changing circumstances, for example, it may need to move the mainlobe of the beampattern to follow a moving signal of interest, or it may need to generate a new null to counteract a new source of interference. Similarly, if a source of interference disappears, the constraints of the system are altered and a better optimal solution may be possible. Therefore, in these circumstances the array needs to be adaptive, i.e. it needs to be able to re-evaluate the constraints and to re-solve the optimization problem to find a new optimal solution. Further, in circumstances where the auditory scene changes rapidly, such as teleconferencing, the beamformer ideally needs to operate in real time; with people starting and stopping speaking all the time, sources of interest and sources of interference are constantly changing in number and direction.

A number of studies have been conducted in this field. To give a few examples, Meyer and Elko [J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in Proc. ICASSP, vol. 2, May 2002, pp. 1781-1784] presented the application and analysis of sound field spherical harmonics decomposition in a spherical microphone array beampattern design, which is symmetric around the look direction, and steerable in 3-D space without changing the shape of the beampattern. See also

- 4 -

WO2006/110230. As an extension to these studies, Rafaely [B. Rafaely, "Phase-mode versus delay-and-sum spherical microphone array processing," IEEE Signal Process. Lett., vol. 12, no. 10, pp. 713-716, Oct.2005] applied the commonly used delay-and-sum beampattern design method to a spherical microphone array, that is, applying array weights and compensating for the delays at the free field microphones due to a single plane wave. This approach results in high robustness, but at the cost of decreased directivity at lower frequencies. In another study, Rafaely et al also achieved sidelobe control for a given mainlobe width and array order, using a classical Dolph-Chebyshev pattern design approach, to improve the directional analysis of a sound field [B. Rafaely, A. Koretz, R. Winik, and M. Agmon, "Spherical microphone array beampattern design for improved room acoustics analysis," in Proceedings of the International Symposium on Room Acoustics, Sept. 2007, p. S42]. By imposing a white noise gain (WNG) constraint into beampattern synthesis, Li and Duraswami [Z. Y. Li and R. Duraishwami, "Flexible and optimal design of spherical microphone arrays for beamforming," IEEE Trans. Audio Speech Lang. Process., vol. 15, no. 2, pp. 702-714, Feb.2007], presented array weights optimization methods to find the balance between beamforming directivity and robustness, which is useful in practical applications. While the studies mentioned above considered only symmetrical beam patterns, Rafaely [B. Rafaely, "Spherical microphone array with multiple nulls for analysis of directional room impulse responses," in Proc. ICASSP, Apr. 2008, pp. 281-284] extended the beampattern design methods to non-symmetric cases for a spherical microphone array. This approach was formulated in both the space domain and the spherical harmonics domains, and included a multiple null-steering method, in which fixed nulls in the beampattern were formed and steered to the interferences coming from known outside beam directions, in order to achieve better signal to noise ratio.

In "Modal Analysis Based Beamforming for Nearfield or Farfield Speaker Localization in Robotics", Argentieri et al, *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp 866-871, convex optimization techniques were employed and a spherical harmonics framework was used to analyse the problem, but the wavefield was not decomposed into spherical harmonics.

In the above studies of spherical harmonics domain beamforming however, multiple deep nulls in the beampatterns could not be adaptively formed and steered to suppress the dynamic interferences coming from arbitrary outside beam directions. Such interference suppression is often desired in speech enhancement and multiple-channel acoustic echo cancellation for video

- 5 -

or teleconference applications, and analysis for directional room impulse response (i.e. acoustic analysis of a room through impulse generation and reflection analysis). Additionally, the above studies were unable to effectively include multiple beamforming performance parameters, such as sidelobe control and robustness constraints into a single optimization algorithm, so it has not so far been possible to obtain the global optimum solution for all of these mutually correlated parameters.

The main difficulty is that optimization algorithms are computationally intensive. As the applications described above, e.g. teleconferencing, are consumer applications, the algorithm must be executable with readily available consumer computing power in a reasonable time. It must also be noted that these applications are based in real time and need to be adaptive in real time. It is therefore very difficult to optimize all of the desired parameters, while maintaining real time operation. The requirements for real time operation can vary depending upon the application of the array. However, in voice pick up applications like teleconferencing, the array has to be able to adapt at the same rate as the dynamics of the auditory scene change. As people tend to speak for periods of several seconds at a time, a beamformer which takes a few seconds (up to about 5 seconds) to re-optimize the beampattern is useful. However, it is preferred that the system be able to re-optimize the beampattern (i.e. recalculate the optimum weightings) in a time scale of the order of a second so as not to miss anything which has been said. Most preferably, the system should be able to re-optimize the weightings several times per second so that as soon as a new signal source (such as a new speaker) is detected, the beamformer ensures that an appropriate array gain is provided in that direction.

It should be noted that, as computing power is still increasing exponentially according to Moores' Law, advances in computing power will rapidly decrease the amount of time to perform the necessary calculations and in the future it is expected that real time applications will be carried out with a significantly increased rate of re-optimizing.

As there are several parameters which affect the choice of beam pattern in a given scenario, an optimal solution for one of these parameters will not necessarily be optimal for the others. Therefore a compromise has to be made between them. Finding the best (optimal) compromise between these factors depends on the requirements of the system. These can be formulated as constraints upon the optimization problem. For example, one might require the system to have a certain directivity or a gain above a chosen threshold level. Alternatively, one might require the

- 6 -

sidelobe levels to be below a certain threshold or one might require that the system has a certain robustness. As discussed above, optimization is a computationally intensive process, and it becomes increasingly more intensive with every constraint added. Therefore, in practice it is normally unfeasible to apply more than a single constraint to the system if the optimal solution is to be found in a reasonable time.

In the studies performed so far, optimization algorithms have been limited to only one or two constraints. In some cases, the constraints have each been solved separately, one by one in individual stages, but it has not been possible to obtain a global optimum solution.

There remains a need to provide a method of finding a global optimum beam pattern for a spherical array while applying multiple constraints to the system.

According to a first aspect of the invention, there is provided a method of forming a beam pattern in a beamformer of the type in which the beamformer receives input signals from a sensor array, decomposes the input signals into the spherical harmonics domain, applies weighting coefficients to the spherical harmonics and combines them to form an output signal, wherein the weighting coefficients are optimized for a given set of input parameters by convex optimization.

By expressing the objective function and the constraints as convex functions, it becomes possible to apply the techniques of convex optimization. Convex optimization has the benefits of guaranteeing that a global minimum will be found if it exists, and that it can be found fast and efficiently using numerical methods.

In previous studies, in order to easily form a regular or irregular, and frequency independent beam pattern, array weight design approaches have always utilized the inversion of the mode amplitudes $b_n(ka)$ (discussed in more detail later) in the spherical harmonics domain to decouple frequency-dependent components. However, $b_n(ka)$ has small values at certain ka and n values, and its inversion may damage the robustness of the beamformer in practical implementations. In the present invention, by directly making the more general weights $w^*(k)$ the targets of the optimization framework, the optimization problem can be formulated as a convex optimization problem, i.e. one where the objective function and the constraints are all convex functions. The advantages of convex optimization, as discussed above, are that there are

- 7 -

fast (i.e. computationally tractable) numerical solvers which can rapidly find the optimum values of the optimization variables. Further, as discussed above, convex optimization will always result in a global optimum solution rather than a local optimum solution. Thus, with the above formulation, the beamformer of the invention can adaptively optimize the array beampattern in real time even with the application of multiple constraints.

The technique of convex optimization has been known for a long time. Various numerical methods and software tools for solving convex optimization problems have also been known for some time. However, convex optimization can only be used when the objective function and the optimization constraints are all convex functions, that is a function f is convex if

$f(ax + by) \leq af(x) + bf(y)$ for all x, y , and all a, b , with $a + b = 1$, $a \geq 0$ and $b \geq 0$. It is therefore not always possible to solve a given optimization problem using convex optimization techniques. First, the problem has to be formulated in a manner in which convex optimization can be applied. In other words, one has to take a property of the system which it is desired to minimise and formulate it as a convex function. Further all the constraints on the optimization problem must be formulated as either convex equalities / inequalities or linear equalities. By formulating the beamforming problem as a convex optimization problem, the present invention permits the use of a number of extremely efficient algorithms which make real time solution of multi-constraint beamforming problems computationally tractable.

Preferably, the sensor array is a spherical array in which the sensors' positions are located on a notional spherical surface. The symmetry of such an arrangement leads to simpler processing.

A number of different spherical sensor array arrangements may be used with this invention.

Preferably, the sensor array is of a form selected from the group of: an open sphere array, a rigid sphere array, a hemisphere array, a dual open sphere array, a spherical shell array, and a single open sphere array with cardioid microphones.

The array size can vary a great deal depending on the applications and the wavelengths involved. However, for microphone arrays used in voice pick up applications, the sensor array preferably has a largest dimension between about 8 cm and about 30 cm. In the case of a spherical array, the largest dimension is the diameter. A larger sphere has the benefit of handling low frequencies well, but to avoid spatial aliasing for high frequencies, the distance between two microphones should be smaller than half the wavelength of the highest frequency. Therefore if

- 8 -

the microphone number is finite, the smaller sphere means a shorter distance between microphones and less spatial aliasing issue. It will be appreciated that in high frequency applications such as ultrasound imaging where frequencies of 5 to 100 MHz can be expected, the sensor array size will be significantly smaller. Similarly, in sonar applications, the array size may be significantly larger.

Preferably, the sensor array is an array of microphones. Microphone arrays can be used in numerous voice pick-up, teleconferencing and telepresence applications for isolating and selectively amplifying the voices of the different speakers from other interference noises and background noises. Although the examples described in this specification concern microphone arrays in the context of teleconferencing, it will be appreciated that the invention lies in the underlying technique of beamforming and is equally applicable in other audio fields such as music recording as well as in other fields such as sonar, e.g. underwater hydrophone arrays for location detection or communication, and radiofrequency applications such as radar with antennas for sensors.

In preferred embodiments, the optimization problem, and optionally also constraints, are formulated as one or more of: minimising the output power of the array, minimising the sidelobe level, minimising the distortion in the mainlobe region and maximising the white noise gain. One or more of these requirements can be selected as input parameters for the beamformer. Furthermore, any of the requirements can be formulated as the optimization problem. Any of the requirements can also be formulated as further constraints upon the optimization problem. For example, the problem can be formulated as minimising the output power of the array subject to minimising the sidelobe level or the problem can be formulated as minimising the sidelobe level subject to minimising the distortion in the mainlobe region. Several constraints may be applied if desired, depending upon the particular beamforming problem.

In some preferred embodiments, the optimization problem is formulated as minimising the output power of the array. This is the parameter which will be globally minimised subject to any constraints which are applied to the system. Thus, in the absence of constraints to the contrary in any given region (direction) of the beam pattern, the optimization algorithm aims to reduce the output power of the array gain in that region by reducing the array gain. This has the general benefit of minimising the gain as much as possible in all regions except those where gain is desired.

Preferably the input parameters include a requirement that the array gain in a specified direction be maintained at a given level, so as to form a main lobe in the beam pattern. With the general tendency of the optimization algorithm to reduce gain as described above, a requirement that the gain be maintained at a given level in a specified direction ensures that a main lobe (i.e. a region of high gain and therefore signal amplification rather than signal attenuation) is present in the beam pattern.

More preferably, the input parameters include requirements that the array gain in a plurality of specified directions be maintained at a given level, so as to form multiple main lobes in the beam pattern. In other words, the directivity of the array is optimized by applying multiple constraints such that the gain of the array is maintained at a selected level in a plurality of directions. In this way multiple main lobes can be formed in the array's beam pattern and multiple source signal directions can be provided with higher gain than the remaining directions.

Yet more preferably, individual required gain levels are provided for each of the plurality of specified directions, so as to form multiple main lobes of different levels in the beam pattern. In other words, the optimization constraints are such as to apply different levels of signal maintenance (i.e. array gain) in different directions. For example, the array gain can be maintained at a higher or lower level in one direction than in other directions. In this way the beamformer can focus on multiple source signals, and at the same time equalise the levels of those signals. For example, if there were three source signals which it were desired to capture, with two of those signals being stronger than the third, the system can form three main lobes in the beam pattern, with the lobe directed to the weaker signal having a stronger gain than the lobes directed to the stronger signals, thereby amplifying the weaker source more and equalising the signal strengths for the three sources.

Preferably the beamformer formulates the or each requirement as a convex constraint. More preferably, the beamformer formulates the or each requirement as a linear equality constraint. With the constraints formulated in this way, the problem becomes a second order cone programming problem which is a subset of convex optimization problems. The numerical solution of second order programming problems has been studied in detail and a number of fast and efficient algorithms are available for solving convex second order cone problems.

- 10 -

Preferably the beamformer formulates the or each main lobe requirement as a requirement that the array output for a unit magnitude plane wave incident on the array from the specified direction is equal to a predetermined constant. In other words, the beamforming pattern is constrained such that the array output will provide a specific gain for an incident plane wave from the specified direction. This form of constraint is a linear equality and thus can be applied to a second order cone programming problem as above.

In preferred embodiments of the invention, the input parameters include a requirement that the array gain in a specified direction is below a given level, so as to form a null in the beampattern. In other words, the beamformer optimization problem is subjected to an optimization constraint that the array gain in at least one direction is below a selected threshold. This enables minimization of the sidelobe region of the beampattern, thus restricting the size of the secondary maxima of the system. It also allows creation of "notches" in the beampattern, creating a particularly low gain in the selected direction(s) for blocking interference signals.

More preferably, the input parameters include requirements that the array gain in a plurality of specified directions is below a given level, so as to form multiple nulls in the beampattern. In other words, the beamformer optimization problem is subjected to optimization constraints that the array gain in a plurality of directions is below a corresponding threshold. In this way, multiple nulls can be formed in the beampattern, thereby allowing suppression of multiple interference sources.

Still more preferably, individual maximum gain levels are provided for each of the plurality of specified directions, so as to form multiple nulls of different depths in the beampattern. In this way, different levels of constraint can be applied to different regions of the beam pattern. For example, the side lobes can be kept generally below a certain level, but with more stringent constraints being applied in regions where notches or nulls are desired for blocking interference signals. By applying the most stringent constraints only where they are required, the freedom of the beampattern is affected less, allowing the remainder of the pattern to minimise more uniformly.

Preferably, the beamformer formulates the or each side lobe requirement as a convex constraint. More preferably, the beamformer formulates the or each side lobe requirement as a second order cone constraint. As above, with the constraints formulated in this way, the problem becomes a

- 11 -

second order cone programming problem which is a subset of convex optimization problems. The numerical solution of second order programming problems has been studied in detail and a number of fast and efficient algorithms are available for solving convex second order cone problems.

Most preferably, the beamformer formulates the or each side lobe requirement as a requirement that the magnitude of the array output for a unit magnitude plane wave incident on the array from the specified direction is less than a predetermined constant. As above, this form of constraint is a convex inequality and thus can be applied to a second order cone programming problem as above.

Preferably, the input parameters include a requirement that the beampattern has a specified level of robustness. In applications where it is vital that the desired source signal be picked up, it is desirable to ensure that the system does not fail merely due to minor mis-alignments, random noise or other unexpected interference. In other words, it is desired that the system be resilient to errors to a certain extent. Preferably, the level of robustness is specified as a limitation on a norm of a vector comprising the weighting coefficients. More preferably, the norm is the Euclidean norm. As described in more detail below, minimising the norm of the weighting coefficients vector maximises the white noise gain of the array and thus increases the robustness of the system.

Preferably, the weighting coefficients are optimized by second order cone programming. As described above, second order cone programming is a subset of convex optimization techniques which has been studied in much detail and fast and efficient algorithms are available for solving such problems rapidly. Such numerical algorithms can converge on the global minimum of the problem very quickly, even when numerous constraints are applied on the system.

Preferably one or more weighting coefficients are optimized for each order n of spherical harmonic, but within each order of spherical harmonics, said weighting coefficients are common to all degrees $m=-n$ to $m=n$ of said order n . By reducing the number of weighting coefficients in this manner, the beampattern is confined to being rotationally symmetric about the look direction. However, such a beampattern is useful in a number of circumstances and the reduction in the number of coefficients simplifies the optimization problem and allows for faster computation of the solution.

In some preferred embodiments the input signals may be transformed into the frequency domain before being decomposed into the spherical harmonics domain. In some preferred embodiments the beamformer may be a broadband beamformer in which the frequency domain signals are divided into narrowband frequency bins and wherein each bin is optimized and weighted separately before the frequency bins are recombined into a broadband output. In other preferred embodiments, the input signals may be processed in the time domain and the weighting coefficients may be the tap weights of finite impulse response filters applied to the spherical harmonic signals.

The choice of processing domain will depend on the circumstances of the particular scenario, i.e. the particular beam forming problem. For example, the expected frequency spectrum to be received and processed may influence the choice between the time domain and the frequency domain, with one domain giving a better solution or being computationally more efficient.

Processing in the time domain is particularly advantageous in some instances because it is inherently broadband in nature. Therefore, with such an implementation, there is no need to perform a computationally intensive fourier transform into the frequency domain before optimization and a corresponding computationally intensive inverse fourier transform back to the time domain after optimization. It also avoids the need to split the input into a number of narrowband frequency bins in order to obtain a broadband solution. Instead a single optimization problem may be solved for all weighting coefficients. In some embodiments, the weighting coefficients will take the form of finite impulse response (FIR) filter tap weights.

In principle, from the viewpoint of beamforming performance, the time domain and the frequency domain implementations can give the same beamforming performance if the FIR length equals the FFT length. The time domain may have a significant advantage over the frequency domain in some real implementations since no FFT and inverse FFT will be needed. However from the viewpoint of optimization complexity, assuming that the FIR and FFT have the same length L , the computational complexity of optimizing a set of FIRs (i.e. L FIR coefficients for each channel) by a single optimization, would be much higher than that of optimizing a set of array weights (i.e. a single weight for each channel) by L sub-band optimizations. Therefore, each approach may have advantages in different situations.

- 13 -

According to a second aspect, the present invention provides a beamformer comprising: an array of sensors, each of which is arranged to generate a signal; a spherical harmonic decomposer which is arranged to decompose the input signals into the spherical harmonics domain and to output the decomposed signals; a weighting coefficients calculator which is arranged to calculate weighting coefficients to be applied to the decomposed signals by convex optimization based on a set of input parameters; and an output generator which combines the decomposed signals with the calculated weighting coefficients into an output signal.

Such a beamformer implements all the benefits of the beamforming method described above. Moreover, all of the preferred features described above in relation to the beamforming method also apply to this implementation of the beamformer. As discussed above, in the time domain implementation, the output generator may comprise a number of finite impulse response filters.

Preferably, the beamformer further comprises a signal tracker which is arranged to evaluate the signals from the sensors to determine the directions of desired signal sources and the directions of unwanted interference sources. Such algorithms can run in parallel with the beamforming optimization algorithms, using the same data. While the localization algorithms pick out the directions of signals of interest and the directions of sources of interference, the beamformer forms an appropriate beampattern for amplifying the source signals and attenuating the interference signals.

As described above, this description is predominantly concerned with signal processing in the spherical harmonics domain. However, the techniques described herein are also applicable to the other domains, particularly the space domain. Although convex optimization has been used in some applications in space domain processing, it is believed to be a further inventive concept to formulate the problem for a spherical array. Therefore, according to a further aspect of the invention, there is provided a method of forming a beampattern in a beamformer for a spherical sensor array of the type in which the beamformer receives input signals from the array, applies weighting coefficients to the signals and combines them to form an output, wherein the weighting coefficients are optimized for a given set of input parameters by convex optimization. The inventors have recognised that the techniques and formulations developed in relation to the spherical harmonics domain, also apply to processing of a spherical array in the space domain and that it is therefore also possible, with this invention, to carry out multiple constraint optimization in real time in the space domain.

According to a further aspect, the invention provides a method of forming a beampattern in a beamformer of the type in which the beamformer receives input signals from a sensor array, applies weighting coefficients to the signals and combines them to form an output signal, wherein the weighting coefficients are optimized for a given set of input parameters by convex optimization, subject to constraints that the array gain in a plurality of specified directions be maintained at a given level, so as to form multiple main lobes in the beampattern, and wherein each requirement is formulated as a requirement that the array output for a unit magnitude plane wave incident on the array from the specified direction is equal to a predetermined constant.

As discussed above, the applicability of the methods derived in this description allow multiple constraints to be applied to the optimization problem without slowing the system up so much that it is of little practical use. Therefore, with the techniques and formulations of this invention, it is possible to apply multiple-main lobe formation and directivity constraints at the same time as applying multiple null forming and steering constraints, robustness constraints, and main-lobe beam-width constraints.

Preferably the beamformer is capable of operating in real time or quasi-real time. It will be appreciated that if the environment (e.g. the acoustic environment in audio applications) is fixed, it is not necessary to update the array weights during run time. Instead, a single set of optimized weights can be calculated in advance (e.g. at system startup or upon a calibration instruction) and need not be changed during operation. However, this set up does not make use of the full power of the invention. Preferably therefore, the array dynamically changes the optimum weights by re-solving the optimization problem according to the changing environment and constraints. As described above, the system can preferably re-optimize the array weights in real time or quasi-real time. The definition of real time may vary from application to application. However, in this description we mean that the array is capable of re-optimizing the array weights and forming a new optimized beam pattern in under a second. By quasi-real time, we mean an optimization time of up to about 5 seconds. Such quasi-real time may still be useful in situations where the dynamics of the environment do not change so rapidly, e.g. acoustics in a lecture where the number and direction of sources and interferences change only infrequently.

In real time or quasi-real time operation, the optimization operations preferably run in the background in order to gradually and continuously update the weights. Alternatively, sets of

- 15 -

weights for certain situations can be pre-calculated and stored in memory. The most appropriate set of weights can then be simply loaded into the system upon a change in environment.

However, it will be appreciated that this implementation does not make full use of the power and speed of this invention for actual optimization in real time.

The beamformer of the present invention can operate well in the space domain as well as in the spherical harmonics domain. The choice of domain will depend on the particular application of the array, the geometry of the array, the characteristics of the signals that it is expected to handle and the type of processing which is required of it. Although the space domain and the spherical harmonics domain are generally the most useful, other domains (e.g. the cylindrical harmonics domain) may also be used. In addition, the processing can be done in the frequency domain or the time domain. In particular, time domain processing with spherical harmonic decomposition is also useful. Preferably therefore the sensor signals are decomposed into a set of orthogonal basis functions for further processing. Most preferably, the orthogonal basis functions are the spherical harmonics, i.e. the solutions to the wave equation in spherical co-ordinates, and the wave field decomposition is performed by a spherical Fourier transform. The spherical harmonics domain is particularly well suited to spherical or near spherical arrays.

According to a further aspect, the present invention provides a method of optimizing a beampattern in a beamformer in a sensor array in which the input signals from the sensors are weighted and combined to form an array output signal, and wherein the sensor weights are optimized by expressing the array output power as a convex function of the sensor weights and minimizing the output power subject to one or more constraints, wherein the one or more constraints are expressed as equalities and/or inequalities of convex functions of the sensor weights.

It can be seen that the method of the present invention provides a general solution to the beamforming problem. A large number of constraints can be applied simultaneously in a single optimization problem, with one global optimum solution. However, if fewer constraints are applied, the results of the previous studies described above can be replicated. The present invention can therefore be seen as a more general solution to the problem.

A more detailed analysis of preferred forms of the system will now be discussed.

- 16 -

Since spatial over-sampling is typically employed in practice, the following analysis concentrates on spherical harmonics domain processing, which tends to be more efficient. However, it will be appreciated that the techniques discussed in relation to the spherical harmonic domain weighting functions applies in the same manner to an analysis in the space domain and results in an analogous convex optimization problem.

A few derivations of background material and useful results are given in the Annex to this application. The equation numbers in the following description follow on from those of the annex.

From previous studies, in order to easily form a regular or irregular, and frequency independent beam pattern, array weight design approaches have always utilized the inversion of $b_n(ka)$ in the spherical harmonics domain to decouple frequency-dependent components. However, as $b_n(ka)$ has small values at certain ka and n values, and its inversion will damage the robustness in practical implementations, we directly make the more general weights $w^*(k)$ the targets of our optimization framework.

This next section develops the results derived in the annex, using matrix formulations and derives the convex optimization problem and the corresponding constraints of the invention.

We use the notation

$$\mathbf{x} = \text{vec}(\{[x_{nm}]_{m=-n}^n\}_{n=0}^N) = [x_{00}, \dots, x_{nm}, \dots, x_{NN}]^T, \quad (16)$$

where $\text{vec}(\cdot)$ denotes stacking all the entries in the parentheses to obtain an $(N+1)^2 \times 1$ column vector and $(\cdot)^T$ denotes the transpose.

Using this notation, we can further define

$$\mathbf{w} = \text{vec}(\{[w_{nm}]_{m=-n}^n\}_{n=0}^N), \quad (17)$$

$$\mathbf{b} = \text{vec}(\{[b_n]_{m=-n}^n\}_{n=0}^N), \quad (18)$$

$$\mathbf{Y} = \text{vec}(\{[Y_n^m]_{m=-n}^n\}_{n=0}^N), \quad (19)$$

- 17 -

$$\mathbf{p} = \text{vec}(\{[p_{nm}]_{m=-n}^n\}_{n=0}^N). \quad (20)$$

Note that (18) means that $\mathbf{b}$ has repetitions of b_n from the $(n^2 + 1)$ through $(n + 1)^2$ entries. From (9), it is seen that $\mathbf{p}$ can be viewed as the modal array manifold vector.

We can write (14) in vector notation as

$$y(ka) = \mathbf{w}^H(k) \mathbf{x}(ka) = \mathbf{x}^H(ka) \mathbf{w}(k), \quad (21)$$

where $(\cdot)^H$ denotes the Hermitian transpose.

In the following description, the optimization problem is formulated as minimizing the array output power in order to suppress any interferences coming from outside beam directions, while the signal from the mainlobe direction is maintained and the sidelobes are controlled.

Furthermore, for the purpose of improving the beamformer's robustness, a white noise gain constraint is also applied to limit the norm of array weights to a specified constant.

The array output power is given by

$$P_0(\omega) = E[y(ka)y^*(ka)] = \mathbf{w}^H(k) E[\mathbf{x}(ka)\mathbf{x}^H(ka)] \mathbf{w} = \mathbf{w}^H(k) \mathbf{R}(\omega) \mathbf{w}(k), \quad (22)$$

where $E[\cdot]$ denotes the statistical expectation of the quantity in the brackets, and $\mathbf{R}(\omega)$ is the covariance matrix (spectral matrix) of $\mathbf{x}$.

The directivity pattern, denoted by $H(ka, \Omega)$, is a function of the array's response to a unit input signal from all angles of interest. Thus,

$$\begin{aligned} H(ka, \Omega) &= \sum_{s=1}^M \alpha_s p(ka, \Omega, \Omega_s) w^*(k, \Omega_s) \\ &= \sum_{n=0}^N \sum_{m=-n}^n p_{nm}(ka, \Omega) w_{nm}^*(k) = \mathbf{w}^H(k) \mathbf{p}(ka, \Omega). \end{aligned} \quad (23)$$

- 18 -

Assuming that the signal sources are uncorrelated from each other, the covariance matrix of $\mathbf{x}$ has the following form

$$\begin{aligned} \mathbf{R}(\omega) &= \mathbf{E}[\mathbf{x}(ka)\mathbf{x}^H(ka)] \\ &= \beta^2 \sigma_0^2 \mathbf{p}(ka, \Omega_0) \mathbf{p}^H(ka, \Omega_0) + \sum_{d=1}^D \sigma_d^2 \mathbf{p}(ka, \Omega_d) \mathbf{p}^H(ka, \Omega_d) + \mathbf{Q}(\omega), \end{aligned} \quad (24)$$

where $\{\sigma_d^2\}_{d=0}^D$ are the powers of the $D+1$ uncorrelated signals, and $\mathbf{Q}(\omega) = \mathbf{E}[\mathbf{N}(\omega)\mathbf{N}^H(\omega)]$ is the noise covariance matrix with $\mathbf{N} = \text{vec}(\{[N_{nm}]_{m=-n}^n\}_{n=0}^N)$.

We now consider a special case of noise field: isotropic noise, i.e., noise distributed uniformly over a sphere. Isotropic noise with power spectral density $\sigma_n^2(\omega)$ can be viewed as if there are an infinite number of uncorrelated plane waves arriving at the sphere from all directions Ω with uniform power density $\sigma_n^2(\omega)/(4\pi)$. Thus, by integrating the covariance matrix over all directions, the isotropic noise covariance matrix is given by

$$\mathbf{Q}_{iso}(\omega) = \frac{\sigma_n^2(\omega)}{4\pi} \int_{\Omega \in S^2} \mathbf{p}(ka, \Omega) \mathbf{p}^H(ka, \Omega) d\Omega, \quad (25)$$

Using (7), (18) and (19), (25) can be rewritten as

$$\begin{aligned} \mathbf{Q}_{iso}(\omega) &= \frac{\sigma_n^2(\omega)}{4\pi} \int_{\Omega \in S^2} [\mathbf{b}(ka) \circ \mathbf{Y}(\Omega)] [\mathbf{b}(ka) \circ \mathbf{Y}(\Omega)]^H d\Omega \\ &= \frac{\sigma_n^2(\omega)}{4\pi} \text{diag}\{\mathbf{b}(ka) \circ \mathbf{b}^*(ka)\} \\ &= \frac{\sigma_n^2(\omega)}{4\pi} \text{diag}\{|b_0(ka)|^2, |b_1(ka)|^2, |b_1(ka)|^2, |b_1(ka)|^2, \dots, |b_N(ka)|^2\}, \end{aligned} \quad (26)$$

where $\circ$ denotes the Hadamard (i.e. element-wise) product of two vectors. Note that the spherical harmonic orthonormal property (4) has been employed in the above derivation.

- 19 -

In practical applications, the exact covariance matrix $\mathbf{R}(\omega)$ is unavailable. Therefore, the sample covariance matrix is usually used instead of Eq. (24). The sample covariance matrix is given by:

$$\hat{\mathbf{R}}(\omega) = \frac{1}{I} \sum_{i=1}^I \mathbf{x}(ka, i) \mathbf{x}^H(ka, i)$$

where I is the number of snapshots.

The array gain $G(k)$ is defined to be the ratio of the signal-to-noise ratio (SNR) at the output of the array to the SNR at an input sensor.

$$G(k) = \frac{\sigma_0^2 |\mathbf{w}^H(k) \mathbf{p}(ka, \Omega_0)|^2}{\mathbf{w}^H(k) \mathbf{Q}(\omega) \mathbf{w}(k)} \bigg/ \frac{\sigma_0^2}{\sigma_n^2} = \frac{|\mathbf{w}^H(k) \mathbf{p}(ka, \Omega_0)|^2}{\mathbf{w}^H(k) \boldsymbol{\rho}(\omega) \mathbf{w}(k)}, \quad (27)$$

where $\boldsymbol{\rho}(\omega) = \mathbf{Q}(\omega) / \sigma_n^2(\omega)$ is the normalized noise covariance matrix.

A common measure of performance of an array is the directivity. The directivity factor $D(k)$, or directive gain, can be interpreted as the array gain against isotropic noise. Replacing $\mathbf{Q}$ in (27) by $\mathbf{Q}_{iso}$ gives the directivity factor

$$D(k) = \frac{\sigma_n^2(\omega) |\mathbf{w}^H(k) \mathbf{p}(ka, \Omega_0)|^2}{\mathbf{w}^H(k) \mathbf{Q}_{iso}(\omega) \mathbf{w}(k)} = \frac{4\pi \left| \sum_{n=0}^N \sum_{m=-n}^n p_{nm}(ka, \Omega) w_{nm}^*(k) \right|^2}{\sum_{n=0}^N |b_n(ka)|^2 \sum_{m=-n}^n |w_{nm}(k)|^2}. \quad (28)$$

The directivity index (DI) is then defined as $DI(k) = 10 \log_{10} D(k)$ dB.

There are many performance measures by which one may assess the capabilities of a beamformer. Commonly used array performance measures are directivity, array gain, beamwidth, sidelobe level, and robustness.

- 20 -

The trade-off among these conflicting performance measures represents the beamformer design optimization problem. In the method of this invention, the optimization problem is directed to minimizing the output power subject to a distortionless constraint on the signal of interest (SOI) (i.e. to form the main lobe in the beampattern) together with any number of other desired constraints, such as sidelobes and robustness constraints. Taking the array weights vector $\mathbf{w}(k)$ as the optimization variable, the multi-constraint beamforming optimization problem may be formulated as

$$\begin{aligned} \min_{\mathbf{w}} \mathbf{w}^H(k) \mathbf{R}(\omega) \mathbf{w}(k), \\ \text{subject to } H(ka, \Omega_0) = 4\pi / M, \\ |H(ka, \Omega)| \leq \varepsilon \cdot 4\pi / M, \forall \Omega \in \Omega_{SL}, \\ WNG(k) \geq \zeta(k), \end{aligned} \quad (29)$$

where Ω_{SL} is the sidelobe region, and ε and ζ are user parameters to control the sidelobes and the white noise gain (i.e., array gain against white noise) WNG , respectively. A white noise gain constraint has been commonly used to improve the robustness of a beamformer. The look direction (i.e. the direction of the main lobe) is Ω_0 , the SOI's direction of arrival.

The white noise gain (WNG) is given by

$$WNG(k) = \frac{1}{\sum_{s=1}^M |w(k, \Omega_s)|^2}. \quad (30)$$

Using (15), WNG can be rewritten as

$$WNG(k) = \frac{1}{\sum_{s=1}^M |w(k, \Omega_s)|^2} = \frac{4\pi / M}{\sum_{n=0}^N \sum_{m=-n}^n |w_{nm}(k)|^2} = \frac{4\pi / M}{\mathbf{w}^H(k) \mathbf{w}(k)}. \quad (31)$$

It is seen that the white noise gain is inversely proportional to the norm of the weight vector. In order to improve the beamformer's robustness, the denominator, or norm of array weights may be limited to a certain threshold.

- 21 -

Due to the correlation between responses at neighbouring directions, the sidelobe region Ω_{SL} can be approximated using a finite number of grid points in direction, $\Omega_l \in \Theta_{SL}$, $l = 1, \dots, L$. The choice of L is determined by the required accuracy of approximation.

Using (23) and (31), (29) now takes the form

$$\begin{aligned} & \min_{\mathbf{w}} \mathbf{w}^H(k) \mathbf{R}(\omega) \mathbf{w}(k), \\ & \text{subject to } \mathbf{w}^H(k) \mathbf{p}(ka, \Omega_0) = 4\pi / M, \\ & \left| \mathbf{w}^H(k) \mathbf{p}(ka, \Omega_l) \right| \leq \varepsilon \cdot 4\pi / M, \Omega_l \in \Theta_{SL}, l = 1, \dots, L, \\ & \|\mathbf{w}(k)\| \leq \sqrt{\frac{4\pi}{M\zeta(k)}}. \end{aligned} \quad (32)$$

where $\|\cdot\|$ denotes the Euclidean norm.

Second Order Cone Programming is a subclass of the general convex programming problems where a linear function is minimized subject to a set of second-order cone constraints and possibly a set of linear equality constraints. The problem can be described as

$$\begin{aligned} & \min_{\mathbf{y}} \mathbf{b}^T \mathbf{y}, \\ & \text{subject to } \|\mathbf{A}_i \mathbf{y} + \mathbf{b}_i\| \leq \mathbf{c}_i^T \mathbf{y} + d_i, \quad i = 1, 2, \dots, I, \\ & \mathbf{F} \mathbf{y} = \mathbf{g}, \end{aligned}$$

where $\mathbf{b} \in C^{\alpha \times 1}$, $\mathbf{y} \in C^{\alpha \times 1}$, $\mathbf{A}_i \in C^{(\alpha_i - 1) \times \alpha}$, $\mathbf{b}_i \in C^{(\alpha_i - 1) \times 1}$, $\mathbf{c}_i \in C^{\alpha \times 1}$, $\mathbf{c}_i^T \mathbf{y} \in \mathfrak{R}$, $d_i \in \mathfrak{R}$, $\mathbf{F} \in C^{g \times \alpha}$, $\mathbf{g} \in C^{g \times 1}$ with $\mathfrak{R}$ and C being the set of real and complex numbers (or matrices) respectively.

Taking the optimization problem defined in (32) above, and omitting the arguments ω and k temporarily for convenience, let

$$\mathbf{R} = \mathbf{U}^H \mathbf{U} \quad (32.1)$$

- 22 -

be the Cholesky factorization of $\mathbf{R}$. We obtain

$$\mathbf{w}^H \mathbf{R} \mathbf{w} = (\mathbf{U} \mathbf{w})^H (\mathbf{U} \mathbf{w}) = \|\mathbf{U} \mathbf{w}\|^2 \quad (32.2)$$

Introducing a new scalar non-negative variable y_1 , and defining $\mathbf{y} = [y_1, \mathbf{w}^T]^T$ and $\mathbf{b} = [1, \mathbf{0}^T]^T$, where $\mathbf{0}$ is the vector of zeros of a conformable dimension, the optimization problem (32) can be rewritten as

$$\begin{aligned} & \min_{\mathbf{y}} \quad \mathbf{b}^T \mathbf{y} \\ & \text{subject to} \quad \left| \mathbf{0} \quad \mathbf{p}^H(ka, \Omega_0) \right| \mathbf{y} = 4\pi / M, \\ & \quad \left\| \begin{bmatrix} \mathbf{0} & \mathbf{U} \end{bmatrix} \mathbf{y} \right\| \leq \begin{bmatrix} 1 & \mathbf{0}^T \end{bmatrix} \mathbf{y}, \\ & \quad \left| \left[\mathbf{0} \quad \mathbf{p}^H(ka, \Omega_l) \right] \mathbf{y} \right| \leq \varepsilon \cdot 4\pi / M, \quad \Omega_l \in \Theta_{SL}, \quad l = 1, \dots, L, \\ & \quad \left\| \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{y} \right\| \leq \sqrt{\frac{4\pi}{M\zeta(k)}}, \end{aligned} \quad (32.3)$$

where $\mathbf{I}$ is an identity matrix. Thus, the optimization problem (32) has been rewritten in the form of Second Order Cone Programming problem. Numerical methods can therefore be used to find the solution to this problem efficiently. After solving the optimization problem, the only parameters of interest in the vector of variables $\mathbf{y}$ are given by its subvector $\mathbf{w}$.

It can therefore be seen that this optimization problem has been formulated as a convex second-order cone programming (SOCP) problem where a linear function is minimized subject to a set of second-order cone constraints and possibly a set of linear equality constraints. This is a subclass of the more general convex programming problems. SOCP problems are computationally tractable and can be solved efficiently using known numerical solvers. An example of such a numerical solver is the SeDuMi solver (<http://sedumi.ie.lehigh.edu/>) available for MATLAB. The global optimal numerical solution of an SOCP problem is guaranteed if it exists, i.e. if a global minimum exists for the problem, the numerical solving algorithm will find it. Further, as the techniques are highly computationally tractable, many constraints can be included in the optimization problem while maintaining a real-time optimization. SOCP is more efficient in computation than general convex optimization and so it is highly preferred for real time applications.

- 23 -

Concerning computational complexity, when interior-point methods are used to solve the SOCP problem derived in (32.3) above, the number of iterations to decrease the duality gap to a constant fraction of itself is bounded above by $O(\sqrt{I+1})$ (here the term “1” is due to the equality constraint), and the amount of computation per iteration is $O[\alpha^2(\sum_i \alpha_i + g)]$.

For the optimization problem (32.2), the amount of computation per iteration is

$$O\{[(N+1)^2 + 1]^2[1 + ((N+1)^2 + 1) + 2L + ((N+1)^2 + 1)]\} = O\{[(N+1)^2 + 1]^2[3 + 2(N+1)^2 + 2L]\}$$

and the number of iterations is $O(\sqrt{L+3})$. The algorithm converges typically in less than 10 iterations (a well-known and widely accepted fact in the optimization community).

Before going on to describe preferred embodiments of the invention, it should be noted that the above analysis is all based on the assumption that the signal sources are in the far-field, so that they may be approximated by plane waves incident on the array.

It should also be noted that the analysis is based on a narrowband beamformer design. The broadband beamformer can be simply realized by decomposing the frequency band into narrower frequency bins and processing each bin with the narrowband beamformer.

If implemented in the time domain, then in order to achieve a broadband beamformer, the proper time delays and weights are applied to each of the sensors for each sub-band, in order to form the beampattern, or, alternatively an FIR-and-weight method can be used to achieve broadband beamforming in the time domain. However, if implemented in the frequency domain, then for each narrow frequency bin, complex weights are applied to each of the sensors. The above description focuses on the frequency domain implementation and optimizes the complex weights for each frequency. A more detailed description of a time domain implementation follows.

The above approach bases the signal model in the frequency domain, where the complex-valued modal transformation and array processing are employed. In order to achieve a broadband beamformer, which is very important for speech and audio applications, the broadband array signals are decomposed into narrower frequency bins using the discrete Fourier transform (DFT), then each frequency bin is independently processed using the narrowband beamforming algorithm, and then an inverse DFT is employed to synthesise the broadband output signal. Since

- 24 -

the frequency-domain implementation is performed with block processing, it might be unsuitable for time-critical speech and audio applications due to its associated time delay.

It is well known that, in classical element space array processing, the broadband beamformer can be implemented in the time domain using the filter-and-sum structure in which a bank of finite impulse response (FIR) filter are placed at the output of sensors, and the filter outputs are summed together to produce the final output time series. The main advantage of the time-domain filter-and-sum implementation is that the beamformer can be updated at run time when each new snapshot arrives. The key point of the filter-and-sum beamformer design is how to calculate the FIR filters' tap weights, in order to achieve the desired beamforming performance.

The spherical array modal beamforming can also be implemented in the time domain with the real-valued modal transformation and the filter-and-sum beamforming structure. WO 03/061336 proposed a novel time domain implementation structure for spherical array modal beamformer, within the spherical harmonics framework. In that implementation, the number of the signal processing channels is reduced significantly, the real and imaginary parts of spherical harmonics are employed as the spherical Fourier transform basis to convert the time domain broadband signals to the real-valued spherical harmonics domain, and the look direction of the beamformer can be tactfully decoupled from its beampattern shape. To achieve a frequency independent beampattern, WO 03/061336 proposed to employ inverse filters to decouple the frequency-dependent components in each signal channel, however, such kind of inverse filtering could damage the system robustness (J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield", in *Proc.ICASSP*, vol.2, May 2002, pp.1781-1784.) . Moreover, since no systematic performance analysis framework has been formulated for such a filter-and-sum modal beamforming structure, all the mutually conflicting broadband beamforming performance measures, such as directivity factor, sidelobe level, and robustness, etc. cannot be effectively controlled.

Here, a broadband modal beamforming framework implemented in the time domain is presented. This technique is based on a modified filter-and-sum modal beamforming structure. We derive the expression for the array response, the beamformer output power against both isotropic noise and spatially white noise, and the mainlobe spatial response variation (MSRV) in terms of the FIR filters tap weights. With the aim of achieving a suitable trade-off among multiple conflicting performance measures (e.g., directivity index, robustness, sidelobe level, mainlobe response

- 25 -

variation, etc.), we formulate the FIR filters' tap weights design problem as a multiply-constrained optimization problem which is computationally tractable.

In addition, in the arrangement described here, a steering unit is described. With the steering unit, the number of signal processing channels is reduced, and the modal beamforming approach is computationally more efficient compared to a classical element space array processing. The steering unit reduces the computational complexity by forming a beam pattern which is rotationally symmetric about the look direction. Although not as general as the asymmetric beam pattern discussed above, such a configuration is still frequently useful. It will be appreciated however that the steering unit is not an essential component of the time domain beamformer discussed below and it can be omitted if the more general beam pattern formation is desired.

In the following, we will reformulate some of the results previously derived for the frequency domain approach and add in a beam steering unit. We assume that the time series received at the s th microphone is $x_s(t)$ and the frequency-domain notation is $x(f, \Omega_s)$. The discrete spherical Fourier transform (spherical Fourier coefficients) of $x(f, \Omega_s)$, is given by

$$x_{nm}(f) = \sum_{s=1}^M \alpha_s x(f, \Omega_s) [Y_n^m(\Omega_s)]^* \quad (T5)$$

Using (T5), the sound field is transformed from the time or frequency domain into the spherical harmonics domain.

We assume each microphone has a weighting, denoted by $w^*(f, \Omega_s)$. The array output, denoted by $y(f)$, can be calculated as:

$$y(f) = \sum_{s=1}^M \alpha_s x(f, \Omega_s) w^*(f, \Omega_s) = \sum_{n=0}^N \sum_{m=-n}^n x_{nm}(f) w_{nm}^*(f), \quad (T6)$$

where $w_{nm}^*(f)$ are the spherical Fourier coefficients of $w^*(f, \Omega_s)$. The second summation term in (T6) can be viewed as weighting in the spherical harmonics domain.

As before, we use the notation

$$\mathbf{x}_b = \text{vec}(\{[x_{nm}]_{m=-n}^n\}_{n=0}^N) = [x_{00}, \dots, x_{nm}, \dots, x_{NN}]^T, \quad (T7)$$

- 26 -

where $\text{vec}(\cdot)$ denotes stacking all the entries in the parentheses to obtain an $(N+1)^2 \times 1$ column vector and $(\cdot)^T$ denotes the transpose.

We can rewrite (T6) in vector notation as

$$y(f) = \mathbf{w}_b^H(f) \mathbf{x}_b(f), \quad (\text{T8})$$

where $\mathbf{w}_b = \text{vec}(\{[w_{nm}]_{m=-n}^n\}_{n=0}^N)$.

The array output power is given by

$$P_{out}(\omega) = E[y(f)y^*(f)] = \mathbf{w}_b^H(f) E[\mathbf{x}_b(f)\mathbf{x}_b^H(f)] \mathbf{w}_b(f) = \mathbf{w}_b^H(f) \mathbf{R}_b(f) \mathbf{w}_b(f), \quad (\text{T9})$$

where $E[\cdot]$ denotes the statistical expectation of the quantity in the brackets, $\mathbf{R}_b(f)$ is the covariance matrix (spectral matrix) of $\mathbf{x}_b$.

The directivity pattern, denoted by $B(f, \Omega)$, is a function of the array's response to a unit input signal from all angles of interest Ω . Thus,

$$B(f, \Omega) = \sum_{s=1}^M \alpha_s p(ka, \Omega, \Omega_s) w^*(f, \Omega_s) = \sum_{n=0}^N \sum_{m=-n}^n p_{nm}(ka, \Omega) w_{nm}^*(f). \quad (\text{T10})$$

By applying Parseval's relation for the spherical Fourier transform to the weights, we have

$$\sum_{s=1}^M \alpha_s |w(f, \Omega_s)|^2 = \sum_{n=0}^N \sum_{m=-n}^n |w_{nm}(f)|^2. \quad (\text{T11})$$

Intuitively, we want the microphones distributed uniformly on the spherical surface. However, true equidistant spatial sampling is only possible for arrangements that are constructed according to five regular polyhedrons geometries, i.e., tetrahedron, cube, octahedron, dodecahedron, and icosahedron. An arrangement that provides a close-to-uniform sampling scheme has been used, in which 32 microphones are located at the center of the faces of a truncated icosahedron.

Another example of specific, simple, close-to-uniform grid shown to behave well with spherical array is Fliege grid. In these close-to-uniform cases, $\alpha_s \cong 4\pi / M$.

In order to form a beampattern with rotational symmetry around the look direction Ω_0 , the array weights take the form

$$w_{nm}^*(f) = \sqrt{\frac{4\pi}{2n+1}} c_n(f) Y_n^m(\Omega_0). \quad (\text{T12})$$

- 27 -

where $\sqrt{\frac{4\pi}{2n+1}}Y_n^m(\Omega_0)$ act as the steering units that are responsible for steering the look direction by Ω_0 and $c_n(f)$ act as pattern generation.

Using (T12) in (T6) gives

$$y(f) = \sum_{n=0}^N \left[\sqrt{\frac{4\pi}{2n+1}} \sum_{m=-n}^n x_{nm}(f) Y_n^m(\Omega_0) \right] c_n(f). \quad (T13)$$

According to (T5) and (T13), we get the modal beamformer structure as depicted in Fig. 20.

First, the sound field data $x(f, \Omega_s)$ are transformed from the time or frequency domain into the spherical harmonics domain data $x_{nm}(f)$. Then, the harmonics domain data $x_{nm}(f)$ are directly fed to the modal beamformer (steering, weighting, and summing). This is a difference to that presented by Meyer and Elko in "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield" in *Proc. ICASSP*, vol. 2, May 2002, pp. 1781-1784, where the spherical harmonics, which have been compensated for b_n , are fed to a modal beamformer instead. This modification is presented to avoid a bad robustness of the beamformer caused by the compensation unit.

Using (T12), (5) and (7) in (T10) gives

$$\begin{aligned} B(f, \Omega) &= \sum_{n=0}^N \sum_{m=-n}^n p_{nm}(ka, \Omega) w_{nm}^*(f) = \sum_{n=0}^N \tilde{c}_n(f) b_n(ka) \sum_{m=-n}^n [Y_n^m(\Omega)]^* Y_n^m(\Omega_0) \\ &= \sum_{n=0}^N \tilde{c}_n(f) b_n(ka) \frac{2n+1}{4\pi} P_n(\cos \Theta) = \sum_{n=0}^N c_n(f) b_n(ka) \sqrt{\frac{2n+1}{4\pi}} P_n(\cos \Theta), \end{aligned} \quad (T14)$$

where P_n is the Legendre polynomial and Θ is the angle between Ω and Ω_0 .

The robustness is an important measure of array performance and is commonly quantified by the white noise gain (WNG), i.e., array gain against white noise. Using (T11) and assuming that $\alpha_s \cong 4\pi / M$, WNG is given by

$$\begin{aligned} WNG(f) &= \frac{1}{\sum_{s=1}^M |w(f, \Omega_s)|^2} \cong \frac{4\pi / M}{\sum_{n=0}^N \sum_{m=-n}^n |w_{nm}(f)|^2} \\ &= \frac{4\pi / M}{\sum_{n=0}^N \frac{4\pi}{2n+1} c_n^*(f) c_n(f) \sum_{m=-n}^n Y_n^m(\Omega_0) [Y_n^m(\Omega_0)]^*} = \frac{4\pi / M}{\sum_{n=0}^N c_n^*(f) c_n(f)} = \frac{4\pi / M}{\mathbf{c}^H(f) \mathbf{c}(f)}, \end{aligned} \quad (T15)$$

- 28 -

where $\mathbf{c} = [c_0, \dots, c_n, \dots, c_N]^T$ is an $(N+1) \times 1$ column vector.

For the Maximum-DI modal beamformer and the Maximum-WNG modal beamformer, we have

$$[c_n(f)]_{MDI} = \frac{4\pi\sqrt{4\pi(2n+1)}}{M(N+1)^2 b_n(ka)}, \quad (T16)$$

$$[c_n(f)]_{MWNG} = \sqrt{\frac{4\pi}{2n+1}} \frac{4\pi b_n^*(ka)}{M \sum_{n=0}^N |b_n(ka)|^2}. \quad (T17)$$

where the subscript *MDI* and *MWNG* denote the Maximum-DI beamformer and the Maximum-WNG beamformer, respectively.

Up to now, the mathematical analysis of the modal transformation and beamforming has been discussed for complex spherical harmonics. We next consider the time-domain implementation of the broadband modal beamformer. Since the real-valued coefficients are more suitable for a time-domain implementation, we can work with the real and imaginary parts of the spherical harmonics domain data.

We assume that the sampled broadband time series received at the *s*th microphone is

$x_s(l) = x_s(t)|_{t=lT_s}$, where T_s is the sampling interval. Considering that $Y_n^m(\Omega)$ is independent of frequency, similar to (T5), the broadband spherical harmonics domain data is given

$$x_{nm}(l) = \sum_{s=1}^M \alpha_s x_s(l) [Y_n^m(\Omega_s)]^*, \quad l = 1, 2, \dots, \tilde{L}, \quad (T18)$$

where $x_{nm}(l)$ is the time-domain notation of $x_{nm}(f)$ in (T5), i.e., the inverse Fourier transform of $x_{nm}(f)$, and $\tilde{L}$ is the length of the input data.

Filter-and-sum structure has been used in broadband beamforming in classical element space array processing, in which each sensor feeds an FIR filter and the filter outputs are summed to produce the beamformer output time series. Using the analogy to classical array processing, we can apply the filter-and-sum structure to a modal beamformer. That is, we place a bank of real-valued FIR filters at the output of the steering unit the filters act as the role of complex weighting $c_n(f)$ in a broadband frequency band. An advantage of the modal beamformer with the steering unit is that it is computationally efficient since only $N+1$ FIR filters are required, in contrast to the classical element space beamformer, which requires M filters. Note that $M \geq (N+1)^2$. It

- 29 -

should be noted that the steering unit is an optional feature of this invention and if it is not used, a FIR filter is used for each of the $(N+1)^2$ spherical harmonics ($Y_n^m(\Omega)$).

Let $\mathbf{h}_n$ be the impulse response of the FIR filter corresponding to the spherical harmonics of order n , i.e., $\mathbf{h}_n = [h_{n1}, h_{n2}, \dots, h_{nL}]^T$, $n = 0, \dots, N$. Here, L is the length of the FIR filter.

Performing the inverse Fourier transform to (T13) and considering that the response of the filter $\mathbf{h}_n$ over the working frequency band is approximately equal to $c_n(f)$, the time-domain

beamformer output, denoted by $y(l)|_{l=1}^{\tilde{L}}$, can be given by

$$\begin{aligned} y(l)|_{l=1}^{\tilde{L}} &= \sum_{n=0}^N \left\{ \left[\sqrt{\frac{4\pi}{2n+1}} \sum_{m=-n}^n \left(\sum_{s=1}^M \alpha_s x_s(l) [Y_n^m(\Omega_s)]^* \right) Y_n^m(\Omega_0) \right]_{l=1}^{\tilde{L}} * \mathbf{h}_n \right\} \\ &= \sum_{n=0}^N \{ x_n(l, \Omega_0) |_{l=1}^{\tilde{L}} * \mathbf{h}_n \}. \end{aligned} \quad (\text{T19})$$

where $*$ denotes the convolution and

$$\begin{aligned} x_n(l, \Omega_0) &= \sqrt{\frac{4\pi}{2n+1}} \sum_{m=-n}^n \left(\sum_{s=1}^M \alpha_s x_s(l) [Y_n^m(\Omega_s)]^* \right) Y_n^m(\Omega_0) \\ &= \sqrt{\frac{4\pi}{2n+1}} \left\{ \tilde{x}_{n0}(l) Y_n^0(\Omega_0) + 2 \sum_{m=1}^n \tilde{x}_{nm}(l) \text{Re}[Y_n^m(\Omega_0)] + 2 \sum_{m=1}^n \tilde{x}_{nm}(l) \text{Im}[Y_n^m(\Omega_0)] \right\}, \end{aligned} \quad (\text{T20})$$

where $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ denote the real part and imaginary part, respectively,

$\tilde{x}_{nm}(l) = \sum_{s=1}^M \alpha_s x_s(l) \text{Re}[Y_n^m(\Omega_s)]$ and $\tilde{x}_{nm}(l) = \sum_{s=1}^M \alpha_s x_s(l) \text{Im}[Y_n^m(\Omega_s)]$. Note that the property

$Y_n^{-m}(\Omega) = (-1)^m [Y_n^m(\Omega)]^*$ has been employed in the above derivation.

Using (3) in (T20) gives

$$x_n(l, \Omega_0) = \tilde{x}_{n0}(l) P_n^0(\cos \theta_0) + 2 \sum_{m=1}^n \sqrt{\frac{(n-m)!}{(n+m)!}} P_n^m(\cos \theta_0) [\tilde{x}_{nm}(l) \cos(m\phi_0) + \tilde{x}_{nm}(l) \sin(m\phi_0)]. \quad (\text{T21})$$

According to (T19) and (T21), the time-domain implementation of the broadband modal beamformer can be given in Fig. 21. Note that the predelay T_0 is attached before the FIR filters for each harmonics. This predelay is used to compensate the inherent group delay of a FIR filter, which is typically chosen as $T_0 = -(L-1)T_s/2$. The aim is then to choose the impulse response

- 30 -

(or tap weights) of these FIR filters to achieve the desired frequency-wavenumber response of the modal beamformer.

The complex frequency response of the FIR filter with impulse response $\mathbf{h}_n$ is given by

$$H_n(f) = \sum_{l=1}^L h_{nl} e^{-j(l-1)2\pi T_s} = \mathbf{h}_n^T \mathbf{e}(f), \quad (\text{T22})$$

where $\mathbf{e}(f) = [1, e^{-j2\pi T_s}, \dots, e^{-j(L-1)2\pi T_s}]^T$.

Let $\eta = e^{-j2\pi T_0}$. The total weighting function in the pattern generation unit corresponding to the n th order spherical harmonics at frequency f is given by

$$\hat{c}_n(f) = \eta \mathbf{h}_n^T \mathbf{e}(f), \quad n = 0, 1, \dots, N. \quad (\text{T23})$$

We use $\hat{c}_n(k)$ in (T23) in lieu of $c_n(k)$ in (T14) to obtain

$$B(f, \Omega) = \sum_{n=0}^N b_n(ka) \sqrt{\frac{2n+1}{4\pi}} P_n(\cos \Theta) \eta \mathbf{h}_n^T \mathbf{e}(f). \quad (\text{T24})$$

Let $a_n(f, \Theta) = b_n(ka) \sqrt{\frac{2n+1}{4\pi}} P_n(\cos \Theta) \eta$, $\mathbf{a} = [a_0, \dots, a_n, \dots, a_N]^T$, and define an $(N+1)L \times 1$

composite vector $\mathbf{h} = [\mathbf{h}_0^T, \mathbf{h}_1^T, \dots, \mathbf{h}_N^T]^T$. Eq.(T24) can be rewritten as

$$\begin{aligned} B(f, \Omega) &= \sum_{n=0}^N a_n(f, \Theta) \mathbf{h}_n^T \mathbf{e}(f) = [\mathbf{a}(f, \Theta) \otimes \mathbf{e}(f)]^T \mathbf{h} \\ &= \mathbf{u}^T(f, \Theta) \mathbf{h} = \mathbf{h}^T \mathbf{u}(f, \Theta), \end{aligned} \quad (\text{T25})$$

where $\otimes$ denotes the Kronecker product and $\mathbf{u}(f, \Theta) = \mathbf{a}(f, \Theta) \otimes \mathbf{e}(f)$.

Note that, in the case of $\alpha_s = 4\pi / M$, the array output amplitude in (T6) is the factor $4\pi / M$

higher than the classical array processing, which is $\sum_{s=1}^M x(f, \Omega_s) w^*(f, \Omega_s)$. Therefore, the

distortionless constraint in the spherical harmonics domain becomes

$$\mathbf{h}^T \mathbf{u}(f, 0) = 4\pi / M. \quad (\text{T26})$$

We now consider a special case of noise field: spherically isotropic noise, i.e., noise distributed uniformly over a sphere. Isotropic noise with power spectral density $\sigma_n^2(f)$ can be viewed as if there are an infinite number of uncorrelated plane waves arriving at the sphere from all directions

- 31 -

Ω with uniform power density $\sigma_n^2(f)/(4\pi)$. Thus, by integrating the covariance matrix over all directions, the isotropic noise covariance matrix is given by

$$\mathbf{Q}_{biso}(f) = \frac{\sigma_n^2(f)}{4\pi} \int_{\Omega \in S^2} \mathbf{p}_b(ka, \Omega) \mathbf{p}_b^H(ka, \Omega) d\Omega, \quad (\text{T27})$$

$$\begin{aligned} &= \frac{\sigma_n^2(f)}{4\pi} \int_{\Omega \in S^2} [\mathbf{b}_b(ka) \circ \mathbf{Y}_b^*(\Omega)] [\mathbf{b}_b(ka) \circ \mathbf{Y}_b^*(\Omega)]^H d\Omega \\ &= \frac{\sigma_n^2(f)}{4\pi} \text{diag}\{\mathbf{b}_b(ka) \circ \mathbf{b}_b^*(ka)\} \\ &= \frac{\sigma_n^2(f)}{4\pi} \text{diag}\{|b_0(ka)|^2, |b_1(ka)|^2, |b_1(ka)|^2, |b_1(ka)|^2, \dots, |b_N(ka)|^2\}, \end{aligned} \quad (\text{T28})$$

where $\mathbf{p}_b = \text{vec}(\{[p_{nm}]_{m=-n}^n\}_{n=0}^N)$, $\mathbf{b}_b = \text{vec}(\{[b_n]_{m=-n}^n\}_{n=0}^N)$, $\mathbf{Y}_b = \text{vec}(\{[Y_n^m]_{m=-n}^n\}_{n=0}^N)$, $\circ$ denotes the Hadamard (i.e., element-wise) product of two vectors, and $\text{diag}\{\cdot\}$ denotes a square matrix with the elements of its arguments on the diagonal. Note that the spherical harmonic orthonormal property has been employed in the above derivation.

Consider a special case with only isotropic noise impinging on the microphone array. We use (T9) with $\mathbf{R}_b(f)$ replaced by the isotropic noise covariance matrix $\mathbf{Q}_{biso}(f)$ to obtain the isotropic noise-only beamformer output power, denoted by $P_{isoout}(\omega)$,

$$\begin{aligned} P_{isoout}(f) &= \mathbf{w}_b^H(f) \mathbf{Q}_{biso}(f) \mathbf{w}_b(f) \\ &= \sum_{n=0}^N \sum_{m=-n}^n w_{nm}^*(f) \frac{\sigma_n^2(f) |b_n(ka)|^2}{4\pi} w_{nm}(f) \\ &= \sum_{n=0}^N \frac{\sigma_n^2(f) |b_n(ka)|^2}{2n+1} c_n(f) c_n^*(f) \sum_{m=-n}^n Y_n^m(\Omega_0) [Y_n^m(\Omega_0)]^* \\ &= \sum_{n=0}^N c_n(f) \frac{\sigma_n^2(f) |b_n(ka)|^2}{4\pi} c_n^*(f) \\ &= \mathbf{c}^T(f) \mathbf{Q}_{ciso}(f) \mathbf{c}^*(f), \end{aligned} \quad (\text{T29})$$

where

$$\begin{aligned} \mathbf{Q}_{ciso}(f) &= \frac{\sigma_n^2(f)}{4\pi} \text{diag}\{|b_0(ka)|^2, |b_1(ka)|^2, |b_2(ka)|^2, \dots, |b_N(ka)|^2\} \\ &= \frac{\sigma_n^2(f)}{4\pi} \text{diag}\{\mathbf{b}_c(ka) \circ \mathbf{b}_c^*(ka)\} \end{aligned} \quad (\text{T30})$$

with $\mathbf{b}_c(ka) = [b_0(ka), b_1(ka), b_2(ka), \dots, b_N(ka)]^T$.

- 32 -

Using (T23) and denoting $\hat{\mathbf{c}} = [\hat{c}_0, \dots, \hat{c}_n, \dots, \hat{c}_N]^T$ gives

$$\hat{\mathbf{c}}(f) = [\eta \mathbf{h}_0^T \mathbf{e}(f), \dots, \eta \mathbf{h}_n^T \mathbf{e}(f), \dots, \eta \mathbf{h}_N^T \mathbf{e}(f)]^T = \eta [\mathbf{I}_{(N+1) \times (N+1)} \otimes \mathbf{e}(f)]^T \mathbf{h}. \quad (\text{T31})$$

Using $\hat{\mathbf{c}}(k)$ in lie of $\mathbf{c}(k)$ in (T29) gives

$$\begin{aligned} P_{\text{isoout}}(f) &= \mathbf{c}^T(f) \mathbf{Q}_{\text{ciso}}(f) \mathbf{c}^*(f) \\ &= \mathbf{h}^T [\mathbf{I}_{(N+1) \times (N+1)} \otimes \mathbf{e}(f)] \mathbf{Q}_{\text{ciso}}(f) [\mathbf{I}_{(N+1) \times (N+1)} \otimes \mathbf{e}(f)]^H \mathbf{h} \\ &= \mathbf{h}^T \mathbf{Q}_{\text{hiso}}(f) \mathbf{h}, \end{aligned} \quad (\text{T32})$$

where $\mathbf{Q}_{\text{hiso}}(f) = [\mathbf{I}_{(N+1) \times (N+1)} \otimes \mathbf{e}(f)] \mathbf{Q}_{\text{ciso}}(f) [\mathbf{I}_{(N+1) \times (N+1)} \otimes \mathbf{e}(f)]^H$ is the isotropic noise covariance matrix associated with $\mathbf{h}$.

For a broadband isotropic noise that occupy the frequency band $[f_L, f_U]$ with f_L and f_U being respectively the lower and upper bound frequency, its broadband covariance matrix, denoted by $\overline{\mathbf{Q}}_{\text{hiso}}$, can be given by performing the integration with respect to f over the region $[f_L, f_U]$

$$\overline{\mathbf{Q}}_{\text{hiso}} = \int_{f_L}^{f_U} \mathbf{Q}_{\text{hiso}}(f) df. \quad (\text{T33})$$

where the integration can be approximated by performing summation.

Assume that the spatially white noise has a flat spectrum $\sigma_n^2(f) = 1$ over the frequency band $[f_L, f_U]$. The broadband isotropic noise-only beamformer output power is

$$\overline{P}_{\text{isoout}} = \mathbf{h}^T \overline{\mathbf{Q}}_{\text{hiso}} \mathbf{h}. \quad (\text{T34})$$

Consider another special case with only spatially white noise with power spectral density $\sigma_n^2(f)$ impinging on the microphone array. In the case of $\alpha_s \cong 4\pi / M$, the spatially white noise-only beamformer output power, denoted by $P_{\text{wout}}(f)$, is given by

$$\begin{aligned} P_{\text{wout}}(f) &= \sigma_n^2(f) \left(\frac{4\pi}{M} \right)^2 \sum_{s=1}^M |w(f, \Omega_s)|^2 \cong \frac{4\pi \sigma_n^2(f)}{M} \sum_{n=0}^N \sum_{m=-n}^n |w_{nm}(f)|^2 \\ &= \frac{4\pi \sigma_n^2(f)}{M} \sum_{n=0}^N \hat{c}_n^*(f) \hat{c}_n(f) = \frac{4\pi \sigma_n^2(f)}{M} \sum_{n=0}^N |\mathbf{h}_n^T \mathbf{e}(f)|^2. \end{aligned} \quad (\text{T35})$$

- 33 -

Assume that the spatially white noise has a flat spectrum $\sigma_n^2(f) = 1$ over the whole frequency band $[0, f_s/2]$. The broadband beamformer output power, denoted by $\bar{P}_{wout}$, is given by

$$\begin{aligned}\bar{P}_{wout} &= \int_0^{f_s/2} P_{wout}(f) df = \int_0^{f_s/2} \frac{4\pi}{M} \sum_{n=0}^N |\mathbf{h}_n^T \mathbf{e}(f)|^2 df = \frac{4\pi}{M} \sum_{n=0}^N \int_0^{f_s/2} |\mathbf{h}_n^T \mathbf{e}(f)|^2 df \\ &= \frac{4\pi}{M} \sum_{n=0}^N \mathbf{h}_n^T \mathbf{h}_n = \frac{4\pi}{M} \mathbf{h}^T \mathbf{h}.\end{aligned}\quad (\text{T36})$$

The broadband white noise gain, denoted by $BWNG$, is then defined as

$$BWNG = \frac{(4\pi/M)^2}{\bar{P}_{wout}} = \frac{4\pi/M}{\mathbf{h}^T \mathbf{h}}. \quad (\text{T37})$$

A common measure of performance of an array is the directivity. The directivity factor $D(f)$, or directive gain, can be interpreted as the array gain against isotropic noise, which is given by

$$D(f) = \frac{\sigma_n^2(f)(4\pi/M)^2}{\mathbf{h}^T \mathbf{Q}_{hiso}(f) \mathbf{h}}. \quad (\text{T38})$$

Frequently, we express the directivity factor in dB and refer to it as the directivity index (DI),

$DI(f) = 10 \lg D(f)$, where $\lg(\cdot) = \log_{10}(\cdot)$.

The mainlobe spatial response variation (MSRV), is defined as

$$\gamma_{MSRV}(f, \theta) = |\mathbf{h}^T \mathbf{u}(f, \Theta) - \mathbf{h}^T \mathbf{u}(f_0, \Theta)|, \quad (\text{T39})$$

where f_0 is a chosen reference frequency.

Let $f_k \in [f_L, f_U]$ ($k=1, 2, \dots, K$), $\Theta_j \in \Theta_{ML}$ ($j=1, \dots, N_{ML}$), and $\Theta_i \in \Theta_{SL}$ ($i=1, \dots, N_{SL}$) be a chosen (uniform or nonuniform) grid that approximates the frequency band $[f_L, f_U]$, the mainlobe region Θ_{ML} , and the sidelobe region Θ_{SL} , respectively. We define an $N_{ML}K \times 1$ column vector γ_{MSRV} and an $N_{SL}K \times 1$ column vector $\mathbf{B}_{SL}$, whose entries are respectively given by

$$[\gamma_{MSRV}]_{k+(j-1)K} = \gamma_{MSRV}(f_k, \Theta_j), \quad (\text{T40})$$

$$[\mathbf{B}_{SL}]_{k+(i-1)K} = B(f_k, \Theta_i). \quad (\text{T41})$$

Then, the norm of γ_{MSRV} , i.e., $\|\gamma_{MSRV}\|_q$, can be used as a measure of the frequency-invariant

- 34 -

approximation of the synthesized broadband beampatterns over frequencies. The subscript $q \in \{2, \infty\}$ stands for the l_2 (Euclidean) and l_∞ (Chebyshev) norm, respectively. Similarly, $\|\mathbf{B}_{SL}\|_q$ is a measure of sidelobe behavior.

There are many performance measures by which one may assess the capabilities of a beamformer. Commonly used array performance measures are directivity, MSRV, sidelobe level, and robustness. The trade-off among these conflicting performance measures represents the beamformer design optimization problem. After formulating the broadband spherical harmonics domain beampattern $B(f, \Omega)$ (T25), the broadband isotropic noise-only beamformer output power $\bar{P}_{isoot}$ (T34), the broadband white noise gain $BWNG$ (T37), the mainlobe spatial response variation vector γ_{MSRV} (T40), and the sidelobe behavior vector $\mathbf{B}_{SL}$ (T41), the optimal array pattern synthesis problem for broadband modal beamformer can be formulated as

$$\min_{\mu_\ell} \mu_\ell, \ell = \{1, 2, 3, 4\}, \text{ subject to } B(f_k, \Omega_0) = 4\pi / M, k = 1, 2, \dots, K$$

$$\bar{P}_{isoot} \leq \mu_1, \|\gamma_{MSRV}\|_{q_1} \leq \mu_2, \|\mathbf{B}_{SL}\|_{q_2} \leq \mu_3, BWNG^{-1} \leq \mu_4, \quad (T42)$$

where $q_1, q_2 \in \{2, \infty\}$, and $\{\mu_\ell\}_{\ell=1}^4$ include a cost function and three user parameters. In a similar manner to the frequency domain problem discussed above, the optimization problem (T42) can be seen to be in a convex form and can be formulated as a so-called Second Order Cone Program (SOCP) which can be solved efficiently using an SOCP solver such as SeDuMi.

(T42) is given as a general expression which can be used to formulate an appropriate optimization problem depending on the beamforming objectives. For example, any of the four functions ($l = 1, 2, 3, 4$) can be used as the target function with any of the remaining functions used as further constraints. With $l=1$, the problem is formulated as minimising the output power of the array. With $l=2$, the problem is minimising the distortion in the mainlobe region. With $l=3$, the problem is minimising the sidelobe level and with $l=4$, the problem is maximising the white noise gain (robustness). In each case, the problem can be formulated subject to any or all of the other constraints, e.g. the problem can be formulated with $l=2$ as the objective function and with $l=1, l=3$ and $l=4$ as further constraints upon the problem. It can therefore be seen that this beamformer can be made extremely flexible.

In this arrangement, the filter tap weights are optimized for a given set of input parameters by convex optimization. The input signals from the sensor array are decomposed into the spherical

- 35 -

harmonics domain and then the decomposed spherical harmonic components are weighted by the FIR tap weights before being combined to form the output signal.

It should be noted that, although this description provides examples which are mostly concerned with telephone conferencing, the invention is in no way restricted to telephone conferencing applications. Rather the invention lies in the beamforming method which is equally applicable to other technological fields. These include ambisonics for high end surround sound systems and music recording systems where it may be desired to emphasise or de-emphasise particular regions of a very complex auditory scene. For such applications, the multi-main lobe directionality and level control and the simultaneous option of multiple side lobe constraints of the present invention are especially applicable.

Similarly, the beamformer of the present invention can also be applied to frequencies significantly higher or lower than voice band applications. For example, sonar systems with hydrophone arrays for communication and for localization tend to operate at lower frequencies, whereas ultrasound applications, with an array of ultrasound transducers operating typically in the frequency range of 5 to 30 MHz will also benefit from the beamformer of the present invention. Ultrasound beamforming can be used for example in medical imaging and tomography applications where rapid multiple selective directionality and interference suppression can lead to higher image quality. Ultrasound benefits greatly from real time speeds where imaging of patients is affected by constant movement from breathing and heartbeats as well as involuntary movements.

The present invention is also not limited to the analysis of longitudinal sound waves. Beam forming applies equally to electromagnetic radiation where the sensors are antennas. In particular, in radio frequency applications, radar systems can benefit greatly from beamforming. It will be appreciated that these systems also require real time adaptation of the beampattern for example when tracking several aircraft, each of which moves at considerable speed, multi-main lobe forming in real time is highly beneficial.

Further, applications of the present invention include seismic exploration, e.g. for petroleum detection. In this field, it is essential to have a very specific and accurate look direction. Therefore, the ability to apply main lobe width and directionality constraints fast allows faster operation of such systems where large amounts of ground have to be covered.

In one preferred embodiment therefore, the invention comprises a beamformer as described above, wherein the sensor array is an array of hydrophones.

In another preferred embodiment, the invention comprises a beamformer as described above, wherein the sensor array is an array of ultrasound transducers.

In another preferred embodiment, the invention comprises a beamformer as described above, wherein the sensor array is an array of antennas. In some preferred embodiments the antennas are radiofrequency antennas

It will be appreciated that the beamformer of the present invention is largely implemented in software and the software is executed on a computing device (which may be for example a general personal computer (PC) or a mainframe computer, or it may be a specially designed and programmed ROM (Read Only Memory) or it may be implemented in Field Programmable Gate Arrays (FPGAs). On such devices, software may be pre-loaded or it may be transferred onto the system via a data carrier or via transfer over a network. Systems which are connected to a Wide Area Network such as the Internet, may be arranged to download new versions of the software and updates to it.

Therefore, viewed from a further aspect, the present invention provides a software product which when executed on a computer cause the computer to carry out the steps of the above described method(s). The software product may be a data carrier. Alternatively, the software product may comprise signals transmitted from a remote location.

Viewed from another aspect, the invention provides a method of manufacturing a software product which is in the form of a physical carrier, comprising storing on the data carrier instructions which when executed by a computer cause the computer to carry out the method(s) described above.

Viewed from yet another aspect the invention provides a method of providing a software product to a remote location by means of transmitting data to a computer at that remote location, the data comprising instructions which when executed by the computer cause the computer to carry out the method(s) described above.

Preferred embodiments of the invention will now be described, by way of example only, and with reference to the accompanying drawings in which:

Figure 1 is a graph of Directivity Index as a function of ka for the norm-constrained, spherical array beamformer of the first embodiment, of order $N = 4$, for selected values of ζ ;

Figure 2 is a graph of White Noise Gain as a function of ka for the norm-constrained, spherical array beamformer of the first embodiment, of order $N = 4$, for selected values of ζ ;

Figure 3 is a graph of Directivity Index as a function of White Noise Gain for the norm-constrained, spherical array beamformer of the first embodiment, of order $N = 4$, for selected values of ka ;

Figure 4 shows the Directivity patterns of (a) a delay-and-sum beamformer, (b) a pure phase-mode beamformer, and (c) a norm-constrained robust maximum-DI beamformer when $ka = 3$, all arrays being of order $N = 4$ and using 25 microphones;

Figure 5 shows the Directivity pattern as a function of elevation θ for the delay-and-sum beamformer and the norm-constrained beamformer of the first embodiment with $\zeta = M/4$, at frequencies corresponding to $ka = 1, 2$ and 4 ;

Figure 6 shows the Directivity pattern of the norm-constrained beamformer of the second embodiment for the values of $\zeta = M/4$ and $ka = 3$;

Figure 7 shows the Directivity pattern of the robust beamformer with sidelobe control of the third embodiment when $ka = 3$. In (a) the DI is maximized, in (b) a notch is formed around the $(60^\circ, 270^\circ)$ direction with a depth of -40 dB and a width of 30° , and in (c) the output SNR is maximized, which forms a null in the direction of arrival of the interferer at $(60^\circ, 270^\circ)$;

Figure 8 shows beampatterns for (a) robust beamforming with uniform sidelobe control, and (b) robust beamforming with non-uniform sidelobe control and notch forming;

- 38 -

Figure 9 shows beam patterns for (a) robust beamforming with sidelobe control and automatic multi-null steering, and (b) robust beamforming with sidelobe control, multi-mainlobe and automatic multi-null steering;

Figure 10 shows beampatterns for (a) a single beam without sidelobe control, and (b) a single beam with non-uniform sidelobe control;

Figure 11 shows beampatterns for (a) a single beam with uniform sidelobe control and adaptive null steering, and (b) multi-beam without sidelobe control;

Figure 12 shows beampatterns for (a) multi-beam beamforming with sidelobe control and adaptive null steering, and (b) multi-beam beamforming with mainlobe levels control;

Figure 13 shows a 4th order regular beampattern formed with a robustness constraint, but with no side lobe control;

Figure 14 shows a 4th order optimum beampattern formed with a robustness constraint as well as side lobe control constraints;

Figure 15 shows a 4th order optimum beampattern formed with a robustness constraint and side lobe control, and with a deep null steered to the interference coming from the direction (50,90);

Figure 16 shows an optimum multi-main lobe beampattern formed with six distortionless constraints in the directions of the signals of interest;

Figure 17 shows as optimum multi-main lobe beampattern formed with six distortionless constraints in the directions of the signals of interest, a null formed at (0,0) and side lobe control for the lower hemisphere;

Figure 18 is a flowchart schematically showing the method of the invention and apparatus for carrying out that method;

Figure 19 shows practical implementation of the invention in a teleconferencing scenario;

- 39 -

Figure 20 schematically shows a modal beamformer structure operating in the frequency domain and incorporating a steering unit;

Figure 21 schematically shows a time-domain implementation of a broadband modal beamformer incorporating a steering unit and a number of FIR filters;

Figure 22 shows the performance of a modal beamformer using a maximum robustness design. (a) shows the FIR filters' coefficients, (b) shows the weighting function as a function of frequency for time-domain and frequency-domain beamformers using a maximum robustness design, (c) shows the beampattern as a function of frequency and angle, and (d) shows the DI and WNG at various frequencies;

Figure 23 shows the performance of a time-domain modal beamformer using a maximum directivity design. (a) shows the FIR filters' coefficients, (b) shows the weighting function, (c) shows the beampattern, and (d) shows the DI and WNG at various frequencies;

Figure 24 shows the performance of a beamformer using a robust maximal directivity design;

Figure 25 shows the performance of a beamformer with frequency invariant patterns over two octaves;

Figure 26 shows the performance of a beamformer using multiple-constraint optimization; and

Figure 27 shows some experimental results: (a) the received time series at two typical microphones and the spectrogram of the first one, and the output time series for two various steering directions and the spectrogram of the first one for: (b) TDMR, (c) TDMD, and (d) TDRMD modal beamformers, respectively.

Looking first at Figure 18, a preferred embodiment of the system of the present invention is shown schematically as a beamforming system for a spherical microphone array of M microphones.

Microphones 10 (shown schematically in the figure, but in reality arranged into a spherical array, each receive sound waves from the environment around the array and convert these into

- 40 -

electrical signals. The signals from each of the M microphones are first processed by M preamplifiers and M ADCs (Analog to Digital Convertors) and M calibration filters in stage 11. These signals are then all passed to stage 20 where a Fast Fourier Transform algorithm splits the data into M channels of frequency bins. These are then passed to stage 12 where the spherical Fourier transform is taken. Here, the signals are transformed into the spherical harmonics domain of order N , i.e. spherical harmonic coefficients are generated for each of the $(N + 1)^2$ spherical harmonics of order $n = 0, \dots, N$ and of degree $m = -n, \dots, n$.

The spherical harmonics domain information is passed on to stage 13 for constraint formulation and also to stage 16 for post-optimization beam pattern synthesis. In stage 13, the desired parameters of the system are input from the tunable parameters stage 14. In the figure, the desired parameters which can be input include the look direction of the signal, and the main lobe width (14a), the robustness (14b), desired side lobe levels and side lobe regions (14c), and desired null locations and depths (14d).

Stage 13 takes the desired input parameters for the beampattern, combined with the spherical harmonics domain signal information from stage 12 and formulates these into convex quadratic optimization constraints which are suitable for a convex optimization technique. Constraints are formulated for automatic null-steering, main lobe control, side lobe control and robustness. These constraints are then fed into stage 15 which is the convex optimization solver for performing a numerical optimization algorithm such as an interior point method or second order cone programming and determines the optimum weighting coefficients to be applied to the spherical harmonics coefficients in order to provide the optimum beampattern under the input constraints. Note that in the space domain, the transformation to the spherical harmonics domain is not performed and the optimized weighting coefficients are applied directly to the input signals.

These determined weighting coefficients are then passed to stage 16 which combines the coefficients with the data from stage 12 as a weighted sum and finally a single channel Inverse Fast Fourier Transform is performed in stage 17 to form the array output signal.

Turning now to a practical implementation of the invention. Figure 19 shows the invention being put into effect in a teleconferencing scenario. Two conference rooms 30a and 30b are

- 41 -

shown. Each room is equipped with a teleconferencing system which comprises a spherical microphone array 32a and 32b for voice pick up in three dimensions, and a set of loudspeakers 34a and 34b. Each room is shown with four speakers located in the corners of the room, but it will be appreciated that other configurations are equally valid. Each room is also shown with three speaking persons 36a and 36b situated at various positions around the microphone array. The microphone arrays are connected to a beamformer and an associated controller 38a and 38b which carry out the optimization algorithm in order to generate the optimal beampatterns for the microphone arrays 32a,b.

In operation, consider that one of the speaking persons 34a is talking and everybody else is silent. The controller 38a detects the source signal and controls the beamformer to generate a beamforming pattern for the microphone array 32a in room 30a to form a mainlobe (i.e. an area of high gain) in the direction of the speaking person 36a and to minimise the array gain in all other directions.

In room 30b, the beamformer 38b detects sound sources from each of the loudspeakers 34b as interference sources. It is desirable to minimise sound from these directions in order to avoid a feedback loop between the two rooms.

Now if one of the speaking persons 36b in room 30b starts to talk over the person in room 30a, the beamformer in room 30b must immediately form a mainlobe in that speaking person's direction to ensure that his or her voice is safely transmitted to room 30a. Similarly, the beamformer 38a in room 30a must immediately form deep nulls in the beampattern in the direction of the loudspeakers 34a in order to avoid feedback with room 30b.

As the beamformers 38a and 38b are able to create multiple main lobes and multiple deep nulls and can control the directionality of these in real time, the system does not fail even if one of the speaking persons starts to walk around the room while talking. Unexpected interference, such as a police siren passing by the office can also be taken into account by controlling the directionality of the deep nulls in real time. At the same time, the beamformers 38a and 38b aim to minimise the array output power within the bounds of the applied constraints in order to minimise the influence of general background noise such as the building's air conditioning fans.

- 42 -

This system provides high quality spatial 3D audio with full duplex transmission, noise reduction, dereverberation and acoustic echo cancellation

A. Special cases

We next consider several special cases of the above optimization problem (32) and compare these with the results of previous studies.

Special case 1: Maximum directivity, no WNG or sidelobe control. This is formulated as $\varepsilon = 0$, $\zeta = 0$, $\{\sigma_d^2\}_{d=0}^D = 0$, and $\mathbf{Q}(\omega) = \mathbf{Q}_{iso}(\omega)$ in (24). This gives that $\mathbf{R}(\omega) = \mathbf{Q}_{iso}(\omega)$ and the two inequality constraints in (32) are always inactive and can be ignored.

Since the directivity factor can be interpreted as the array gain against isotropic noise, the optimization problem in this case will result in a maximum directivity factor.

The optimization problem in this case resembles a Capon beamformer in classical array processing, and the solution to (32) is easily derived as:

$$\mathbf{w}(k) = \frac{(4\pi / M) \mathbf{Q}_{iso}^{-1}(\omega) \mathbf{p}(ka, \Omega_0)}{\mathbf{p}^H(ka, \Omega_0) \mathbf{Q}_{iso}^{-1}(\omega) \mathbf{p}(ka, \Omega_0)}. \quad (33)$$

Using (7) and (26), and using the fact that

$$\sum_{n=1}^N \sum_{m=-n}^n Y_n^m(\Omega) Y_n^{m*}(\Omega) = \sum_{n=1}^N \frac{2n+1}{4\pi} = \frac{(N+1)^2}{4\pi}, \quad (34)$$

equation (33) can be further transformed to the following form

$$\mathbf{w}(k) = \frac{(4\pi)^2}{M(N+1)^2} \mathbf{Y}^*(\Omega_0) \circ \mathbf{b}^*(ka), \quad (35)$$

- 43 -

where $\circ /$ denotes element-by-element division, i.e., $w_{nm}(k) = \frac{(4\pi)^2}{M(N+1)^2} \frac{Y_n^{m*}(\Omega_0)}{b_n^*(ka)}$. It can be

seen that the weights in (35) are identical to the weights of a pure phase-mode spherical microphone array (See, for example, B. Rafaely, "Phase-mode versus delay-and-sum spherical microphone array processing", IEEE Signal Process. Lett., vol. 12, no. 10, pp. 713-716, Oct.2005 (also cited in the introduction)) except for a scalar multiplier, which does not affect the array gain.

Using (35) in (31) and (28), gives

$$WNG_1(k) = \frac{M}{(4\pi)^2} \frac{(N+1)^4}{\sum_{n=0}^N |b_n(ka)|^2 (2n+1)}, \quad (36)$$

and

$$D_1(k) = (N+1)^2, \quad (37)$$

(Note that these are identical to (11) and (12), respectively in the Rafaely reference cited above, with $d_n \equiv 1$ there). This result confirms that a pure phase-mode spherical microphone array of order N will have a frequency-independent maximum DI of $20 \log_{10}(N+1)$ dB.

Special case 2: Maximum WNG, no directivity or sidelobe control. This is formulated as $\mathbf{R}(\omega) = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix, $\varepsilon = \infty$, and $\zeta = 0$.

Clearly, the optimization problem in this case results in a minimum norm of the weight vector, or maximum white noise gain.

With $\mathbf{Q}_{iso}$ in (33) replaced by $\mathbf{I}$, the solution in this case is found to be:

$$\mathbf{w}(k) = \frac{(4\pi/M) \mathbf{p}(ka, \Omega_0)}{\mathbf{p}^H(ka, \Omega_0) \mathbf{p}(ka, \Omega_0)} = \frac{(4\pi)^2 \mathbf{b}(ka) \circ \mathbf{Y}^*(\Omega_0)}{M \sum_{n=0}^N |b_n(ka)|^2 (2n+1)}, \quad (38)$$

- 44 -

and

$$w_{nm}(k) = \frac{(4\pi)^2 b_n(ka) Y_n^m(\Omega_0)}{M \sum_{n=0}^N |b_n(ka)|^2 (2n+1)}, \quad (39)$$

which in the case of an open sphere configuration is identical to the weights of a delay-and-sum spherical microphone array except for the scalar multiplier.

Moreover, using (38) in (31) and (28), gives

$$WNG_2(k) = \frac{M}{(4\pi)^2} \sum_{n=0}^N |b_n(ka)|^2 (2n+1), \quad (40)$$

and

$$D_2(k) = \frac{\left| \sum_{n=0}^N |b_n(ka)|^2 (2n+1) \right|^2}{\sum_{n=0}^N |b_n(ka)|^4 (2n+1)}. \quad (41)$$

(Note that this is the same result as in (17) and (18) of the above Rafaely reference).

Since the summation in (40) approaches $(4\pi)^2$ with $N \rightarrow \infty$, the delay-and-sum array achieves a frequency-independent constant WNG equal to M , which is a well-known result in classical array processing.

Special case 3: Control of directivity and WNG, no side lobe control. This case is formulated by the criterion $\varepsilon = \infty$.

The optimization problem in this case has a form resembling a white noise gain constrained (or norm-constrained) robust Capon beamforming problem.

It is straightforward to verify that, in the case when $\zeta = WNG_2$, the corresponding solution is a delay-and-sum array as described in Special Case 2. Furthermore, we find that with

- 45 -

$\mathbf{R}(\omega) = \mathbf{Q}_{iso}(\omega)$ and adjusting the value of ζ in the range $(0, WNG_2]$, we can obtain a trade-off between the pure phase-mode and delay-and-sum spherical array processing.

The following preferred embodiments of the invention are simulations of the beamformer described above, and are used to illustrate and evaluate its performance. In the simulations of Figures 1 to 7 below, we consider an open sphere array of order $N = 4$, and assume that the number of microphones, $M = (N + 1)^2$.

The simulations described herein have all been conducted on consumer-grade computer equipment, e.g. a notebook PC with a CPU speed of 2.4 GHz and with 2 GB of RAM. The simulations were conducted in MATLAB and took around 2 to 5 seconds for each narrowband simulation. It will be appreciated that MATLAB code is a high level programming language designed for mathematical analysis and simulation, and that when the optimization algorithms are implemented in a lower level programming language such as C or an assembly language, or if they are implemented in Field Programmable Gate Arrays, significant increases in speed can be expected.

B. Trade-off between pure phase-mode and delay-and-sum array

Let $\mathbf{R}(\omega) = \mathbf{Q}_{iso}(\omega)$ and $\varepsilon = \infty$. The optimization problem (32) becomes a norm-constrained maximum-DI beamforming problem. The spherical array configuration provides three-dimensional symmetry. Without loss of generality, we assume that the look direction is $\Omega_0 = [0^\circ, 0^\circ]$. For given values of ζ , we solve this optimization problem as a function of ka to get the weight vectors $\mathbf{w}(k)$, and insert them into (28) and (31) to get the DI and WNG, respectively. Fig. 1 and Fig. 2 show the DI and WNG, respectively, as a functions of ka for the case where $\zeta = 0, M/2, M/4$ and WNG_2 . The cases with $\zeta = 0$ and $\zeta = WNG_2$ correspond to the pure phase-mode array and delay-and-sum array, respectively. The cases $\zeta = M/2$ and $\zeta = M/4$ correspond, respectively, to robust beamformers with 3 dB and 6 dB degradation in WNG compared to an ideal maximum WNG of M .

Fig. 2 shows that the norm-constrained beamformer yields a WNG to be above the given threshold values, and thus can provide a good robustness. The DI of two normconstrained beamformers, $\zeta = M/2$ and $M/4$, is much higher than the delay-and-sum beamformer.

- 46 -

Although these DI are smaller than that of a pure phase-mode beamformer, they are obtainable. That of the latter, however, is usually not obtainable due to its extreme sensitivity to even small random array errors encountered in real world applications. In addition, the very low WNG observed for two values at about $ka = 3.14$ and 4.50 in Fig. 2 for the pure phase-mode beamformer is a well-known problem for an open-sphere array, which is avoided by using a rigid-sphere array. In summary, this example demonstrates that the norm-constrained beamforming may provide a useful trade-off between the pure phasemode and delay-and-sum array.

It is also seen that, for the case of $\zeta = M/2$ and $M/4$, the weight vector norm constraint is inactive around $ka = 4$ and 5 . This is due to the fact that around these regions, the pure phase-mode beamformer has already provided a considerable WNG. Therefore, these two beamformers are identical to the pure phase-mode beamformer around these regions.

Fig. 3 shows the DI of the norm-constrained beamformer as a function of WNG at frequencies corresponding to $ka = 1, 2, 3$ and 4 . It is seen that, at higher frequency, the array has a good WNG-DI performance. At the lower frequency, its WNG-DI performance reduces significantly.

The three-dimensional array pattern of three beamformers, i.e., the delay-and-sum beamformer, the pure phase-mode beamformer, and a norm-constrained beamformer with $\zeta = M/4$, have been calculated by (23) for the frequency corresponding to $ka = 3$. These results are displayed in Fig. 4, where we have included a normalization factor $M/4\pi$ so the amplitudes of the patterns at the look direction are equal to unity (or to 0 dB). It is seen that the array patterns in this case are symmetric around the look direction. It's also seen that the norm-constrained beamformer yields a narrower mainlobe than the delay-and-sum beamformer. The values of the DI and WNG of these beamformers are also displayed in the figures. The WNG in Fig. 4(c) is exactly $10\log_{10}(M/4) = 7.96$ dB.

Fig. 5 compares the directivity pattern as a function of elevation θ for the delay-and-sum (DAS) beamformer and norm-constrained beamformer with $\zeta = M/4$, at frequencies corresponding to $ka = 1, 2$, and 4 . It is worth noting that the directivity pattern of the pure phase-mode beamformer is frequency independent and, as suggested by Fig. 2, is identical to that of the norm-constrained beamformer with $\zeta = M/4$ at $ka = 4$.

C. Robust beamforming with interference rejection

Consider the special case 3 described above. The noise is assumed to be isotropic noise. A signal and an interferer are assumed to impinge on the array from $(0^\circ, 0^\circ)$ and $(-90^\circ, 60^\circ)$ with the signal(interferer)-to-noise ratio at each sensor of 0 dB and 30 dB, respectively. We assume that exact covariance is known, and expressed by the theoretical array covariance matrix of $\mathbf{R}(\omega)$ (24).

In this case, the optimization problem becomes a norm-constrained robust Capon beamforming problem and results in a beamformer with high array gain at the expense of some degradation in directivity.

Fig. 6 shows the resulting array patterns for the values of $\zeta = M/4$ and $ka = 3$. As expected, the array patterns have deep null in the direction of arrival of the interferer. The array pattern in this case, unlike those by pure phase-mode beamformer and delay-and-sum beamformer shown in Fig. 4, is no longer symmetric around the look direction.

D. Robust beamforming with sidelobe control and interference rejection

Fig. 4 and Fig. 6 show that the sidelobe levels of these array patterns at $ka = 3$ are about from -13.2 dB to -16.3 dB. Such values may be too high for many applications, leading to severe performance degradation in the case of unexpected or suddenly appearing interferers. For applications in such situations we now consider examples of beamformers with sidelobe control.

We first assume isotropic noise with $\mathbf{R}(\omega) = \mathbf{Q}_{iso}(\omega)$ and take a case where $ka = 3$, $\zeta = M/4$ and $\varepsilon = 0.1$, i.e., the desired sidelobe level is -20 dB. The sidelobe region is defined as $\Omega_{SL} = \{(\theta, \phi) | \theta \geq 45^\circ\}$. The solution of the optimization problem of (32) is the norm-constrained maximum DI beamformer with sidelobe control. The resulting array pattern is shown in Fig. 7(a). The sidelobe level is below -20 dB as specified.

Consider now that in addition to sidelobe control, we want to design a notch around the direction $(60^\circ, 270^\circ)$ with depth of -40 dB and the width of 30° . In this case, the desired sidelobe

- 48 -

structure is direction-dependent. By setting $\varepsilon = 0.01$ in the desired notch region while maintaining $\varepsilon = 0.1$ in the other sidelobe region, and solving the optimization problem, the resulting array pattern is shown in Fig. 7(b). It is seen that the prescribed notch is formed and that the low sidelobe level of -20 dB is maintained.

Consider the scenario described in section C above. Assume that we want to control the sidelobes to be below -20 dB, i.e., $\varepsilon = 0.1$. Keep the other parameters the same as those used in section C. The beamformer weight vector is determined by solving the optimization problem (32). The resulting array pattern is shown in Fig. 7(c). Compared to Fig. 4(a), it is seen that the sidelobes by this method are strictly below -20 dB besides the null in the direction of arrival of the interference.

In the following simulations of a rigid sphere array, with order $N = 4$, multiple mainlobe constraints are applied and non-uniform sidelobe constraints are applied. To form multiple mainlobes in the beampattern, each direction of interest must be made subject to a non-distortion constraint. For non-uniform sidelobe control, instead of requiring all sample points in the sidelobe region to be below a given threshold, sidelobe directions can each be subjected to different thresholds. For example, an interference direction can be subjected to a stronger constraint while the remaining directions can be subjected to a less strong threshold. With these extra constraints (for K mainlobe constraints and L sidelobe constraints), the optimization problem (32) can be restated as:

$$\begin{aligned}
 & \min_{\mathbf{w}} \mathbf{w}^H(k) \mathbf{R}(k) \mathbf{w}(k), \\
 & \text{subject to } \mathbf{w}^H(k) \mathbf{p}(ka, \Omega_k) = 4\pi / M, \quad k = 1, \dots, K, \\
 & \quad \left| \mathbf{w}^H(k) \mathbf{p}(ka, \Omega_{sl,l}) \right| \leq \varepsilon_l \cdot 4\pi / M, \quad l = 1, \dots, L, \\
 & \quad \|\mathbf{w}(k)\| < \sqrt{\frac{4\pi}{M\zeta(k)}}.
 \end{aligned} \tag{42}$$

Again, due to the nature of this optimization formulation, convex optimization techniques can be applied, in particular as it is a convex second order cone problem, SOCP techniques can be used to solve it. With these techniques, even with the large number of constraints involved, the problem can still be optimized efficiently and in real time.

Further simulations are used to evaluate the performance of this beamformer. We consider a rigid sphere array of order $N = 4$, and $M = (N + 1)^2$. We assume that the look direction is $[0^\circ, 0^\circ]$ for a single mainlobe case, $ka = 3$, signal and interferer to noise ratios at each sensor are 0 dB and 30 dB, and a WNG constraint is set to 8 dB. Fig. 8(a) shows the array pattern with sidelobe region defined as $\Omega_{SL} = \{(\theta, \phi) \mid \theta \geq 45^\circ\}$ and sidelobe level below -20 dB. Fig. 8(b) shows the performance of non-uniform sidelobe control; a notch around the direction $(60^\circ, 270^\circ)$ with a depth of -40 dB and a width of 30° is formed, and the remaining sidelobe level is still maintained at -20 dB.

In Fig. 9(a), we assume two interferences impinge on array from $(60^\circ, 190^\circ)$ and $(90^\circ, 260^\circ)$, then it is seen that the nulls are automatically formed and steered to the direction of arrival of the interferences with sidelobes strictly below -20 dB. Fig. 9(b) shows the performance of multi-mainlobe formation and automatic multi-null steering with -20 dB sidelobe control, here we assume two desired signals incident on array from $(40^\circ, 0^\circ)$ and $(40^\circ, 180^\circ)$, with three interferences impinging from $(0^\circ, 0^\circ)$, $(45^\circ, 90^\circ)$, and $(50^\circ, 270^\circ)$. Actual directivity index (DI) and WNG values are also calculated for Figures 8 and 9.

In the following analysis, we consider a compact spherical microphone array placed in a room. All signal sources are assumed to be located in the far field of the aperture (so that they may be approximated by plane waves incident on the array), and the early reflections in the room are modelled as point sources while the late reverberation is modelled as isotropic noise. Now we assume that $L + D$ source signals impinge on the sphere from directions $\Omega_1, \Omega_2, \dots, \Omega_L, \Omega_{L+1}, \dots, \Omega_{L+D}$, and in addition noise is present. Then the space domain sound pressure for each microphone position can be written as:

$$\begin{aligned}
 x(ka, \Omega_s) = & \sum_{l=1}^L \left[p(ka, \Omega_l, \Omega_s) S_l(\omega) + \sum_{lr=1}^R p(ka, \Omega_{lr}, \Omega_s) \alpha_{lr} S_{lr}(\omega) \exp(i\omega\tau_{lr}) \right] \\
 & + \sum_{d=1}^D \left[p(ka, \Omega_d, \Omega_s) S_d(\omega) + \sum_{dr=1}^R p(ka, \Omega_{dr}, \Omega_s) \alpha_{dr} S_{dr}(\omega) \exp(i\omega\tau_{dr}) \right] \\
 & + N(\omega, \Omega_s), \quad s = 1, 2, \dots, M,
 \end{aligned} \tag{43}$$

- 50 -

where $\{S_a(\omega)\}_{a=1}^{L+D}$ are the $L + D$ source signal spectrums, $\{S_{lr}(\omega)\}_{lr=1}^R$ and $\{S_{dr}(\omega)\}_{dr=1}^R$ are their R early reflections, α and τ denote the attenuation and propagation time of early reflections, and $N(\omega, \Omega_s)$ is the additive noise spectrum. The first term in (43) corresponds to the L desired signals that it is desired to capture, and the second term in (43) corresponds to D interferences.

The spherical Fourier transform of $x(ka, \Omega_s)$ is given by

$$\begin{aligned} x_{nm}(ka) = & \sum_{l=1}^L \left[p_{nm}(ka, \Omega_l) S_l(\omega) + \sum_{lr=1}^R p_{nm}(ka, \Omega_{lr}) \alpha_{lr} S_{lr}(\omega) \exp(i\omega\tau_{lr}) \right] \\ & + \sum_{d=1}^D \left[p_{nm}(ka, \Omega_d) S_d(\omega) + \sum_{dr=1}^R p_{nm}(ka, \Omega_{dr}) \alpha_{dr} S_{dr}(\omega) \exp(i\omega\tau_{dr}) \right] \\ & + N_{nm}(\omega), \quad n = 0, 1, \dots, N, \quad m = [-n, n], \end{aligned} \quad (44)$$

where $N_{nm}(\omega)$ is the spherical Fourier transform of noise, a N is the spherical harmonics order which satisfies $M \geq (N + 1)^2$ as before.

Array processing can then be performed in either the space domain or the spherical harmonics domain, and the array output $y(ka)$ is calculated as

$$y(ka) = \sum_{s=1}^M \alpha_s x(ka, \Omega_s) w^*(k, \Omega_s) = \sum_{n=0}^N \sum_{m=-n}^n x_{nm}(ka) w_{nm}^*(k), \quad (45)$$

As before, α_s depends on the sampling scheme. For uniform sampling, $\alpha_s = 4\pi / M$.

As with embodiments, in the beamformer of the following embodiments, multiple mainlobe directions are maintained and the sidelobe levels are controlled, while the array output power is minimized in order to adaptively suppress the interferences coming from outside beam directions. Furthermore, for the purpose of improving system robustness, a weight norm constraint (i.e. white noise gain control) is also applied to limit the norm of array weights to a chosen threshold.

- 51 -

To ensure that the L desired signals coming from directions $\Omega_l = \Omega_1, \Omega_2, \dots, \Omega_L$, will be well captured and equalized, we define a $L \times (N+1)^2$ manifold matrix

$$\tilde{\mathbf{P}}_{nm} = [\mathbf{p}(ka, \Omega_1), \mathbf{p}(ka, \Omega_2), \dots, \mathbf{p}(ka, \Omega_L)]^T$$

and a $L \times 1$ vector column containing L desired mainlobe levels

$$\mathbf{A} = [A_1 \cdot 4\pi / M, A_2 \cdot 4\pi / M, \dots, A_L \cdot 4\pi / M]^T$$

where $4\pi / M$ is the normalization factor. Then the problem of multi-beam forming with tractable mainlobe levels can be formulated as a single linear equality constraint:

$$\tilde{\mathbf{P}}_{nm} \mathbf{w}(k) = \mathbf{A}, \quad (46)$$

and the levels for L mainlobe responses can be controlled by setting different A values. This becomes particularly useful in the simple application of equalization of the voice amplitudes of L desired speakers, who have different speech levels. This occurs mainly due to the fact that they sit at different positions in the room.

Similarly to the above description of the embodiments, in order to guarantee all sidelobes strictly below given threshold values ε_j , we can formulate a set of quadratic inequality constraints

$$|\mathbf{p}^H(ka, \Omega_{SL,j}) \mathbf{w}(k)|^2 \leq \varepsilon_j \cdot (4\pi / M)^2, \quad j = 1, 2, \dots, J, \quad (47)$$

where $\Omega_{SL,j}$ denote the sidelobe regions, and they are also utilized to control the beam widths of the multiple mainlobes.

As in the above embodiments, adaptive mainlobe formation and multi-null steering is achieved by minimizing the array output power in run time while applying various constraints. As stated before in (22), the array output power is given by

- 52 -

$$P_0(\omega) = E[\|y(ka)\|^2] = \mathbf{w}^H(k) \mathbf{R}(\omega) \mathbf{w}(k) = \|\mathbf{R}(\omega)^{1/2} \mathbf{w}(k)\|^2, \quad (48)$$

where $E[\cdot]$ denotes the statistical expectation, and $\mathbf{R}(\omega)$ denotes the covariance matrix of $\mathbf{x}$. For simplification, we assume that the early reflections in the room are much lower than direct sound, so that $\mathbf{R}(\omega)$ has the form

$$\mathbf{R}(\omega) = \sum_{a=1}^{L+D} \mathbf{R}_a(\omega) + \mathbf{R}_n(\omega), \quad (49)$$

where $\mathbf{R}_a(\omega)$ is the signal covariance matrix corresponding to the a th signal, and $\mathbf{R}_n(\omega)$ is the noise covariance matrix.

Now, by introducing a variable ξ , the optimization problem can be reformulated as

$$\min_{\mathbf{w}} \xi, \text{ subject to } \|\mathbf{R}(\omega)^{1/2} \mathbf{w}(k)\| \leq \xi. \quad (50)$$

The weight vector norm constraint derived previously in (31) for a single mainlobe also applies to the multi-mainlobe case since it controls the dynamic range of array weights to avoid large noise amplification at the array output.

Combining this with (46), (47) and (50), the optimization problem of (32) can be expressed as

$$\begin{aligned} & \min_{\mathbf{w}} \xi \\ & \text{subject to } \|\mathbf{R}(\omega)^{1/2} \mathbf{w}(k)\| \leq \xi \\ & \tilde{\mathbf{P}}_{nm} \mathbf{w}(k) = \mathbf{A} \\ & \|\mathbf{w}(k)\| < \sqrt{\delta \frac{4\pi}{M}} \\ & |\mathbf{p}^H(ka, \Omega_{SL,j}) \mathbf{w}(k)|^2 \leq \varepsilon_j \cdot (4\pi/M)^2, \quad j = 1, 2, \dots, J. \end{aligned} \quad (51)$$

- 53 -

Thus a single optimization problem has been formulated which accomplishes multiple mainlobe formation with different mainlobe levels, sidelobe control with multiple null formation and steering and a robustness constraint. Further, this optimization problem is a convex second order cone optimization problem and can therefore be solved efficiently using, second order cone programming, in real time.

It will be noted in the above that the weight vector norm constraint has been expressed with the threshold constant δ in the numerator rather than ζ in the denominator. The following simulations indicate values of δ which have been used.

In the following simulations, consider a rigid sphere with $r = 5$ cm is sampled by $M = (N + 1)^2$ microphones, and $ka = 3$. Signal and interferer to noise ratios at each microphone are 0 dB and 30 dB. A uniform grid of 5° is used to discretize the sidelobe region. Unless otherwise stated, the theoretical data covariance matrix $\mathbf{R}(\omega)$ is used in adaptive beamforming examples for convenience.

For single beam cases ($L = 1$), assume order $N = 4$, $A_1 = 1$, the look direction is $[0^\circ, 0^\circ]$, and the WNG constraint is set to 8 dB ($\delta = 0.159$). Fig. 10(a) shows the regular single beam pattern synthesis using (51) without sidelobe control and adaptive null steering constraints. Fig. 10(b) shows the performance of nonuniform sidelobe control. The main sidelobe region is defined as $\Omega_{sl} = \{(\theta, \phi) | \theta \geq 45^\circ\}$ with sidelobe level uniformly below -20 dB ($\varepsilon_j = 0.01$), while defining a notch around the direction $(60^\circ, 270^\circ)$ with depth of -40 dB ($\varepsilon_j = 0.0001$) and the width of 30° . In Fig. 11(a), remove the notch, and assume two interferences impinge on array from $[60^\circ, 190^\circ]$ and $[90^\circ, 260^\circ]$, then it is seen that the nulls are automatically formed and steered to the direction of arrival of the interferences with sidelobes still strictly below -20 dB. Note that actual WNG and directivity index (DI) values are calculated for all the single beam cases.

It is seen that in Fig. 10(b), the mainlobe becomes a little wider, and DI is also 0.3 dB lower than that without sidelobe control. However these costs are acceptable in practical applications. The reason for degradation is that the beamforming performance parameters, i.e., the beamwidth, sidelobe level, DI, and robustness are all mutually correlated. The algorithm illustrated herein provides a suitable compromise among these conflicting objectives.

For multi-beam examples ($L = 3$), we use an array order of $N = 5$ to obtain more degrees of freedom. Assume three desired signals incident on array from $[60^\circ, 0^\circ]$, $[60^\circ, 120^\circ]$ and $[60^\circ, 240^\circ]$. Fig. 11(b) shows the multi-beam forming performance with $A_{1,2,3} = 1$ and $\delta = 0.4$. Fig. 12(a) shows the acceptable performance of multi-beam with adaptive null steering and -20 dB sidelobe control, assuming that interferences come from $[0^\circ, 0^\circ]$, $[65^\circ, 60^\circ]$, $[65^\circ, 180^\circ]$, and $[65^\circ, 300^\circ]$. Next, suppose that the amplitude of the second desired signal is 6 dB lower than the other two signals, and we can just set $A_2 = 2$ and $\delta = 1$, to simply equalize the sound levels. The beam pattern is shown in Fig. 12(b), and shows that we obtain around 6 dB amplitude enhancement for signals coming from the second mainlobe direction.

Figures 13 to 17 show further simulations which illustrate the benefits of the optimal beamformer of the present invention. Figure 13 shows a 4th order regular beampattern formed with a robustness constraint, but with no side lobe control. By contrast, Figure 14 shows a 4th order optimum beampattern obtained according to the invention, formed with a robustness constraint as well as side lobe control constraints. The main lobe is in the region of 45 degrees from the positive z-axis. Figure 15 shows a 4th order optimum beampattern formed in accordance with the invention, with a robustness constraint and side lobe control, and with a deep null steered to the interference coming from the direction (50,90).

Figure 16 shows an optimum multi-main lobe beampattern formed in accordance with the invention with six distortionless constraints in the directions of the signals of interest, thus forming six main lobes in the beampattern. Figure 17 shows an optimum multi-main lobe beampattern formed in accordance with the invention, with six distortionless constraints in the directions of the signals of interest, with a null formed at (0,0) and side lobe control for the lower hemisphere.

Time domain examples

The following provides several numerical examples to illustrate the performances of the time domain approach to array pattern synthesis for a broadband modal beamformer.

In the examples considered below, we consider a rigid spherical array of radius 4.2 cm with

- 55 -

$M = 32$ microphones located at the center of the faces of a truncated icosahedron. An order of $N = 4$ is used for sound field decomposition and $\alpha_s \equiv 4\pi / M$. The sampling frequency is $f_s = 14700$ Hz. The frequency band $[f_L, f_U]$ is discretized using $K = 51$ frequency grids $f_k = f_L \cdot 10^{\lg(f_U / f_L) \cdot (k-1) / (K-1)}$, $k = 1, 2, \dots, K$. The length of the FIR filters is $L = 65$. Unless otherwise stated, we assume $\Theta_{ML} = [0^\circ : 2^\circ : 40^\circ]$ and $\Theta_{SL} = [48^\circ : 2^\circ : 180^\circ]$, which means a uniform grid of 2° is used to discretize the directions.

T.A. Maximum robustness design

Referring to equation (T42), assume that $f_L = 500$ Hz, $f_U = 5000$ Hz. Let $\ell = 4$, $\mu_1 = \infty$, $\mu_2 = \infty$, $\mu_3 = \infty$. The optimization problem becomes

$$\min_{\mathbf{h}} \mathbf{h}^T \mathbf{h}, \text{ subject to } \mathbf{h}^T \mathbf{u}(f_k, 0) = 4\pi / M, \quad k = 1, 2, \dots, K. \quad (\text{T43})$$

A solution of this problem is called a *time-domain Maximum-Robust (TDMR) modal beamformer*. The FIR filter $\mathbf{h}$ is determined by solving the optimization problem (T43) and its subvectors $\mathbf{h}_0, \mathbf{h}_1, \dots, \mathbf{h}_N$ are shown in Fig. 22(a). We substitute $\mathbf{h}$ into (T23) to get $\hat{c}_n(f)$ and display them in Fig. 22(b). For comparison purposes, $[c_n(f_k)]_{MWNG}$, which are calculated using (T17), are also shown in this figure. It is seen that the weights of the time-domain Maximum-Robust modal beamformer, $\hat{c}_n(f)$, approximate that of the frequency-domain Maximum-WNG modal beamformer, $[c_n(f_k)]_{MWNG}$, within the frequency band $[f_L, f_U]$.

Using (T25), the beampattern as a function of frequency and angle are calculated on a grid of points in frequency and angle. The resulting beampatterns are shown in Fig. 22(c), where we have included a normalization factor $M / 4\pi$ so the amplitudes of the patterns at the look direction are equal to unity (or to 0 dB).

The DI and WNG of the are calculated by using (T38) and (T15), respectively. The DI and WNG of the frequency-domain Maximum-WNG modal beamformer are also calculated for comparison purposes. The results are shown in Fig. 22(d) for various frequencies.

T.B. Maximum directivity design

Let $\ell = 1$, $\mu_2 = \infty$, $\mu_3 = \infty$, $\mu_4 = \infty$. The optimization problem (T42) becomes a maximum directivity design problem. The resulting beamformer is referred to as *time-domain Maximum-directivity (TDMD) modal beamformer*.

Assume that $f_L = 500$ Hz, $f_U = 5000$ Hz. The resulting FIR filters $\mathbf{h}_0, \mathbf{h}_1, \dots, \mathbf{h}_N$, the weighting function $\hat{c}_n(f)$, the beampatterns, and the DI and WNG are shown in Fig. 23(a),(b),(c), and (d), respectively. For comparison purposes, the weights function $[c_n(f_k)]_{MDI}$ (T16), and DI and WNG of the frequency-domain Maximum-DI modal beamformer, are also shown in the figures. It is seen that the weights of the time-domain modal beamformer using maximum directivity design approximate that of its frequency-domain counterpart within the frequency band $[f_L, f_U]$.

As compared to Fig. 22(a), (b) and (d), it is seen that the coefficients of the FIR filters and thus the resulting weighting function of the TDMD beamformer are quite large and the WNG at low frequency is too small, all imply that this beamformer lacks robustness.

T.C. Maximal directivity with robustness control

In order to improve the robustness of the beamformer, the broadband white noise gain constraint should be imposed. This can be formulated as $\ell = 1$, $\mu_2 = \infty$, $\mu_3 = \infty$, and μ_4 is a user parameter. The resulting beamformer is referred to as *time-domain Robust Maximal-directivity (TDRMD) modal beamformer*.

Assume that $f_L = 500$ Hz, $f_U = 5000$ Hz, and $\mu_4 = 4\pi / M$. The resulting FIR filters $\mathbf{h}_0, \mathbf{h}_1, \dots, \mathbf{h}_N$, the weighting function $\hat{c}_n(f)$, the beampatterns, and the DI and WNG are shown in Fig. 24(a),(b),(c), and (d), respectively.

It is seen from Fig. 24(d) that the WNG of this beamformer is higher than -3 dB, which at low frequency is much higher than that of the maximum directivity design as shown in Fig. 23. The DI of this beamformer is much higher than that of the maximum robustness design as shown in

Fig. 22. Hence, the results show that this design provides a good tradeoff between the directivity and the robustness.

T.D. Frequency-invariant beamformer

Assume that we want to synthesize a frequency-independent broadband beampattern. We reduce the bandwidth to two octaves so that $f_L = 1250$ Hz, $f_U = 5000$ Hz. Let $\ell = 1$,

$\mu_2 = 10^{-1.5} \cdot 4\pi / M$, $q_1 = 2$, $\mu_3 = \infty$, $\mu_4 = 2\pi / M$, $\Theta_{ML} = [0^\circ : 2^\circ : 180^\circ]$. The results are shown in Fig. 25. It is seen that the expected frequency-independent beampatterns are obtained, and the WNG is moderate.

T.E. Optimal beamformer with multiple constraints

Assume that $f_L = 1250$ Hz, $f_U = 5000$ Hz. Let $\ell = 1$, $\mu_2 = 0.1 \cdot 4\pi / M$, $q_1 = 2$,

$\mu_3 = 10^{-14/20} \cdot 4\pi / M$, $q_2 = \infty$, $\mu_4 = 10^{-4/10} \cdot 4\pi / M$, $\Theta_{ML} = [0^\circ : 2^\circ : 40^\circ]$ and

$\Theta_{SL} = [48^\circ : 2^\circ : 180^\circ]$. The resulting results are shown in Fig. 26. It is seen that all the constraints are guaranteed and the trade-off among multiple performance measures are obtained.

Experimental results

The Eigenmike® microphone array from MH Acoustics was employed, which is a rigid spherical array of radius 4.2 cm with 32 microphones located at the center of the faces of a truncated icosahedron. The experiment was conducted in an anechoic room which is anechoic down to 75Hz, and the Eigenmike® was placed in the center of the room for recording. A loudspeaker, which was located 1.5 meters away from the Eigenmike® roughly in the direction $(20^\circ, 180^\circ)$, was used to play a swept-frequency cosine signal (ranging from 100 Hz to 5 kHz). The sound was recorded by the Eigenmike® with the sampling frequency of 14.7 kHz and 16 bit per sample.

The signals received at two typical microphones (i.e., No. 13 microphone that on the sunny side and No. 31 microphone that on the dark side) are respectively shown in the upper and lower plot of Fig. 27(a). The spectrogram of the signal shown in the upper plot using short-time Fourier

transform is shown in the middle plot.

The TDMR modal beamformer presented in subsection *T.A.* is used. When the beam is steered to the direction of arrival, i.e., $(20^\circ, 180^\circ)$, the beamformer output time series and the spectrogram are shown in the upper and middle plot of Fig. 27(b), respectively. The lower plot of Fig. 27(b) shows the output time series when the beam is steered to another direction $(80^\circ, 180^\circ)$, which is 60° away from the direction of arrival.

We apply the TDMD and TDRMD modal beamformer presented in subsection *T.B.* and *T.C.* to the same microphone array data, respectively. We repeat the process above, the same results as in Fig. 27(b) for the two methods are shown in Fig. 27(c) and (d), respectively.

We look at the upper plots of Fig. 27(b), (c) and (d). It is seen that the output of the TDMRD beamformer is similar as that of the TDMR beamformer. For the TDMD beamformer, however, its magnitude at the lower frequency is much larger. The reason is that the norm of the weights at the lower frequency is very large and leads to a quite large output even to slight mismatches between the presumed and actual array response vectors. In other words, this beamformer is quite sensitive even to slight mismatches.

Comparing the lower plot of Fig. 27(b) with that of Fig. 27(d), it is noted that the magnitude of the time series of the TDMR beamformer is much larger than that of the TDRMD beamformer, especially at the lower frequency, which means that the beamwidth of the former is wide than the latter. This can also be found from the beampatterns shown in Fig. 22 and Fig. 24. Hence, the results presented in Fig. 27 show that the TDRMD beamformer provides a good trade-off between the directivity and the robustness.

The above examples have presented the real-valued time-domain implementation of the broadband modal beamformer in the spherical harmonics domain. The broadband modal beamformer in these examples is composed of the modal transformation unit, the steering unit, and the pattern generation unit, although it will be understood that the steering unit is optional and can be omitted if it is necessary to generate a beam pattern which is not rotationally symmetric about the look direction. The pattern generation unit is independent of the steering direction and is implemented using filter-and-sum structure. The elegant spherical harmonics

- 59 -

framework leads to a more computationally efficient optimization algorithm and implementation scheme than conventional element-space based approaches. The broadband array response, the beamformer output power against both isotropic noise and spatially white noise, and the mainlobe spatial response variation have all been expressed as functions of the FIR filters' tap weights. The FIR filters design problem has been formulated as a multiply-constrained problem, which ensures that the resulting beamformer can provide a suitable trade-off among multiple conflicting array performance measures such as directivity, mainlobe spatial response variation, sidelobe level, and robustness.

It can be seen from all of the above that the problem of optimal beamformer design for spherical microphone arrays has been addressed by formulating the optimization problem as a multiple-constrained convex optimization problem which can be solved efficiently using a Second Order Cone Programming solver. It has been demonstrated that the resulting beamformer can provide a suitable trade-off among multiple performance measures such as directivity index, robustness, array gain, sidelobe level, mainlobe width, and so on as well as providing for multiple mainlobe formation multiple adaptive null forming for interference rejection, both with varying gain constraints for different lobes / regions. It is evident that the approach provides a flexible design tool since it covers the previously studied delay-and-sum beamformer, and the pure phase-mode beamformer as special cases, while also allowing far more complex optimization problems to be solved within the allowable timeframe.

- 60 -
Annex

The following section is some background description of spherical Fourier transforms and spherical-harmonics based beamforming and it derives some results which have been used in this description.

The standard Cartesian (x, y, z) and spherical (r, θ, ϕ) coordinate systems are used. Here, elevation θ and azimuth ϕ are angular displacements in radians measured from the positive z -axis and x -axis of the projection onto the plane $z = 0$, respectively. Consider a unit magnitude plane wave impinging on a sphere of radius a from direction $\Omega_0 = (\theta_0, \phi_0)$ and with a time factor $\exp(i\omega t)$ which is suppressed throughout this application. Here, $i = \sqrt{-1}$, and ω is the temporal radian frequency.

The total sound pressure on the sphere surface at an observation point (a, Ω_s) for a wavenumber k can be written using spherical harmonics as

$$p(ka, \Omega_0, \Omega_s) = \sum_{n=0}^{\infty} b_n(ka) \sum_{m=-n}^n Y_n^{m*}(\Omega_0) Y_n^m(\Omega_s), \quad (1)$$

where $k = \|\mathbf{k}\| = \omega / c$ with c being the sound speed, Y_n^m is the spherical harmonics of order n and degree m , superscript $*$ denotes complex conjugation, and $b_n(ba)$ depends on the sphere configuration, e.g. rigid sphere, open sphere, etc., as given by

$$b_n(ka) = \begin{cases} 4\pi i^n j_n(ka) & \text{open sphere} \\ 4\pi i^n \left(j_n(ka) - \frac{j'_n(ka)}{h'_n(ka)} h_n(ka) \right) & \text{rigid sphere} \end{cases}, \quad (2)$$

where j_n and h_n are the n th order spherical Bessel and Hankel functions, and j'_n and h'_n are their derivatives with respect to their arguments, respectively.

The spherical harmonics are the solutions to the wave equation, or the Helmholtz equation in spherical coordinates. They are given by

$$Y_n^m(\Omega) = Y_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos\theta) e^{im\phi}, \quad (3)$$

where $P_n^m(\cos\theta)$ denotes the associated Legendre function. The spherical harmonics functions are orthonormal and satisfy

$$\int_{\Omega \in S^2} Y_{n'}^{m'}(\Omega) Y_n^{m*}(\Omega) d\Omega = \delta_{n-n'} \delta_{m-m'}, \quad (4)$$

where $\delta_{n-n'}$ and $\delta_{m-m'}$ are the Kronecker delta functions and the integral

$$\int_{\Omega \in S^2} d\Omega = \int_0^{2\pi} \int_0^\pi \sin\theta d\theta d\phi \text{ covers the entire surface of the unit sphere } S^2.$$

The spherical harmonics decomposition, or the spherical Fourier transform of a squared integrable function p on the unit sphere, denoted by p_{nm} , and the inverse transform, are given by

$$p_{nm}(ka, \Omega_0) = \int_{\Omega \in S^2} p(ka, \Omega_0, \Omega) Y_n^{m*}(\Omega) d\Omega, \quad (5)$$

$$p(ka, \Omega_0, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n p_{nm}(ka, \Omega_0) Y_n^m(\Omega). \quad (6)$$

Applying the spherical Fourier transform (5) to a plane wave as expressed by (1) gives the spherical harmonics domain expression of $p(ka, \Omega_0, \Omega)$:

$$p_{nm}(ka, \Omega_0) = b_n(ka) Y_n^{m*}(\Omega_0). \quad (7)$$

Now, to analyze the properties of a spherical array, we assume a signal-of-interest (SOI) plane wave from direction Ω_0 , and D interference plane waves from directions $\Omega_1, \dots, \Omega_d, \dots, \Omega_D$ that impinge on the sphere. Adding uncorrelated noise, the sound pressure on the sphere surface can be written as:

- 62 -

$$x(ka, \Omega_s) = \beta p(ka, \Omega_0, \Omega_s) S_0(\omega) + \sum_{d=1}^D p(ka, \Omega_d, \Omega_s) S_d(\omega) + N(\omega), \quad (8)$$

where $\{S_d(\omega)\}_{d=0}^D$ are the $D+1$ source signals spectra, $N(\omega)$ is the additive noise spectrum, and β is a binary parameter that indicates whether the SOI is present or not.

The spherical Fourier transform of $x(ka, \Omega_s)$ is given by

$$\begin{aligned} x_{nm}(ka) &= \int_{\Omega \in S^2} x(ka, \Omega) Y_n^{m*}(\Omega) d\Omega \\ &= \int_{\Omega \in S^2} [\beta p(ka, \Omega_0, \Omega_s) S_0(\omega) + \sum_{d=1}^D p(ka, \Omega_d, \Omega_s) S_d(\omega) + N(\omega)] Y_n^{m*}(\Omega) d\Omega \\ &= \beta p_{nm}(ka, \Omega_0) S_0(\omega) + \sum_{d=1}^D p_{nm}(ka, \Omega_d) S_d(\omega) + N_{nm}(\omega), \end{aligned} \quad (9)$$

where $N_{nm}(\omega) = \int_{\Omega \in S^2} N(\omega) Y_n^{m*}(\Omega) d\Omega$ denotes the spherical Fourier transform of noise.

Array processing can be carried out in either the space domain or the spherical harmonics domain, respectively by calculating the integral of the product of the array input signal and the array weight function over the entire sphere, or by a similar weighting and summation in the spherical harmonics domain. Denoting the aperture weighting function by w , the array output is given as the integral of the product between array input signal and the complex conjugated weighting function w^* over the entire sphere,

$$y(ka) = \int_{\Omega \in S^2} x(ka, \Omega) w^*(k, \Omega) d\Omega = \sum_{n=0}^{\infty} \sum_{m=-n}^n x_{nm}(ka) w_{nm}^*(k), \quad (10)$$

where w_{nm} are the spherical Fourier transform coefficients of w . Note that the summation term in (10) can be viewed as weighting in the spherical harmonics domain, also called phase-mode processing.

- 63 -

In practice, the sound pressure is spatially sampled at the microphone positions $\Omega_s, s = 1, \dots, M$, where M is the number of microphones. We require that the microphone positions fulfil the following discrete orthonormality condition:

$$\sum_{s=1}^M \alpha_s Y_{n'}^{m'}(\Omega_s) Y_n^{m*}(\Omega_s) = \delta_{n-n'} \delta_{m-m'}, \quad (11)$$

where α_s depends on the sampling scheme. For uniform sampling, in order that

$$\sum_{s=1}^M \alpha_s = \int_{\Omega \in S^2} d\Omega = 4\pi, \text{ we have } \alpha_s \equiv 4\pi / M. \text{ It will be appreciated that alternative spatial}$$

sampling schemes for the positioning of microphones on a sphere are equally valid.

Note that with a finite number of microphones sampling the sphere, the spherical harmonic order N is required to satisfy $M \geq (N+1)^2$ in order to avoid spatial aliasing. In other words, for a given order N , the number of microphones M must be at least $(N+1)^2$.

The discrete spherical Fourier transform (spherical Fourier coefficients) of $x(ka, \Omega_s)$, and the inverse transform, are given by

$$x_{nm}(ka) = \sum_{s=1}^M \alpha_s x(ka, \Omega_s) Y_n^m(\Omega_s), \quad (12)$$

$$x(ka, \Omega_s) = \sum_{n=0}^N \sum_{m=-n}^n x_{nm}(ka) Y_n^m(\Omega_s). \quad (13)$$

To simplify the analysis, in this paper, we assume that the spatial sampling by microphones is perfect and that the aliasing is negligible, thus $\alpha_s \equiv 4\pi / M$.

The corresponding array output $y(ka)$ can be calculated by:

$$y(ka) = \sum_{s=1}^M \alpha_s x(ka, \Omega_s) w^*(k, \Omega_s) = \sum_{n=0}^N \sum_{m=-n}^n x_{nm}(ka) w_{nm}^*(k), \quad (14)$$

- 64 -

where $w^*(k, \Omega_s)$ are the array weights and $w_{nm}^*(k)$ are their spherical Fourier coefficients. Note that, in the case of ideal uniform sampling, the array output amplitude in (14) is the factor

$4\pi / M$ higher than the classical array processing, which is $\sum_{s=1}^M x(ka, \Omega_s) w^*(k, \Omega_s)$.

By using Parseval's relation for the spherical Fourier transform to the weights, we have

$$\sum_{s=1}^M \alpha_s |w(k, \Omega_s)|^2 = \sum_{n=0}^N \sum_{m=-n}^n |w_{nm}(k)|^2, \quad (15)$$

which indicates the factors α_s .

Claims

1. A method of forming a beampattern in a beamformer of the type in which the beamformer receives input signals from a sensor array, decomposes the input signals into the spherical harmonics domain, applies weighting coefficients to the spherical harmonics and combines them to form an output signal, wherein the weighting coefficients are optimized for a given set of input parameters by convex optimization.
2. A method as claimed in claim 1, wherein the sensor array is a spherical array in which the sensors positions are located on a notional spherical surface.
3. A method as claimed in claim 2, wherein the sensor array is of a form selected from the group of: an open sphere array, a rigid sphere array, a hemisphere array, a dual open sphere array, a spherical shell array, and a single open sphere array with cardioid microphones.
4. A method as claimed in claim 1, 2 or 3, wherein the array is designed for voice band applications and has a largest dimension of about 8 cm to about 30 cm.
5. A method as claimed in any preceding claim, wherein the sensor array is a microphone array.
6. A method as claimed in any preceding claim, wherein the optimization problem, and optionally also constraints, are formulated as one or more of: minimising the output power of the array, minimising the sidelobe level, minimising the distortion in the mainlobe region and maximising the white noise gain.
7. A method as claimed in any preceding claim, wherein the optimization problem is formulated as minimising the output power of the array.
8. A method as claimed in any preceding claim, wherein the input parameters include a requirement that the array gain in a specified direction be maintained at a given level, so as to form a main lobe in the beampattern.

- 66 -

9. A method as claimed in claim 8, wherein the input parameters include requirements that the array gain in a plurality of specified directions be maintained at a given level, so as to form multiple main lobes in the beam pattern.
10. A method as claimed in claim 9, wherein individual required gain levels are provided for each of the plurality of specified directions, so as to form multiple main lobes of different levels in the beam pattern.
11. A method as claimed in claim 8, 9 or 10, wherein the beamformer formulates the or each requirement as a convex constraint.
12. A method as claimed in claim 11, wherein the beamformer formulates the or each requirement as a linear equality constraint.
13. A method as claimed in claim 12, wherein the beamformer formulates the or each requirement as a requirement that the array output for a unit magnitude plane wave incident on the array from the specified direction is equal to a predetermined constant.
14. A method as claimed in any preceding claim, wherein the input parameters include a requirement that the array gain in a specified direction is below a given level, so as to form a null in the beam pattern.
15. A method as claimed in any claim 14, wherein the input parameters include requirements that the array gain in a plurality of specified directions is below a given level, so as to form multiple nulls in the beam pattern.
16. A method as claimed in claim 15, wherein individual maximum gain levels are provided for each of the plurality of specified directions, so as to form multiple nulls of different depths in the beam pattern.
17. A method as claimed in claim 14, 15 or 16, wherein the beamformer formulates the or each requirement as a convex constraint.

- 67 -

18. A method as claimed in claim 17, wherein the beamformer formulates the or each requirement as a second order cone constraint.
19. A method as claimed in claim 18, wherein the beamformer formulates the or each requirement as a requirement that the magnitude of the array output for a unit magnitude plane wave incident on the array from the specified direction is less than a predetermined constant.
20. A method as claimed in any preceding claim, wherein the input parameters include a requirement that the beampattern has a specified level of robustness.
21. A method as claimed in claim 20, wherein the level of robustness is specified as a limitation on a norm of a vector comprising the weighting coefficients.
22. A method as claimed in claim 21, wherein the norm is the Euclidean norm.
23. A method as claimed in any preceding claim, wherein the weighting coefficients are optimized by second order cone programming.
24. A method as claimed in any preceding claim, wherein one or more weighting coefficients are optimized for each order n of spherical harmonic, but within each order of spherical harmonics, said weighting coefficients are common to all degrees $m=-n$ to $m=n$ of said order n .
25. A method as claimed in any preceding claim, wherein the input signals are transformed into the frequency domain before being decomposed into the spherical harmonics domain.
26. A method as claimed in claim 25, wherein the beamformer is a broadband beamformer in which the frequency domain signals are divided into narrowband frequency bins and wherein each bin is optimized and weighted separately before the frequency bins are recombined into a broadband output.
27. A method as claimed in any of claims 1 to 24, wherein the input signals are processed in the time domain and wherein the weighting coefficients are the tap weights of finite impulse response filters applied to the spherical harmonic signals.

- 68 -

28. A beamformer comprising:
- an array of sensors, each of which is arranged to generate a signal;
 - a spherical harmonic decomposer which is arranged to decompose the input signals into the spherical harmonics domain and to output the decomposed signals;
 - a weighting coefficients calculator which is arranged to calculate weighting coefficients to be applied to the decomposed signals by convex optimization based on a set of input parameters; and
 - an output generator which combines the decomposed signals with the calculated weighting coefficients into an output signal;
29. A beamformer as claimed in claim 28, further comprising a signal tracker which is arranged to evaluate the signals from the sensors to determine the directions of desired signal sources and the directions of unwanted interference sources.
30. A method of forming a beampattern in a beamformer of the type in which the beamformer receives input signals from a sensor array, applies weighting coefficients to the signals and combines them to form an output signal, wherein the weighting coefficients are optimized for a given set of input parameters by convex optimization, subject to constraints that the array gain in a plurality of specified directions be maintained at a given level, so as to form multiple main lobes in the beampattern, and wherein each requirement is formulated as a requirement that the array output for a unit magnitude plane wave incident on the array from the specified direction is equal to a predetermined constant.
31. A software product which when executed on a computer cause the computer to carry out the steps as claimed in any of claims 1 to 27 or 30.
32. A software product as claimed in claim 31, wherein the software product is a data carrier.
33. A software product as claimed in claim 31, wherein the software product comprises signals transmitted from a remote location.
34. A method of manufacturing a software product which is in the form of a physical carrier, comprising storing on the data carrier instructions which when executed by a computer cause the computer to carry out the method as claimed in any of claims 1 to 27 or 30.

- 69 -

35. A method of providing a software product to a remote location by means of transmitting data to a computer at that remote location, the data comprising instructions which when executed by the computer cause the computer to carry out the method as claimed in any of claims 1 to 27 or 30.

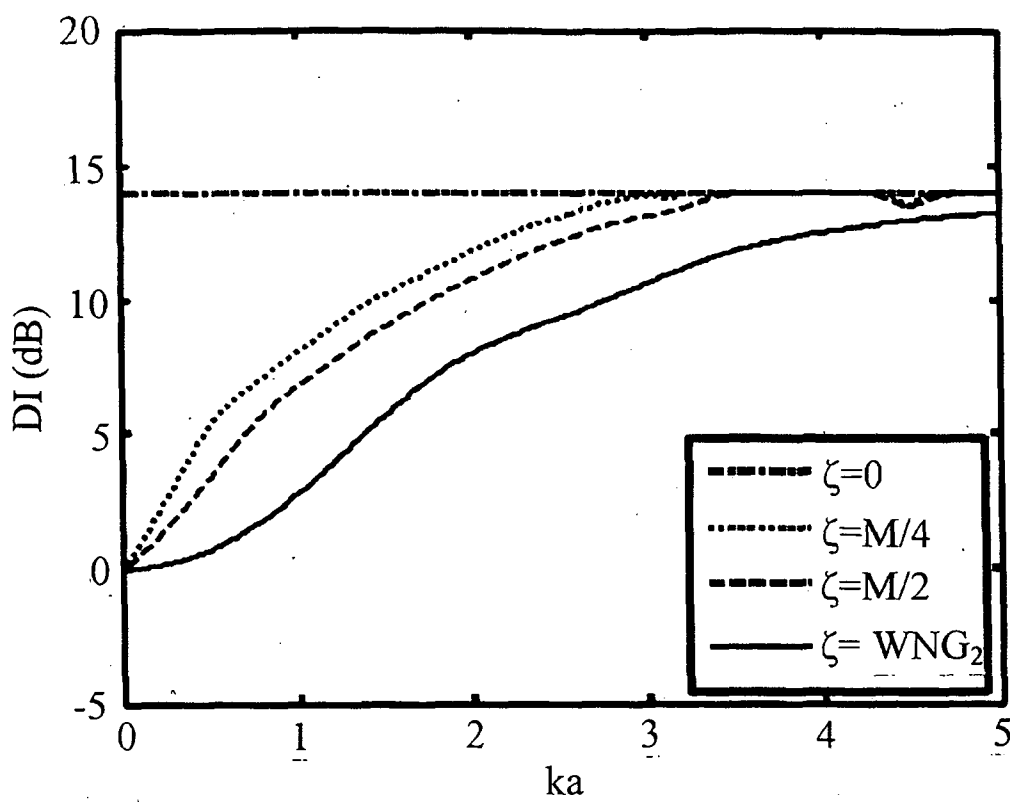


FIG. 1

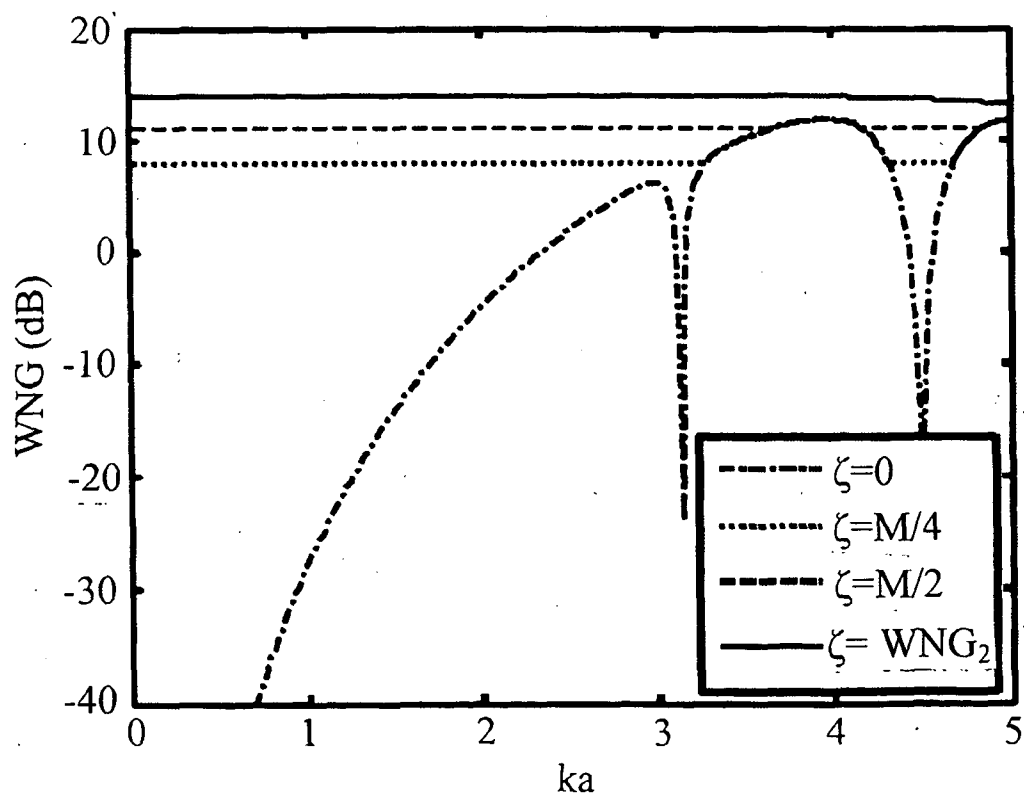


FIG. 2

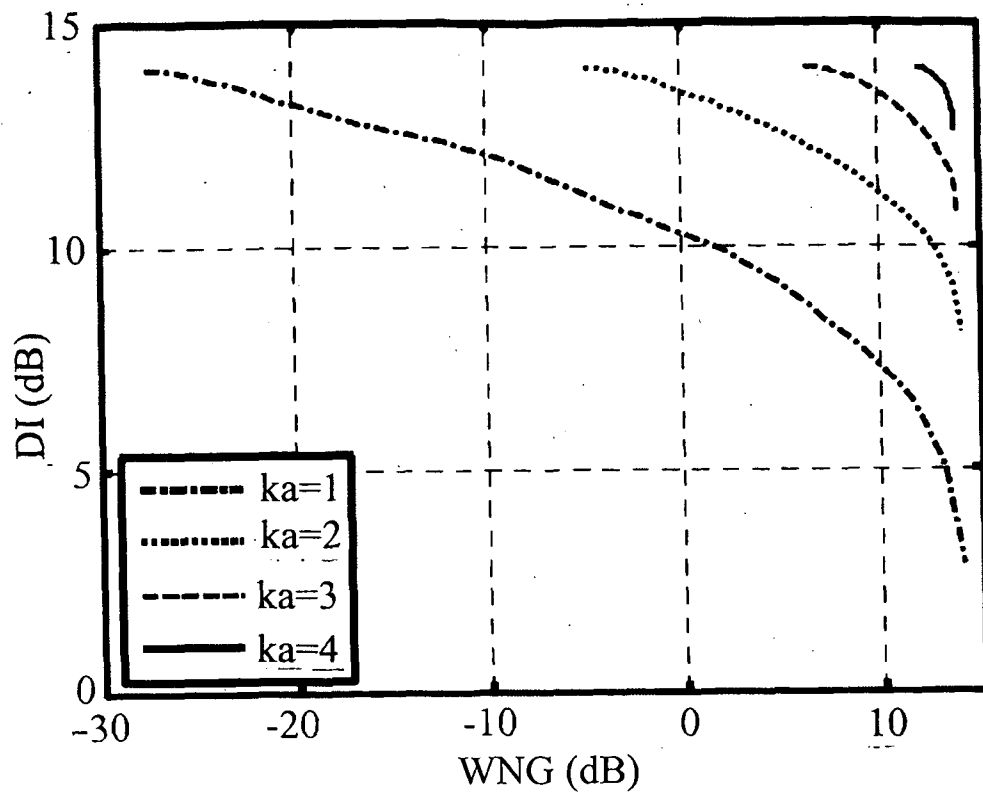


FIG. 3

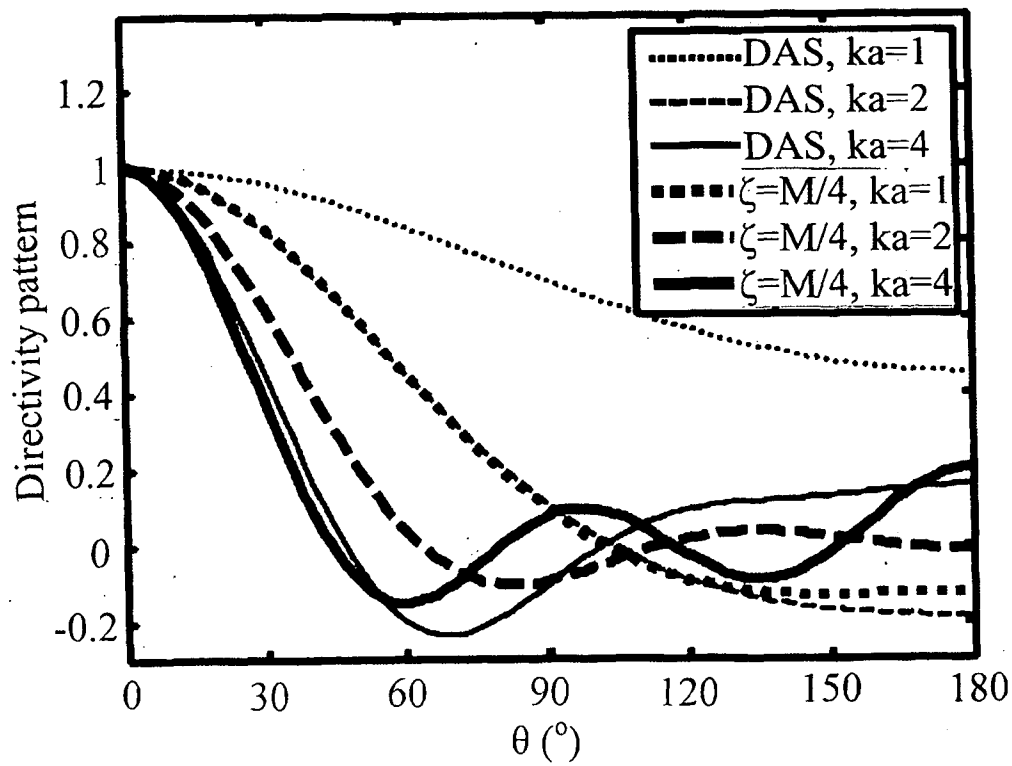


FIG. 5

DI=10.61 dB, WNG=13.97 dB DI=13.98 dB, WNG=6.27 dB DI=13.84 dB, WNG=7.96 dB

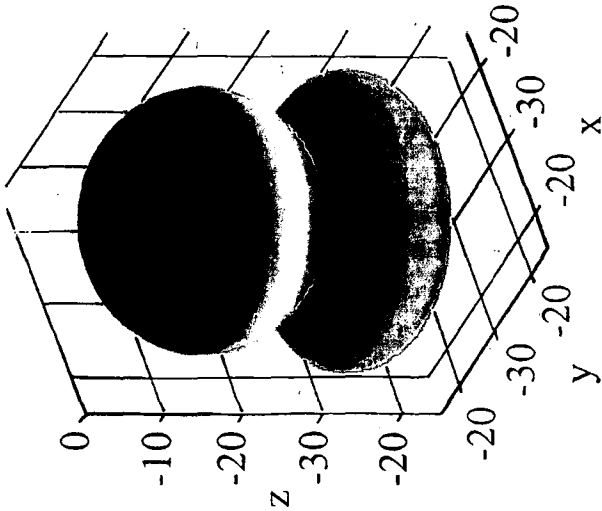


FIG. 4A

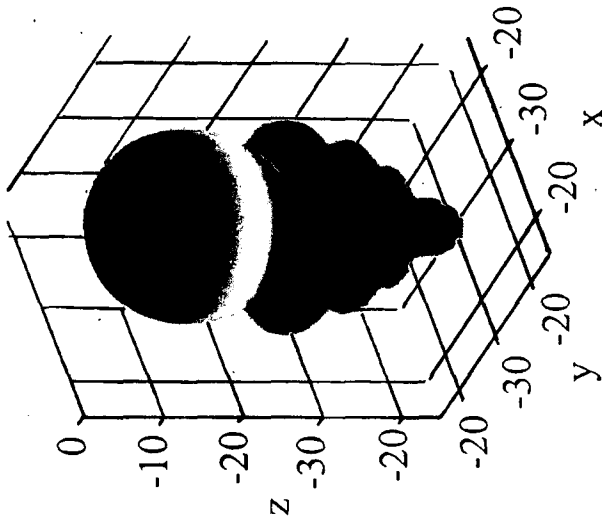


FIG. 4B

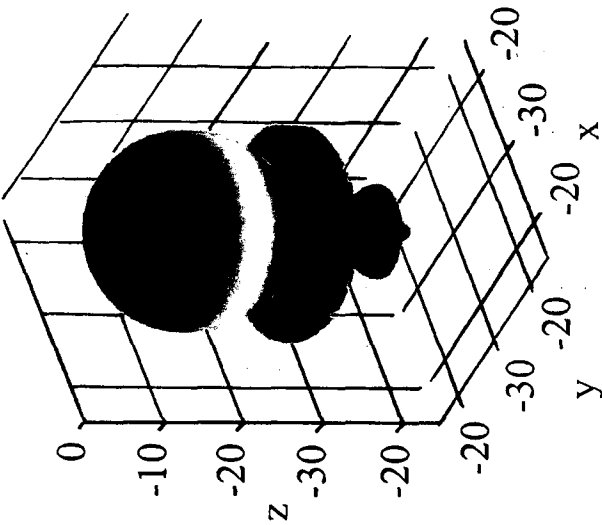
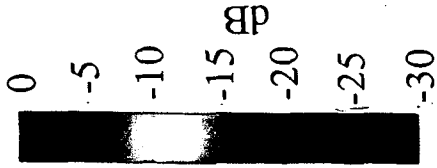


FIG. 4C



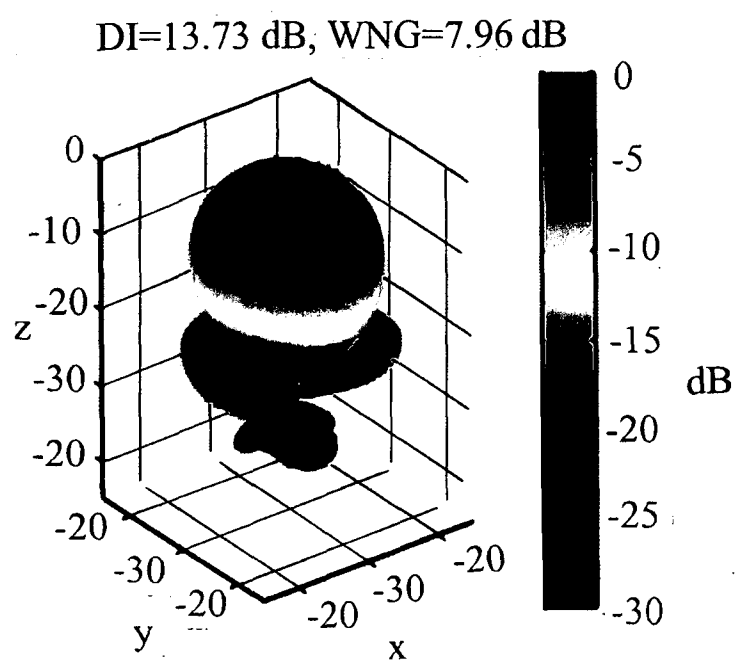
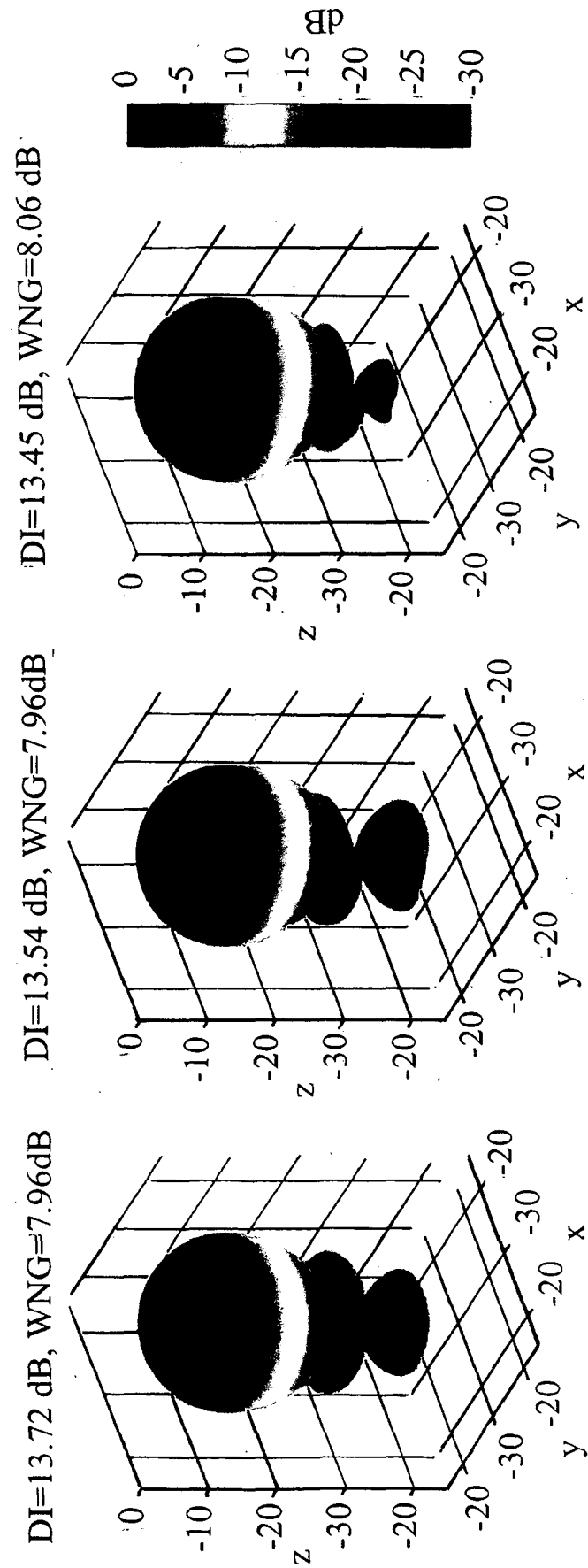


FIG. 6



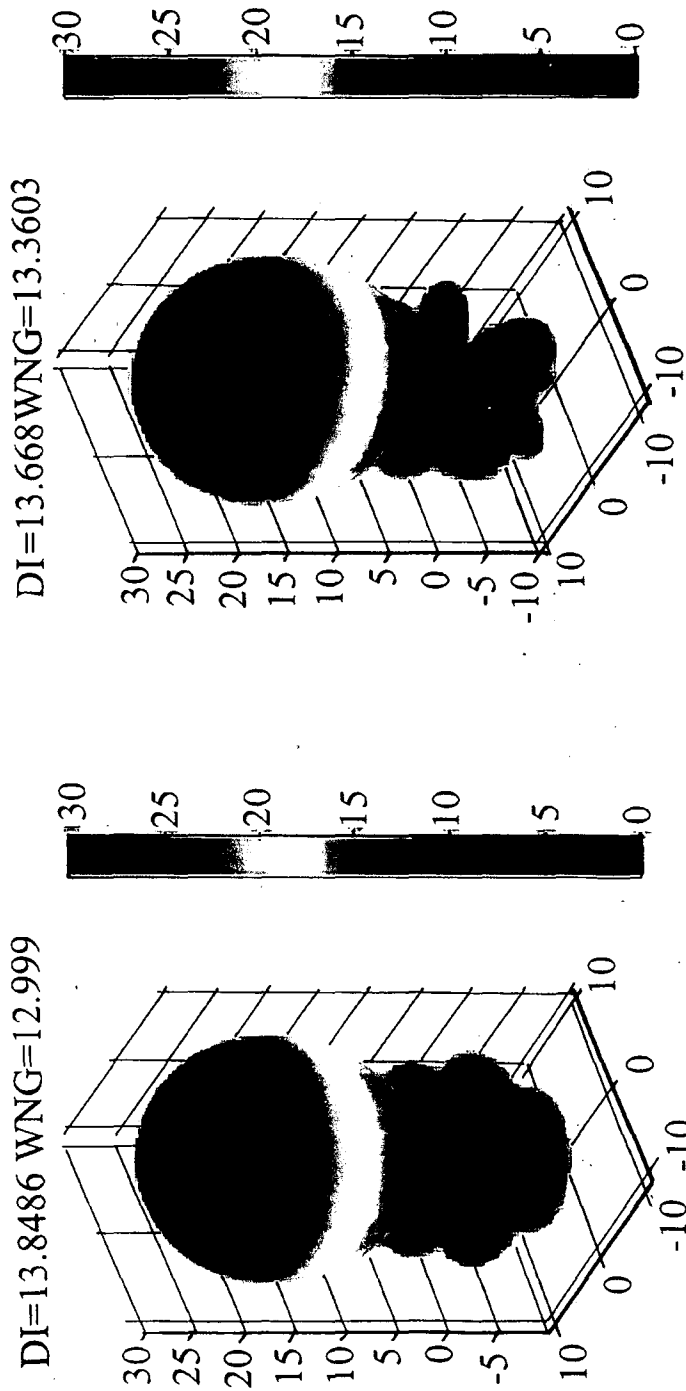


FIG. 8A

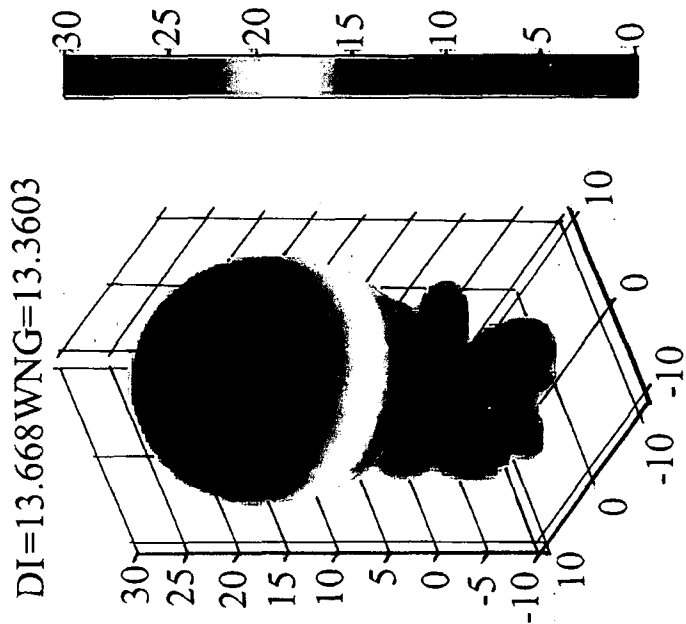


FIG. 8B

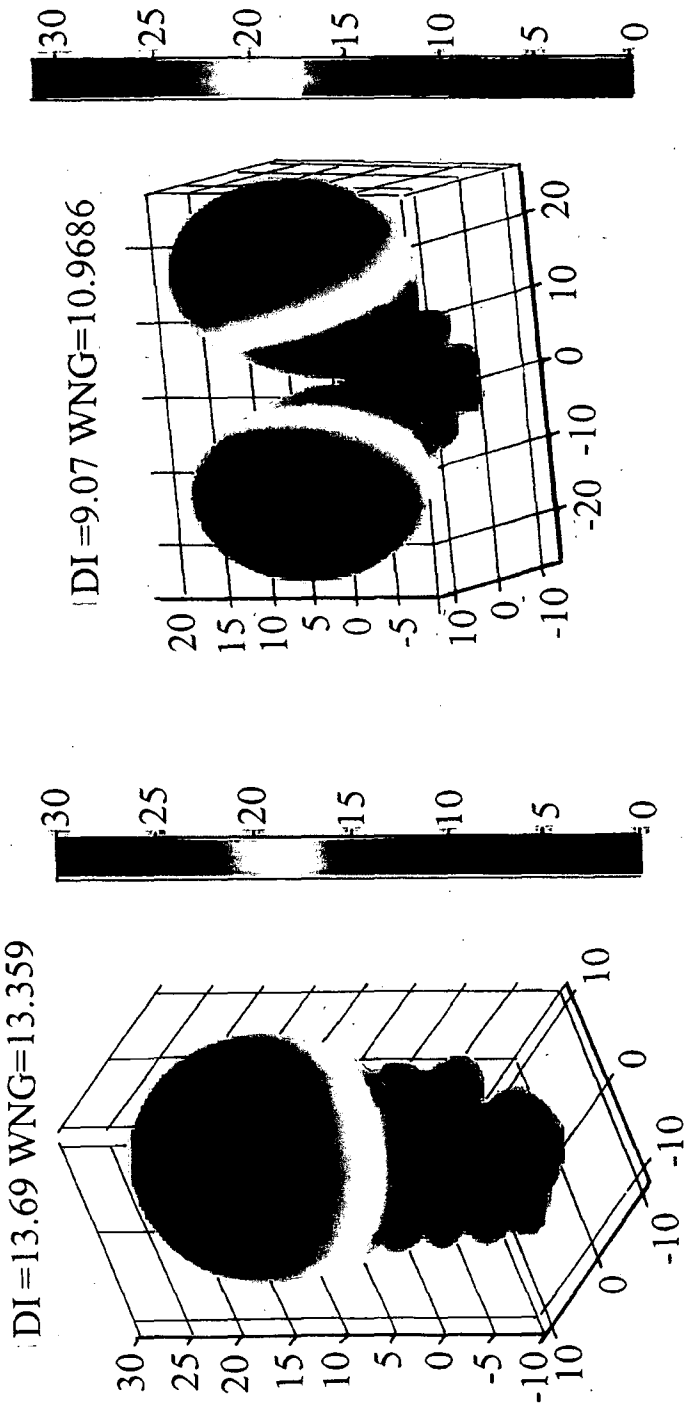


FIG. 9B

FIG. 9A

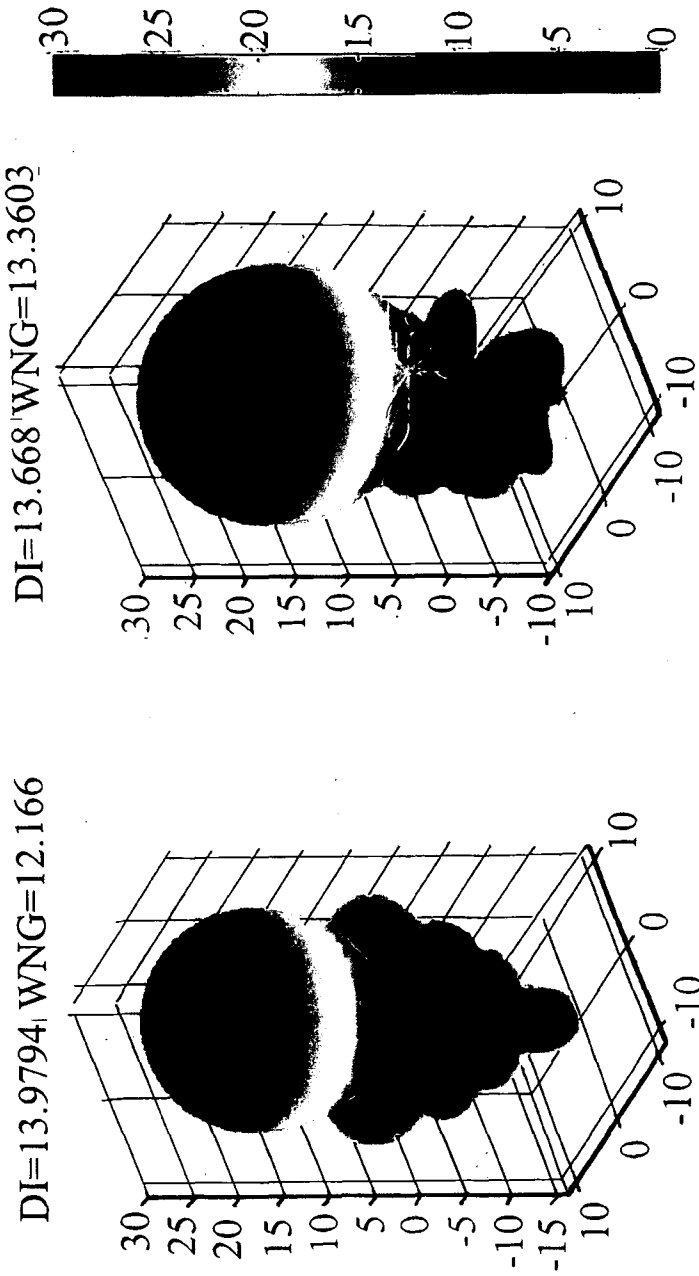


FIG. 10B

FIG. 10A

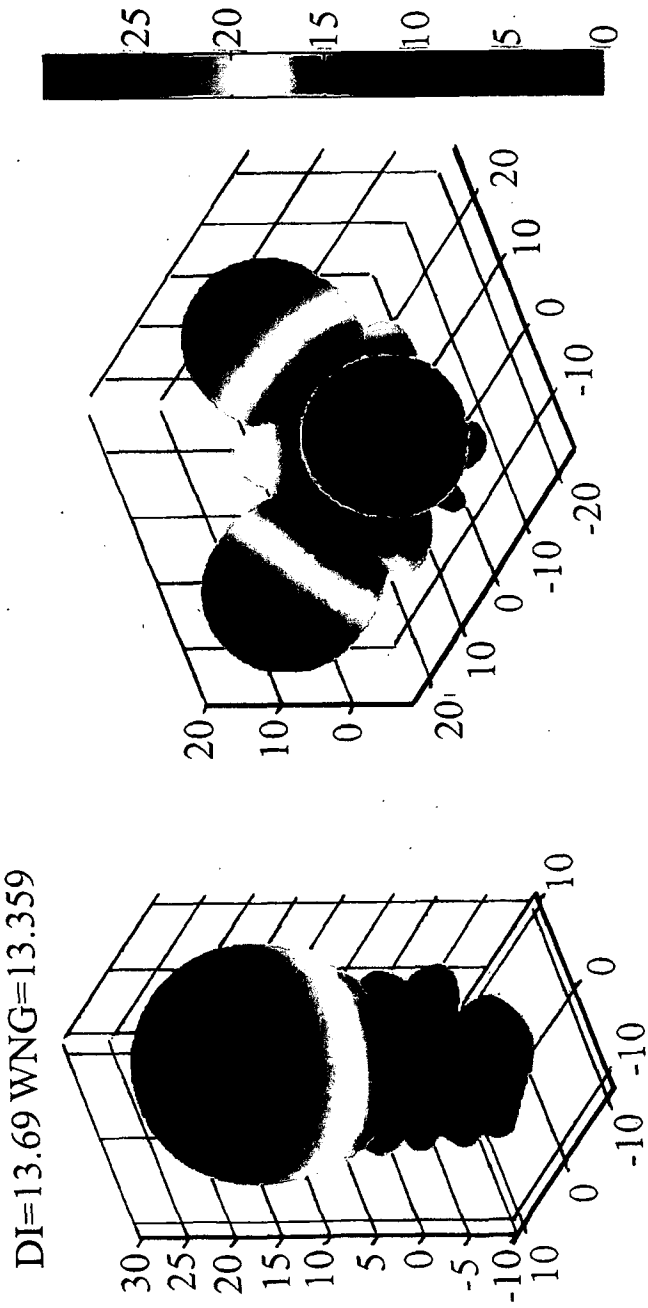


FIG. 11B

FIG. 11A

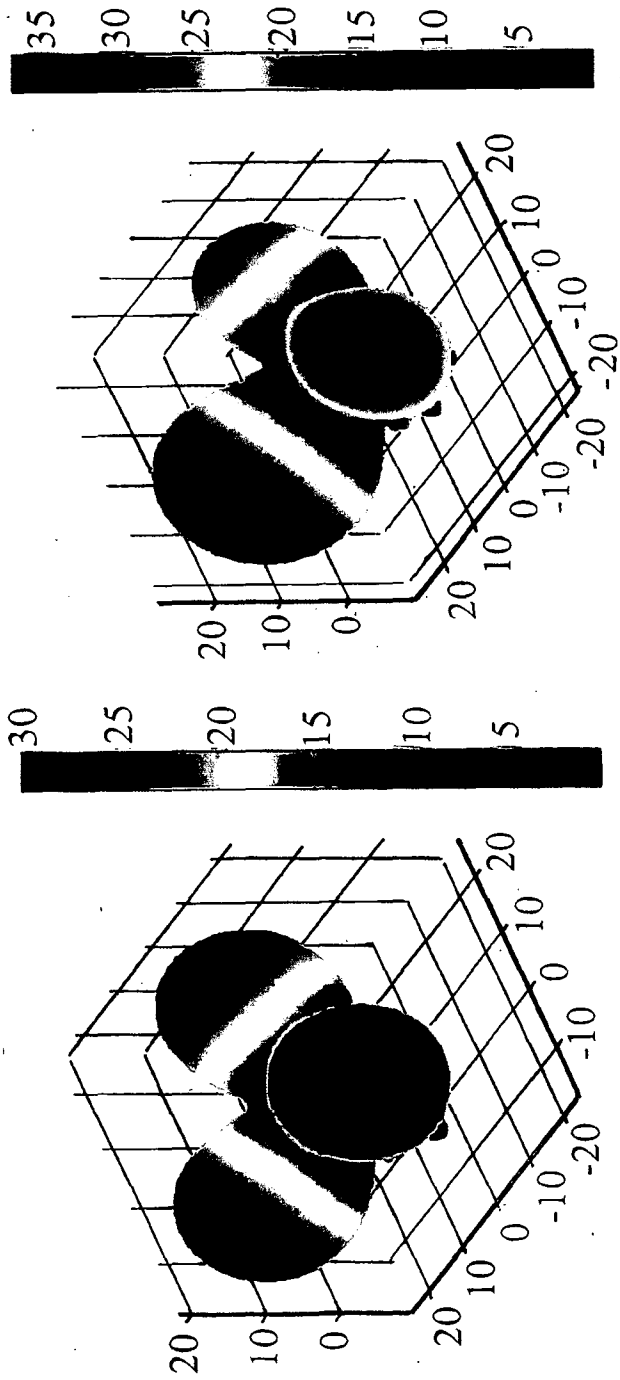


FIG. 12B

FIG. 12A

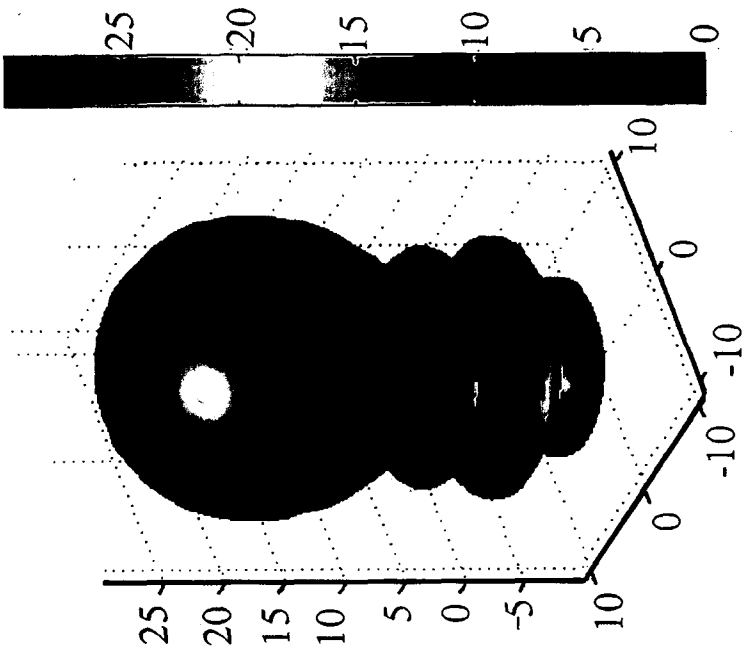


FIG. 14

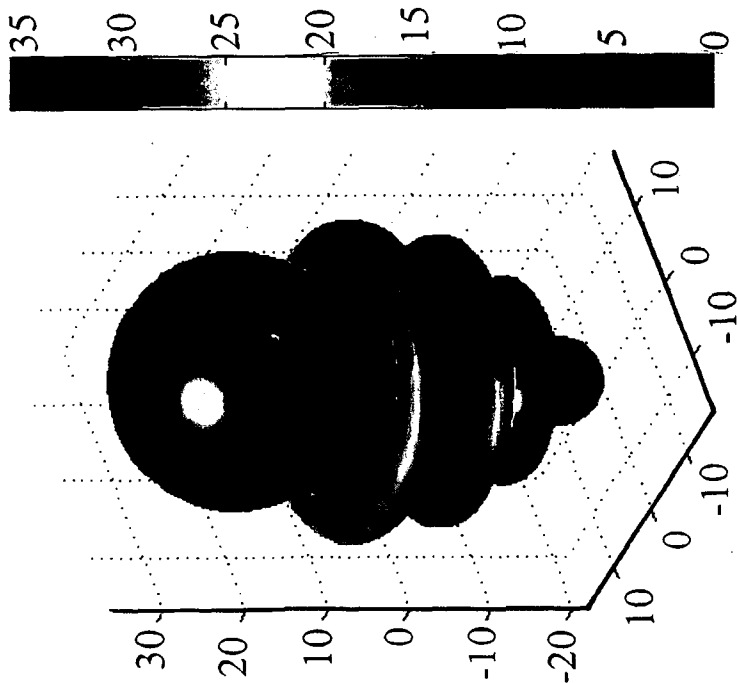


FIG. 13

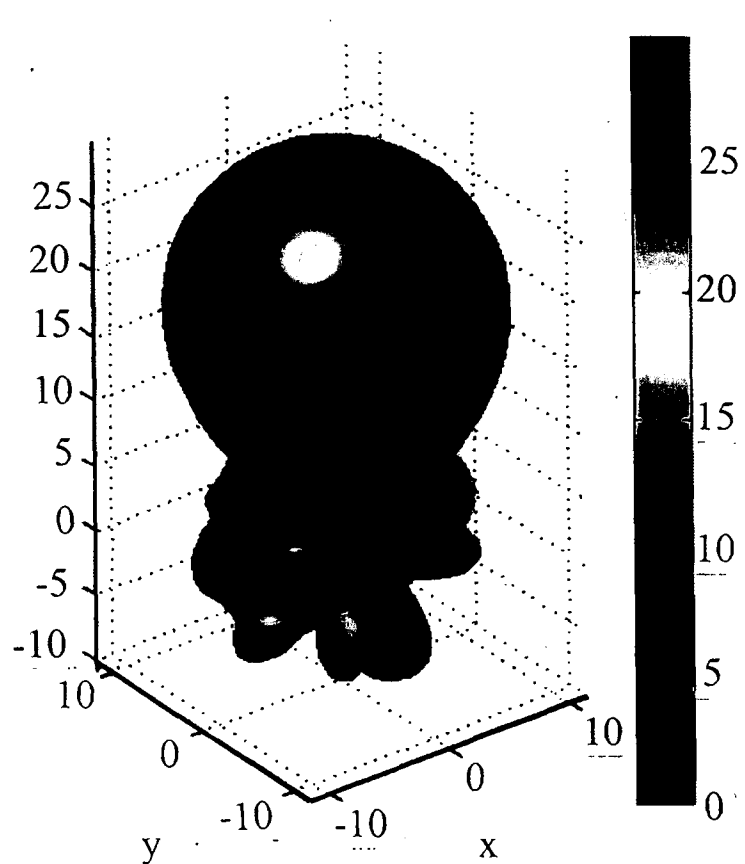


FIG. 15

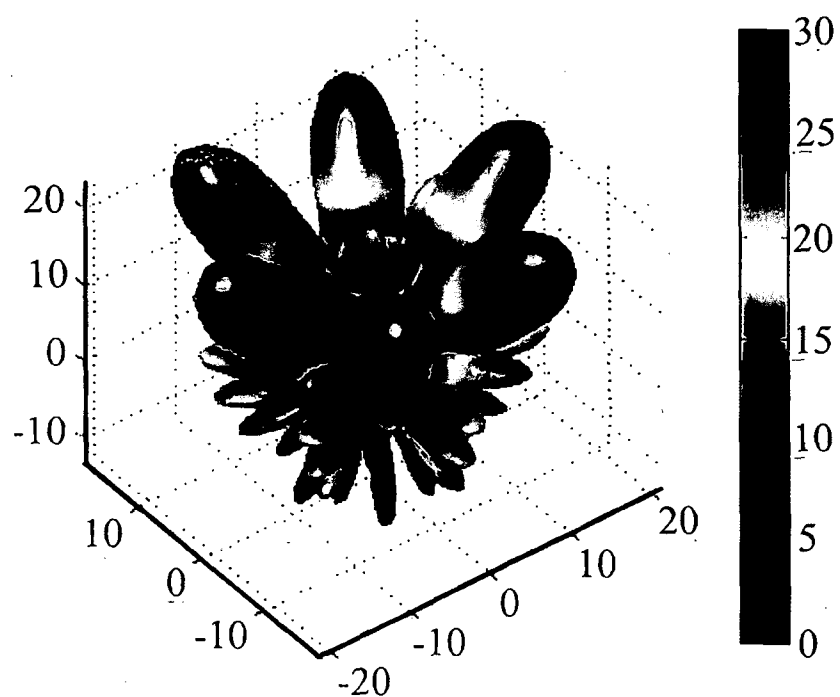


FIG. 16

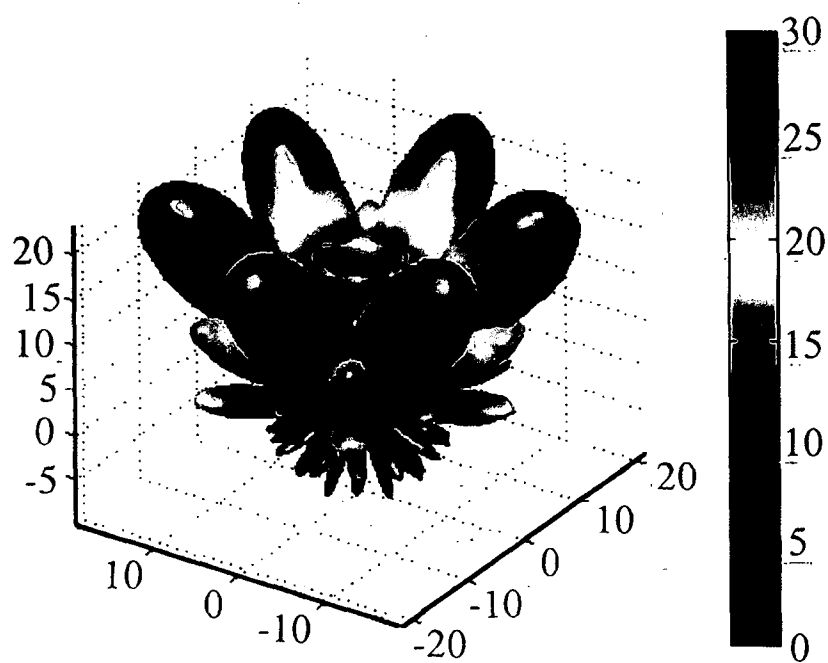


FIG. 17

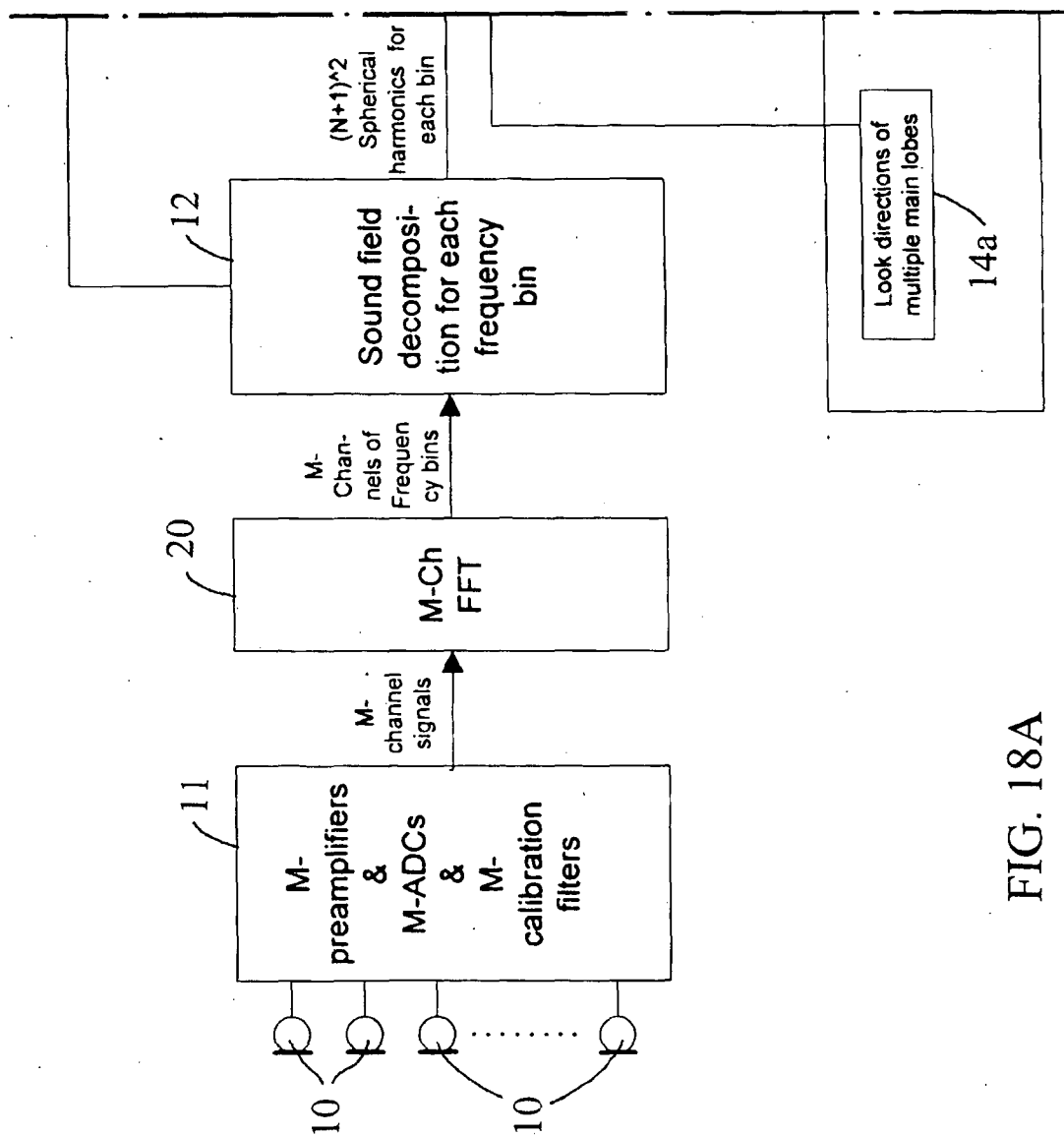


FIG. 18A

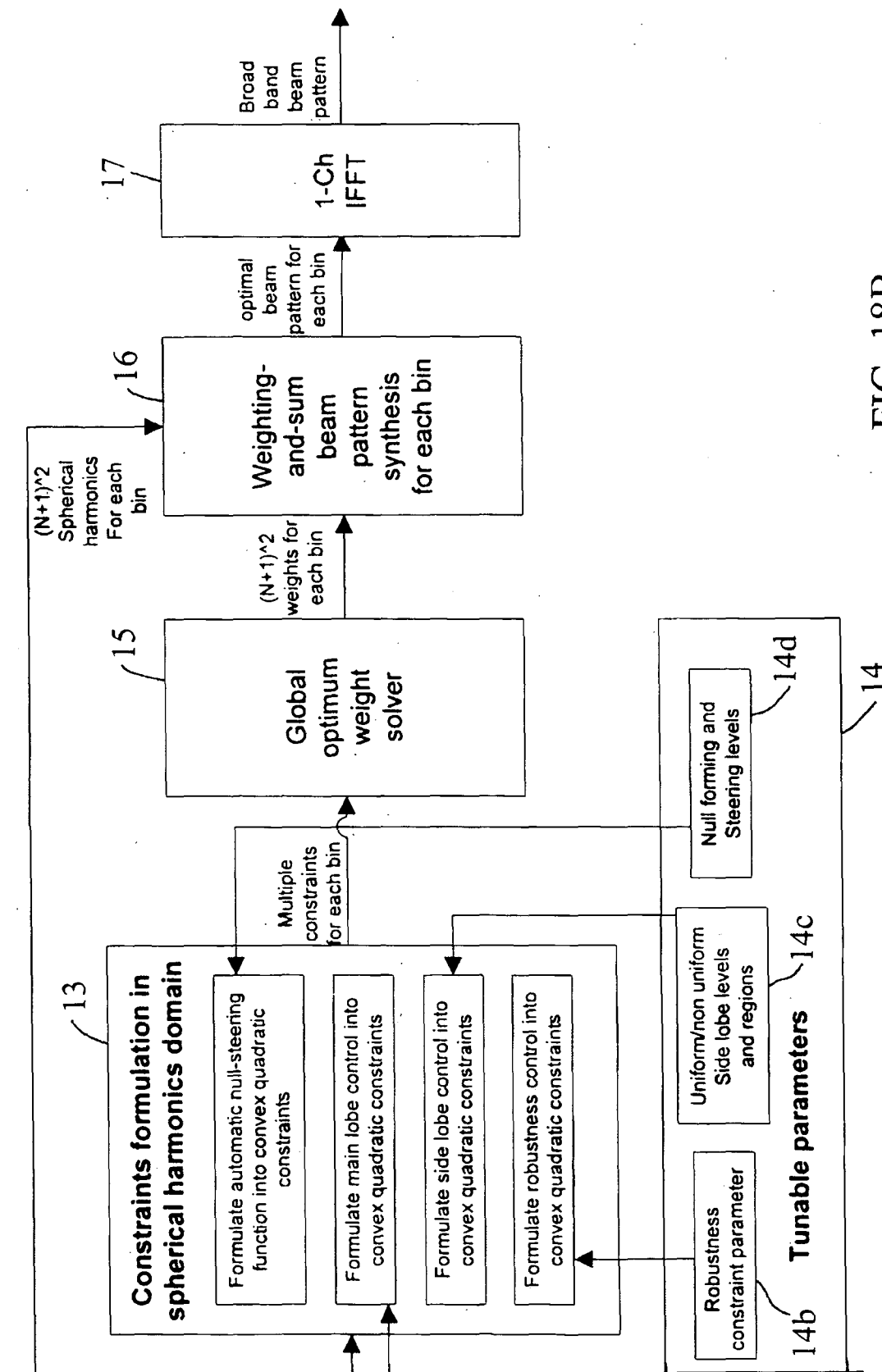


FIG. 18B

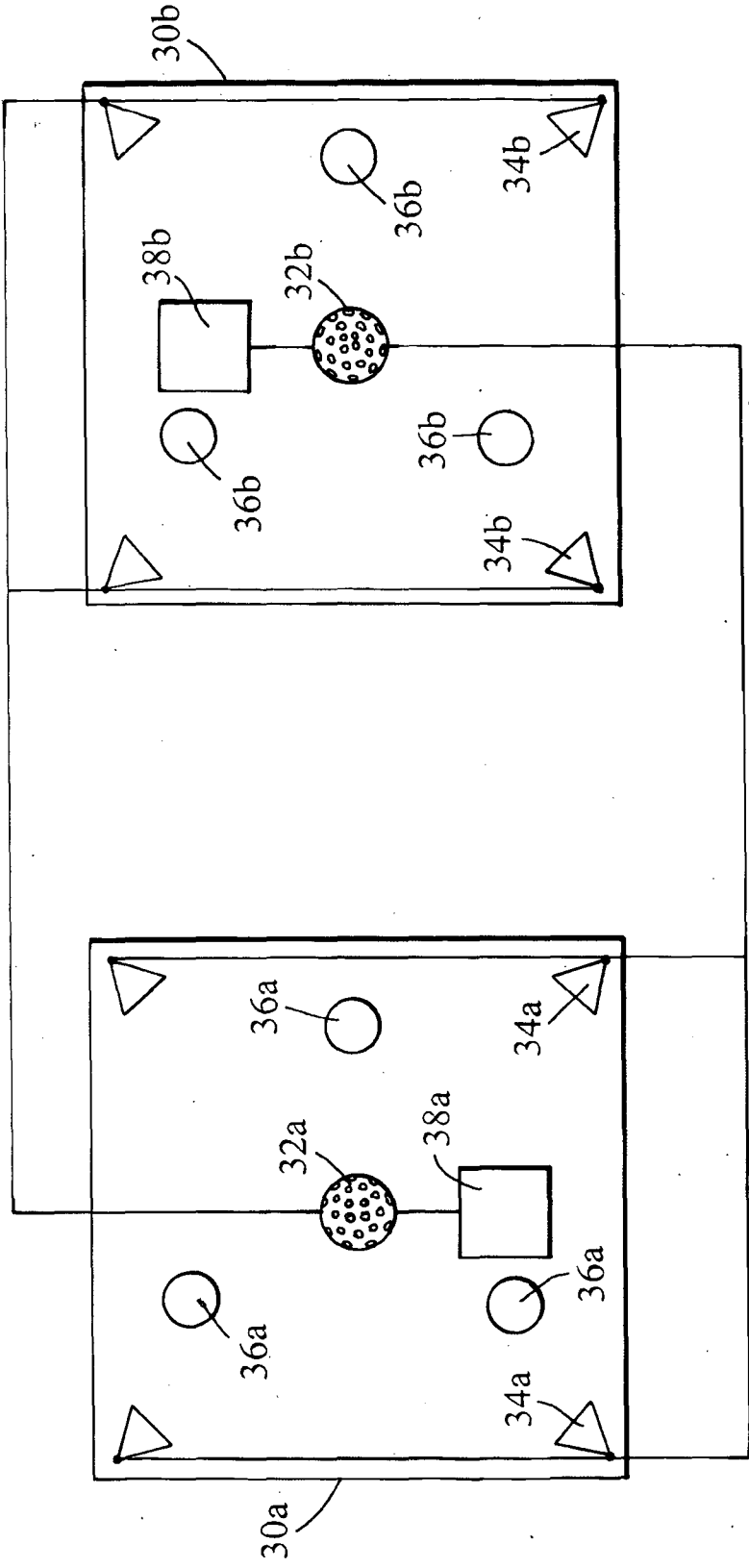


FIG. 19

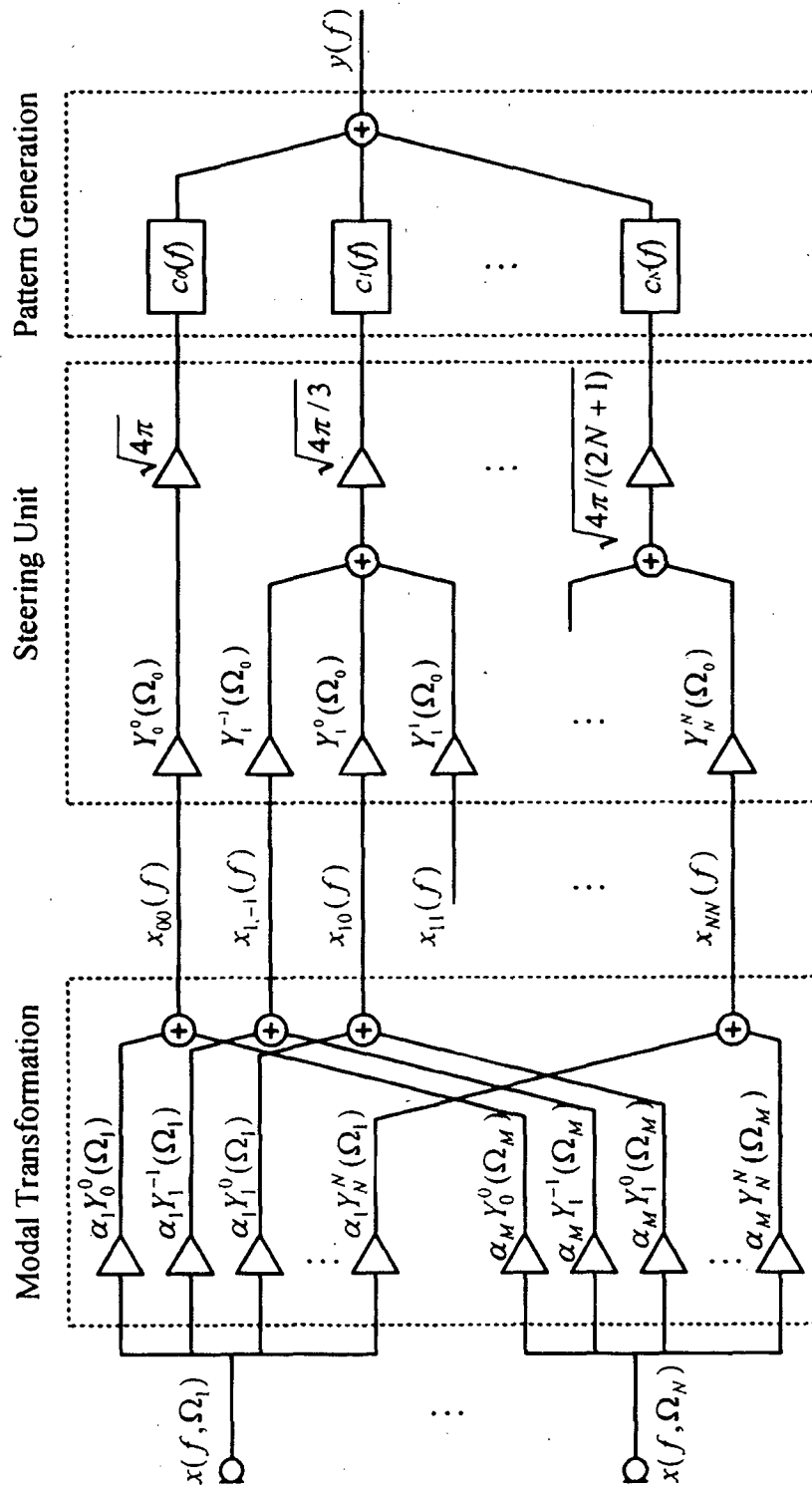


FIG. 20

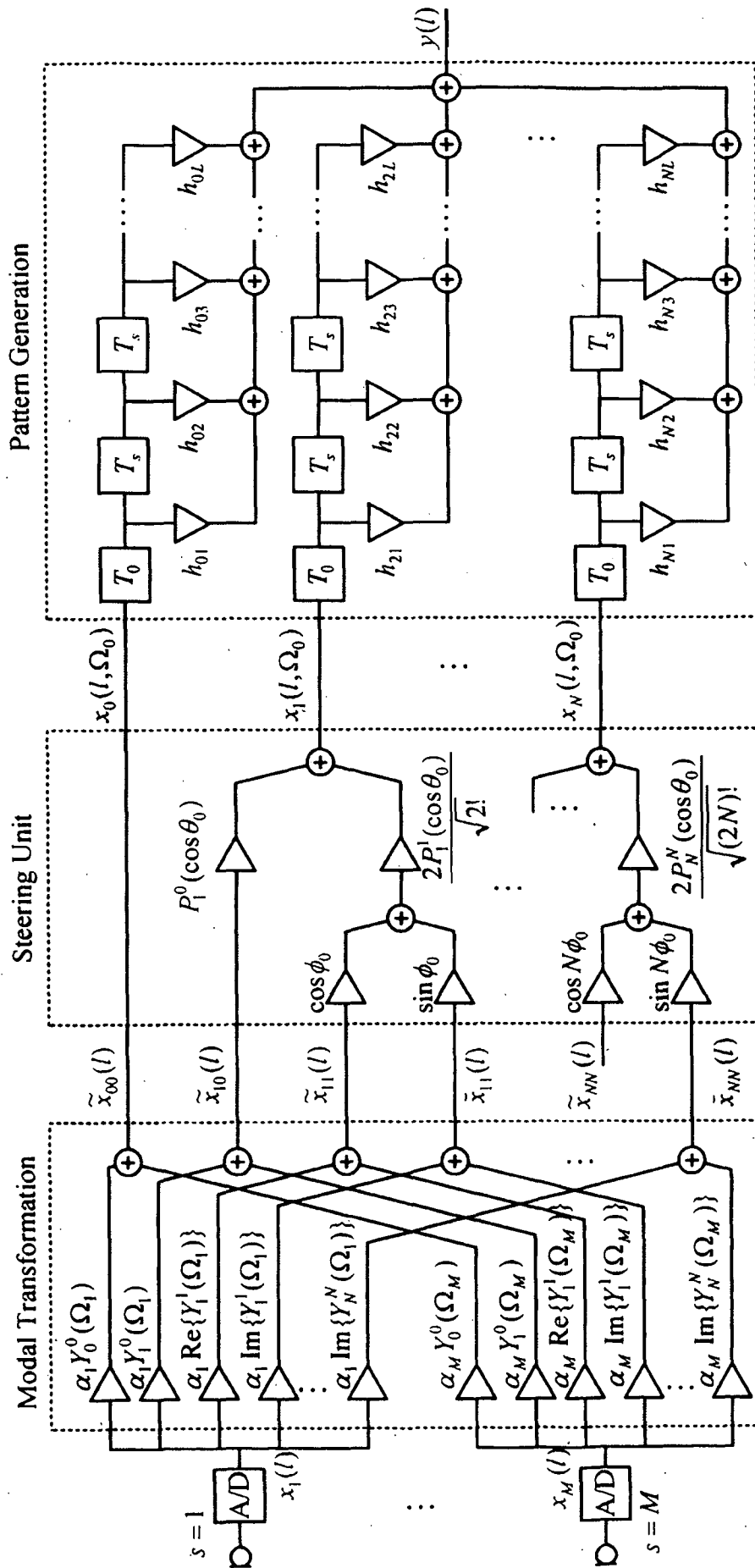


FIG. 21

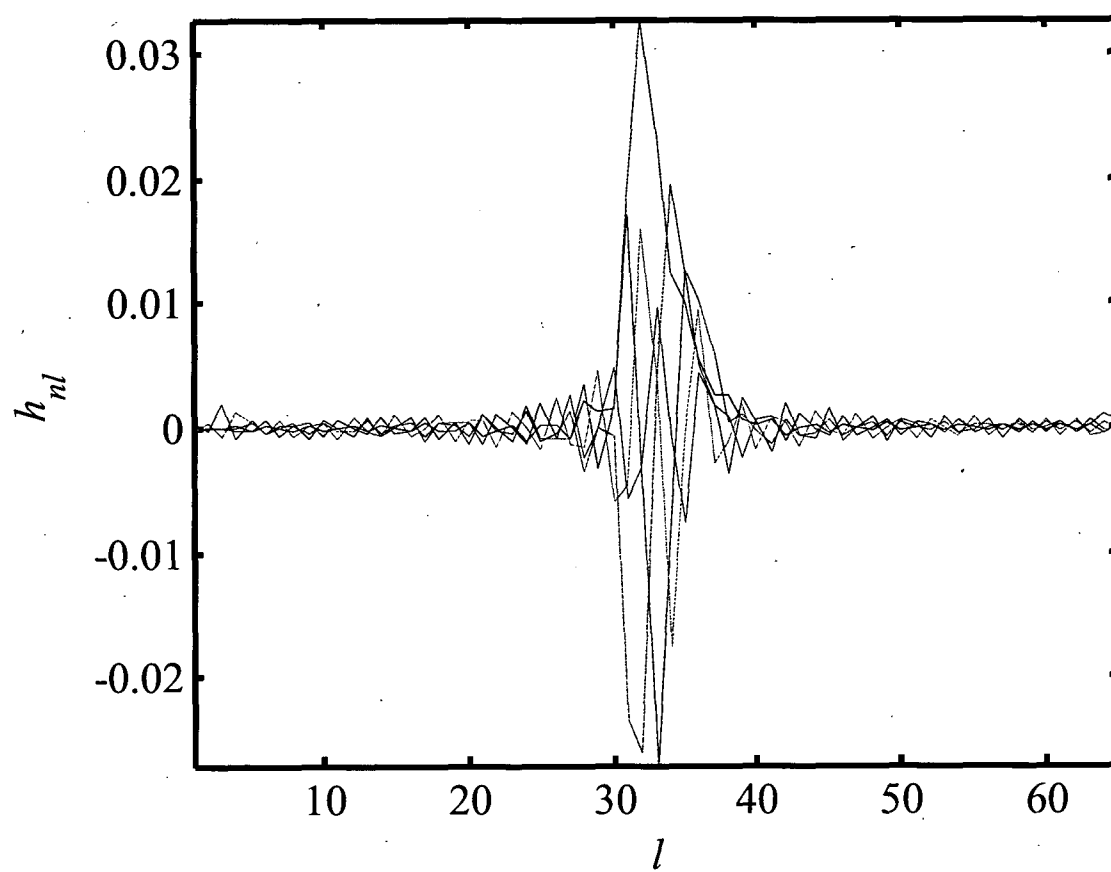


FIG. 22A

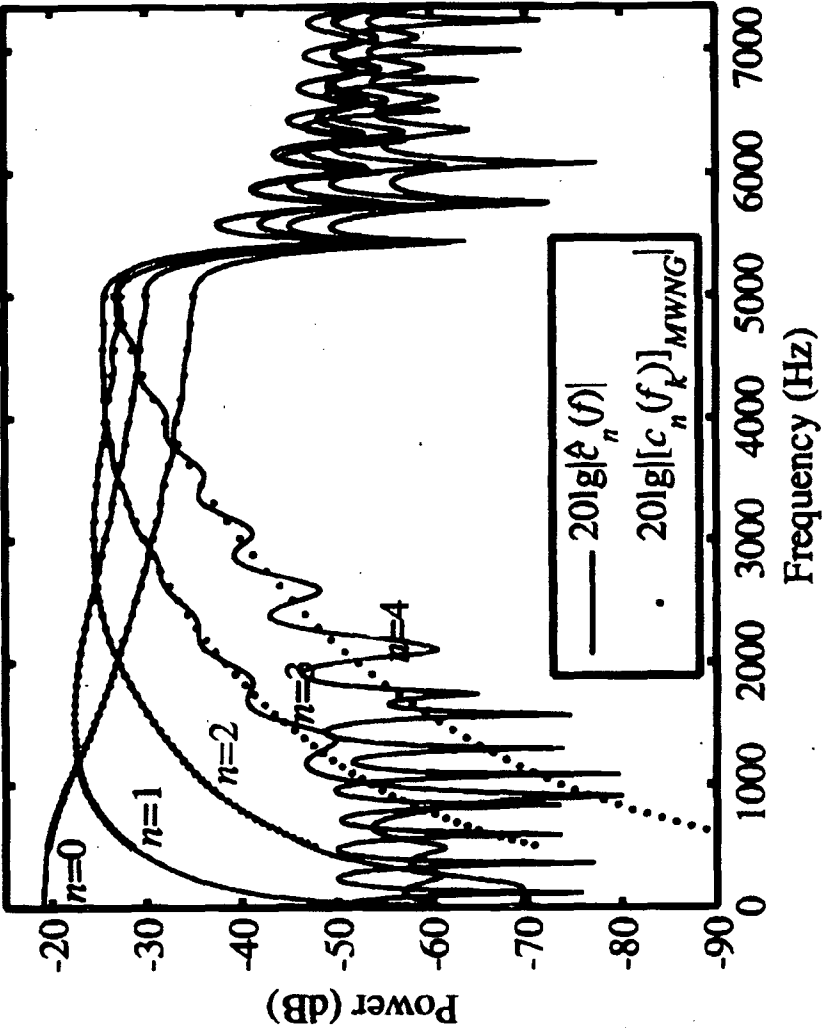


FIG. 22B

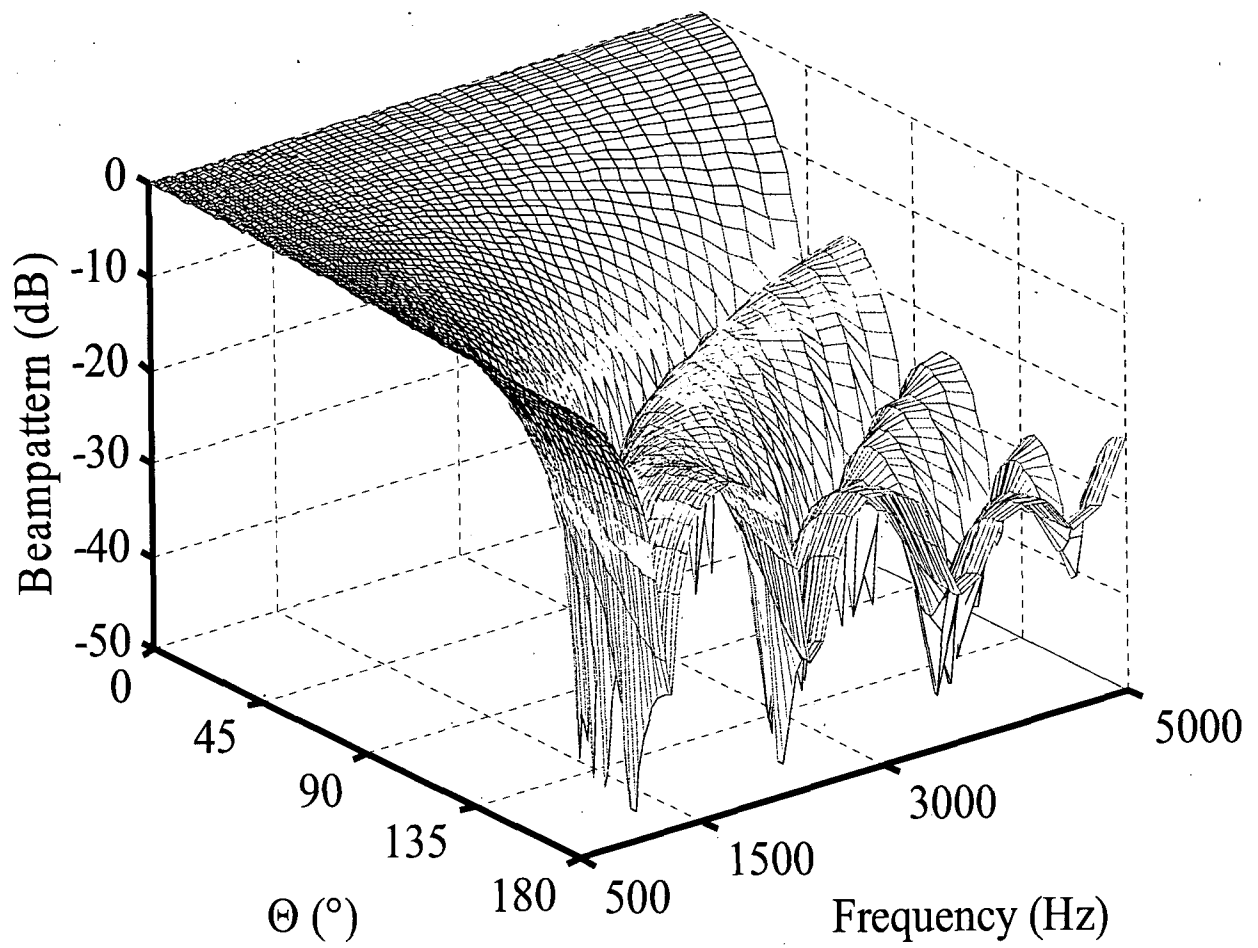


FIG. 22C

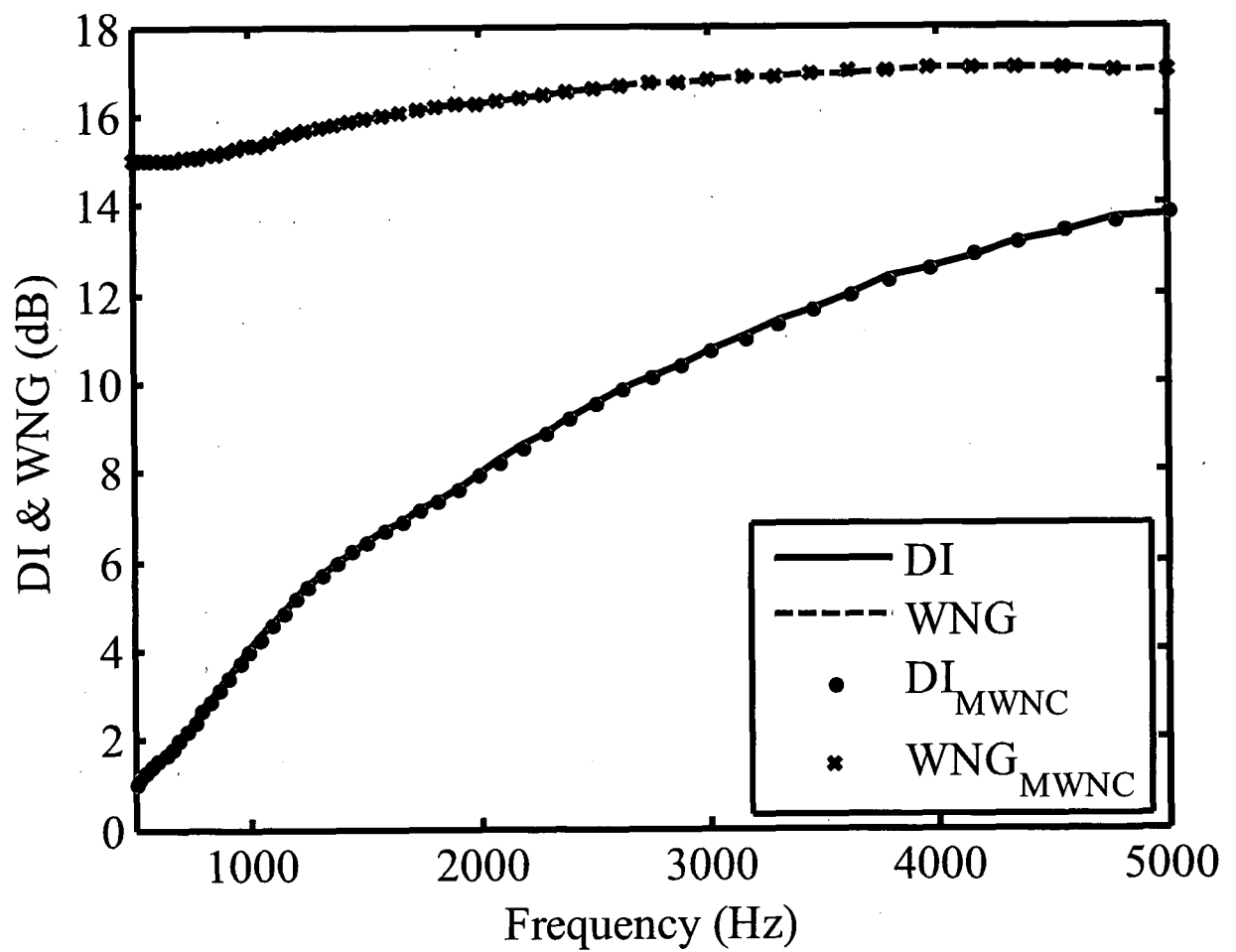


FIG. 22D

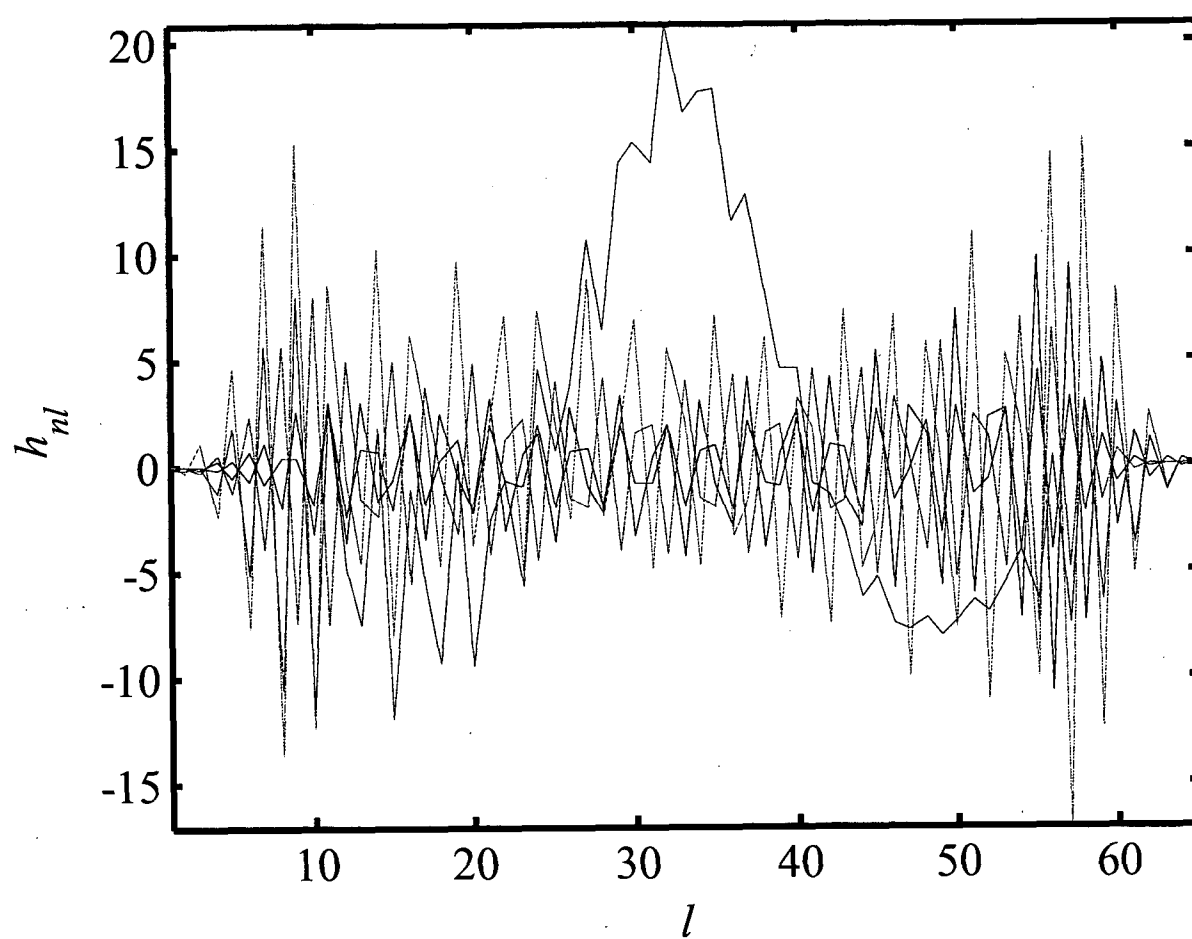


FIG. 23A

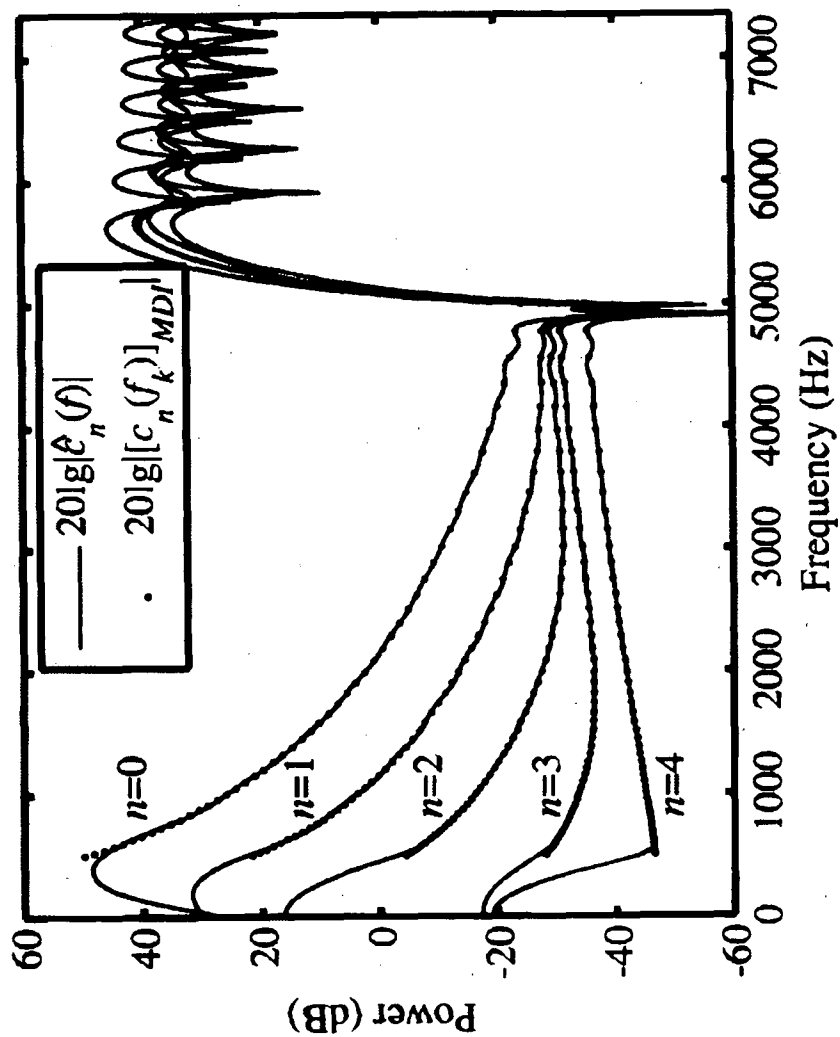


FIG. 23B

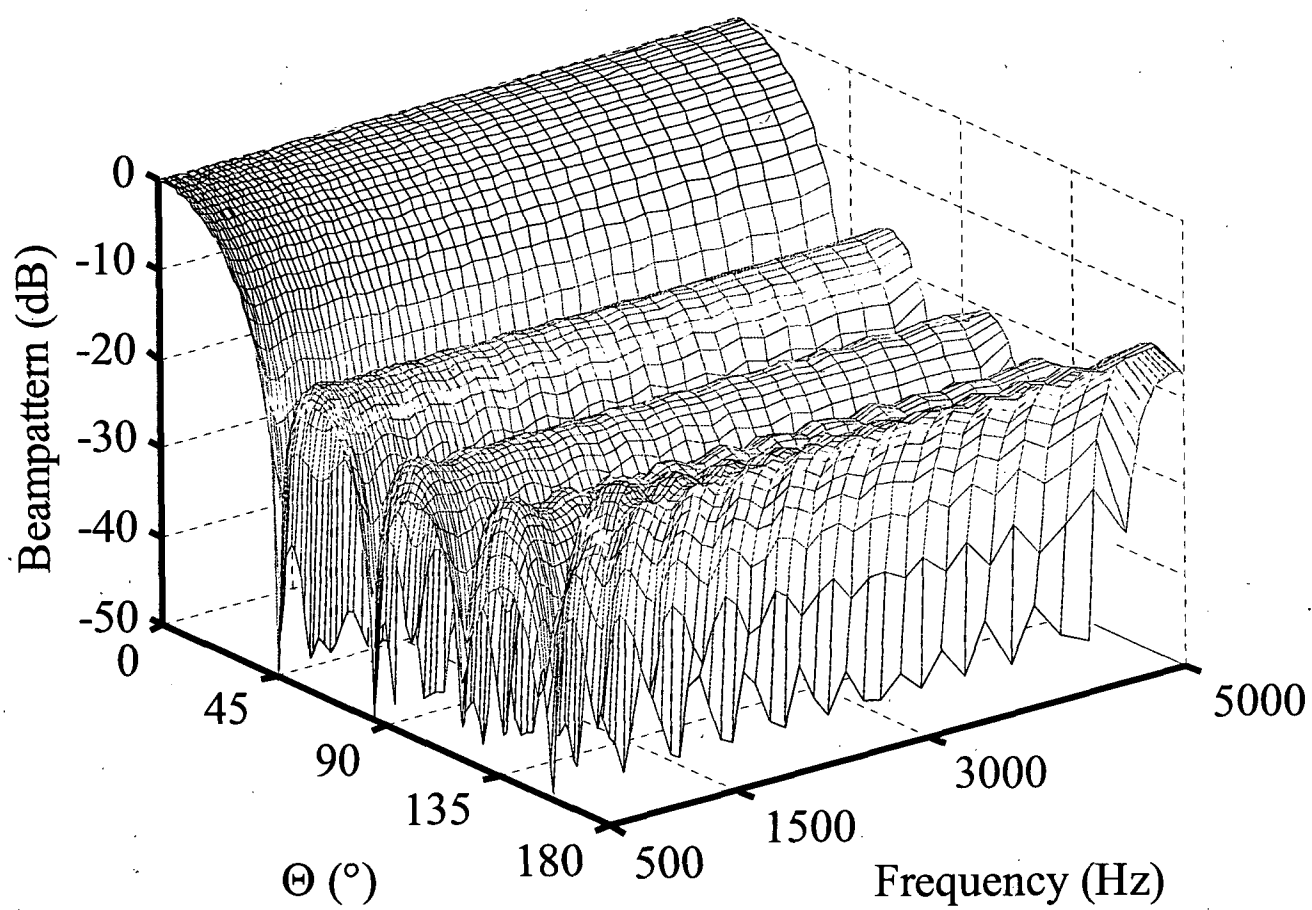


FIG. 23C

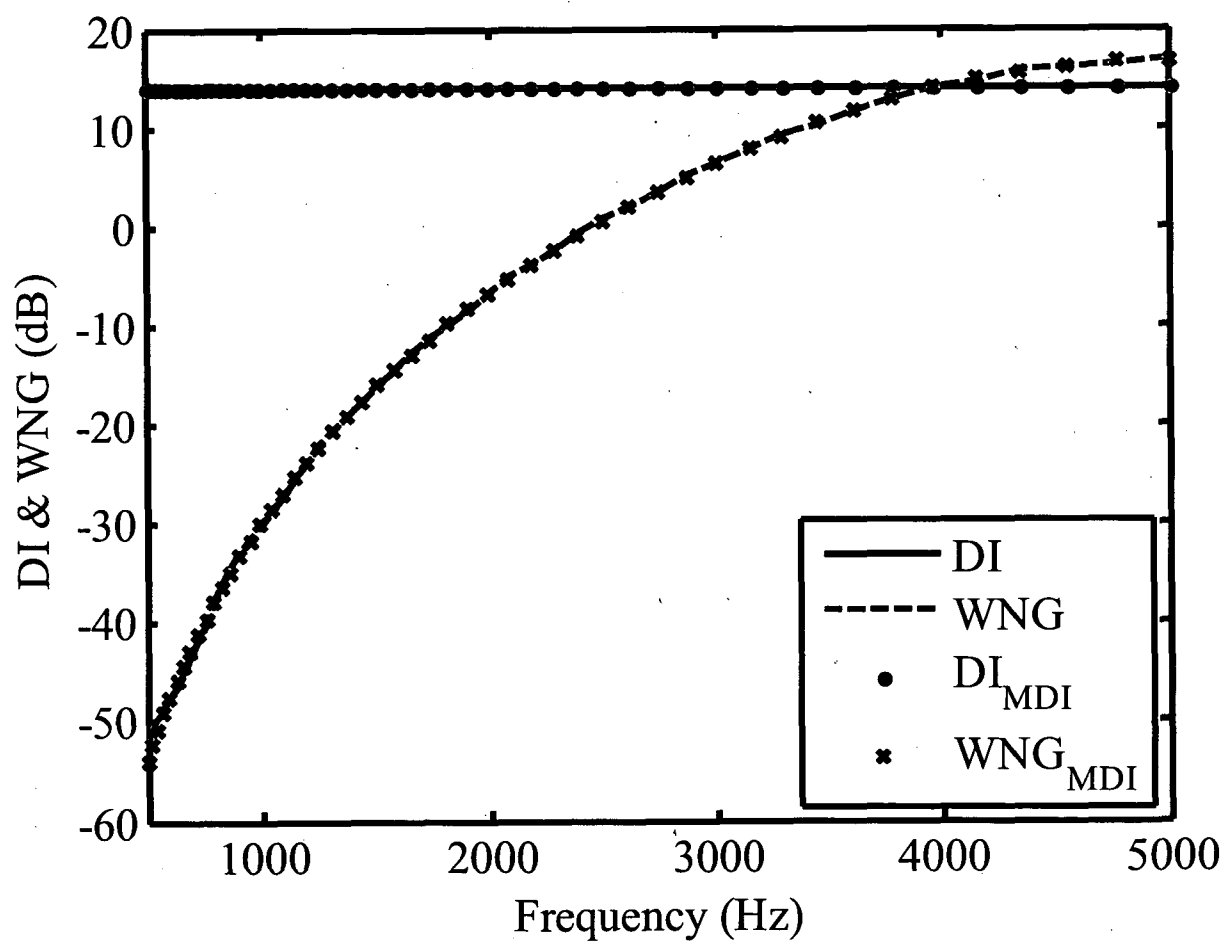


FIG. 23D

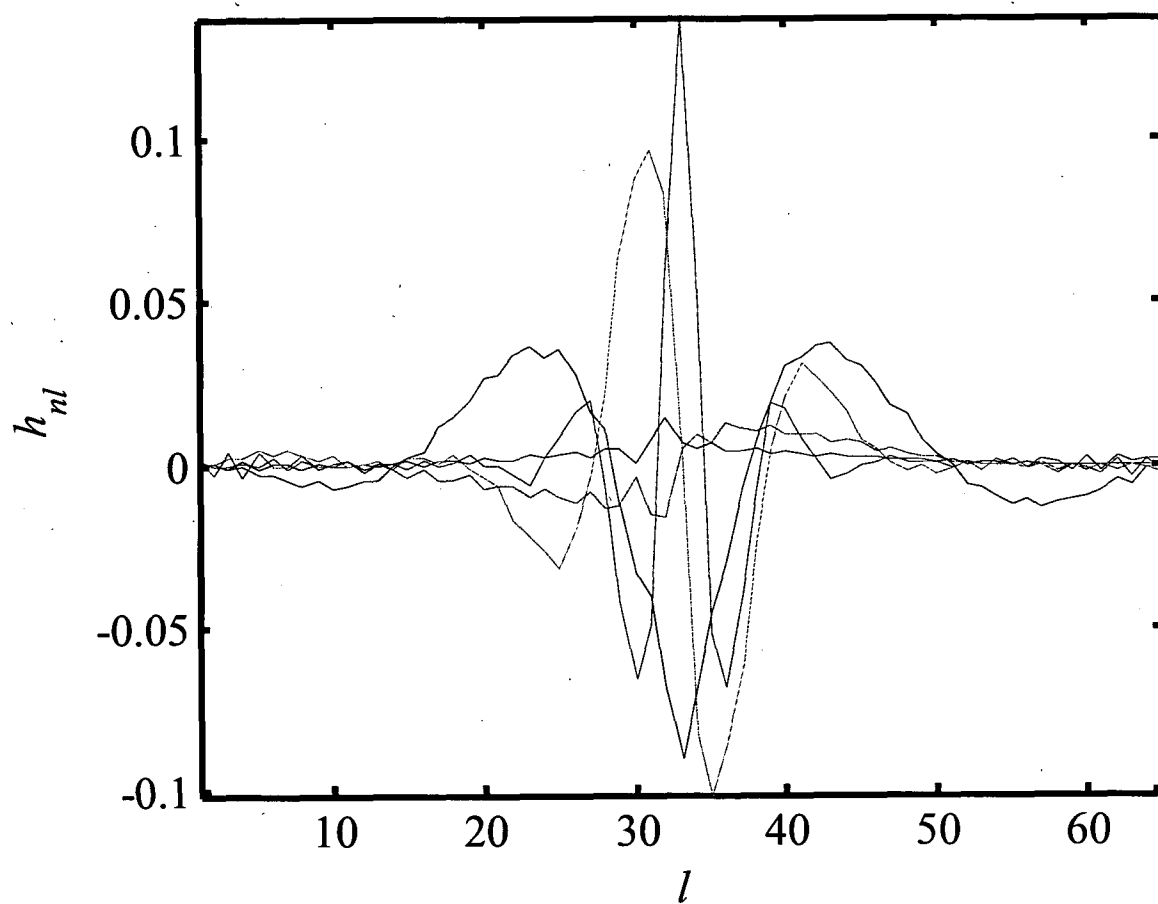


FIG. 24A

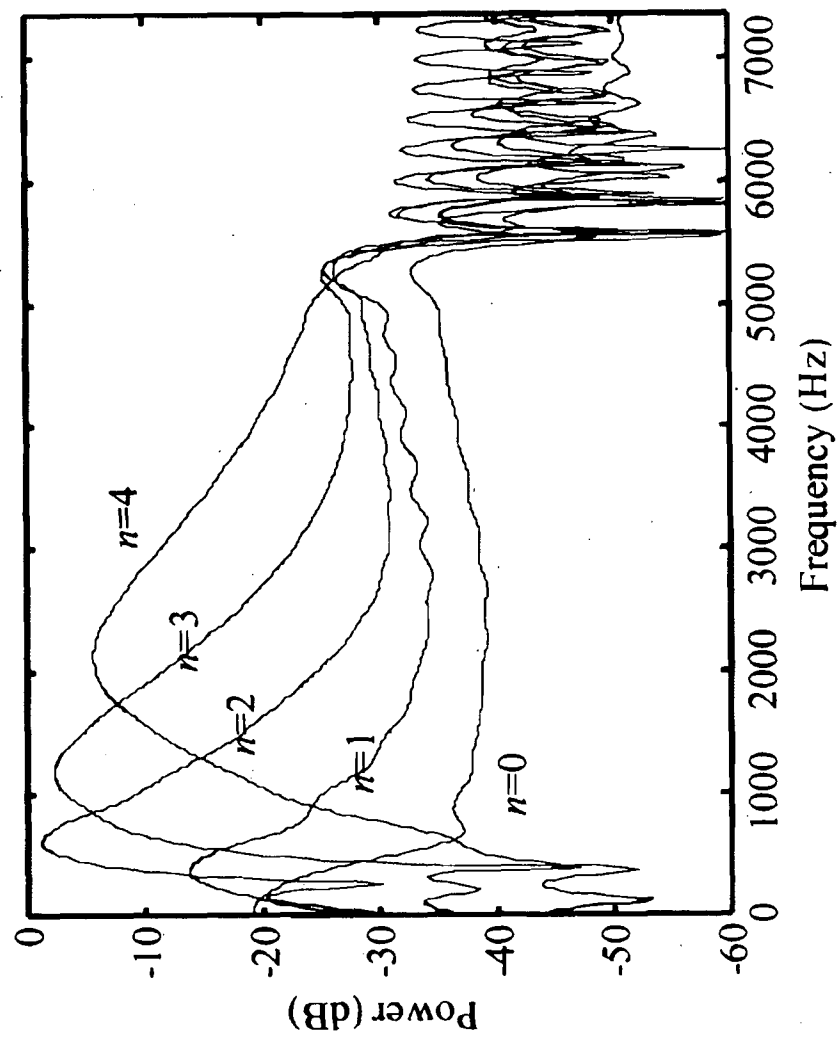


FIG. 24B

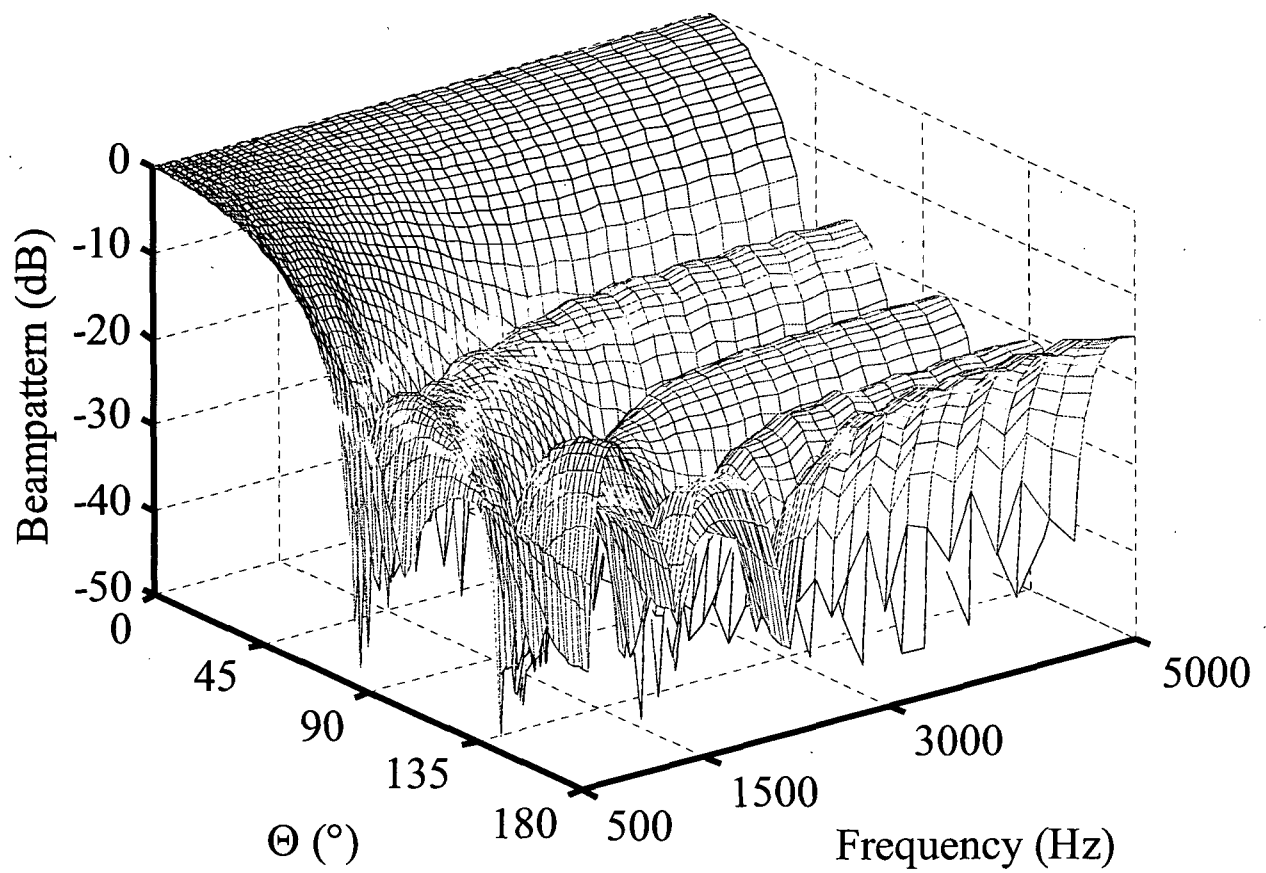


FIG. 24C

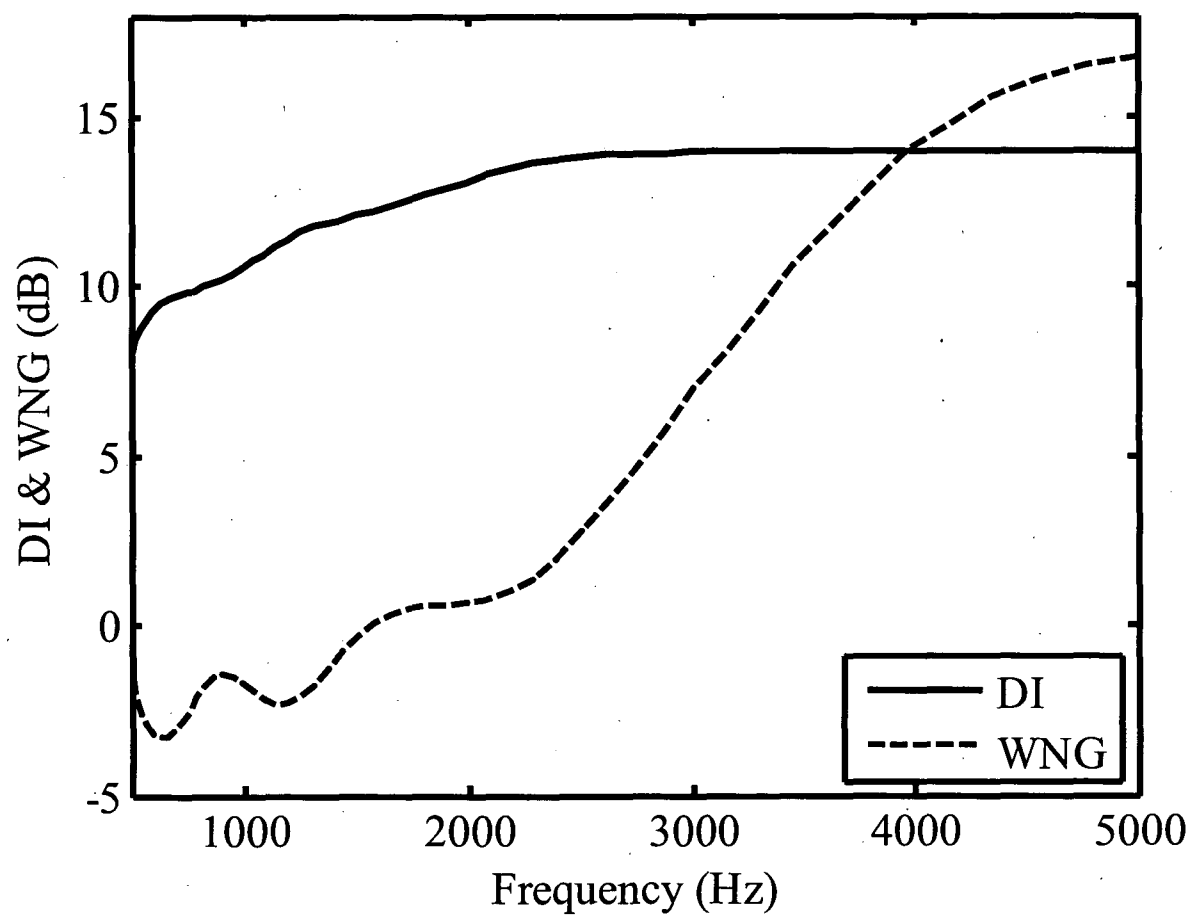


FIG. 24D

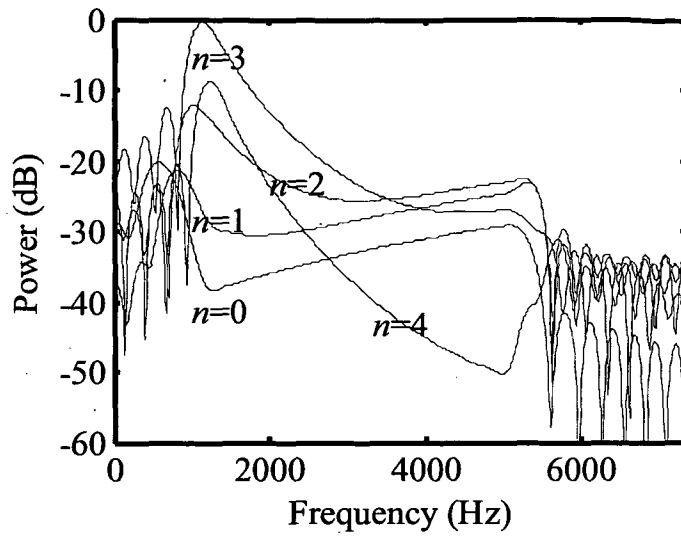


FIG. 25A

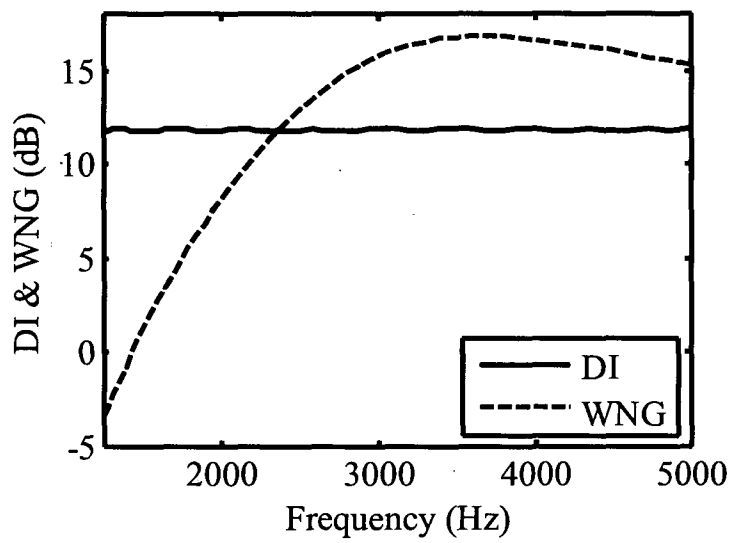


FIG. 25B

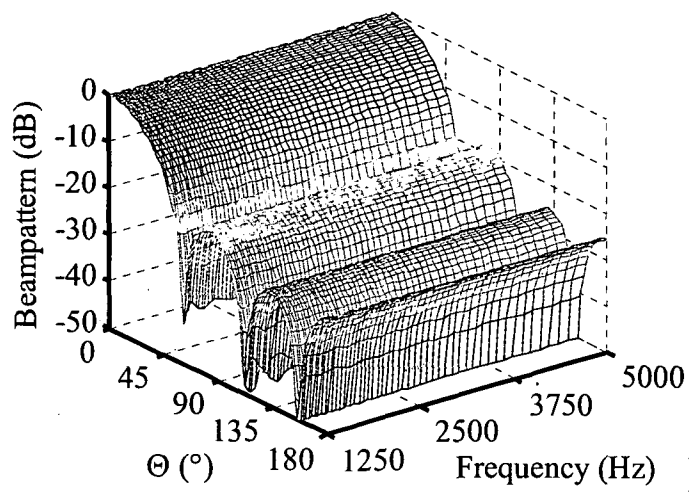


FIG. 25C

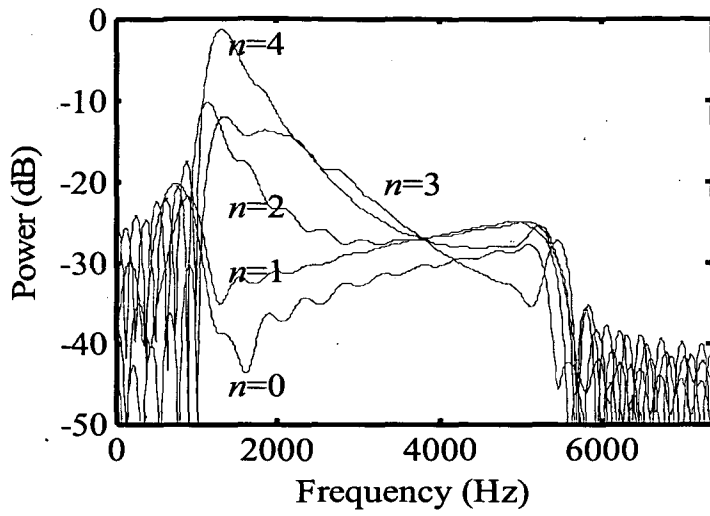


FIG. 26A

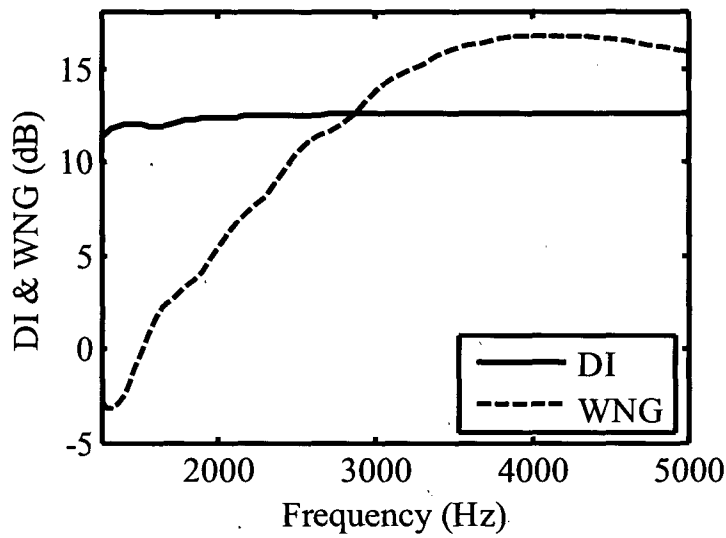


FIG. 26B

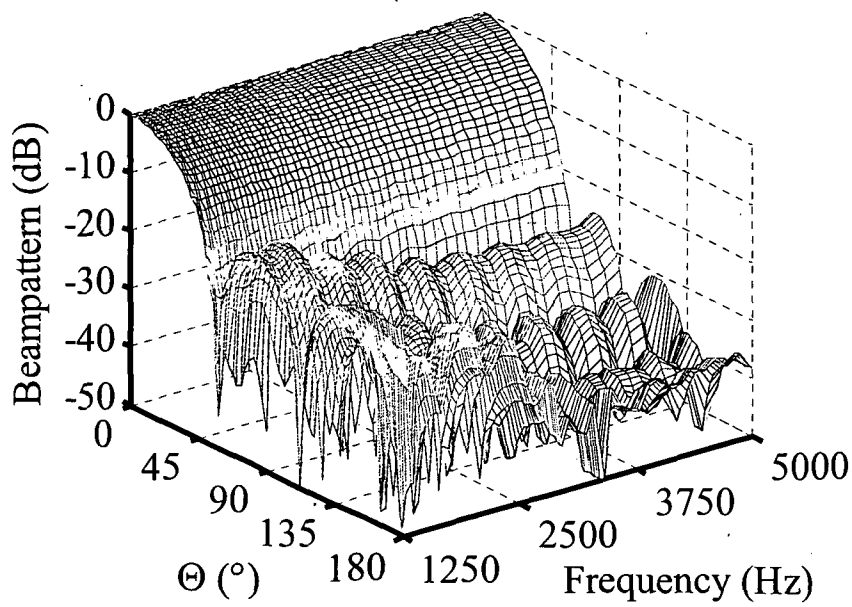


FIG. 26C

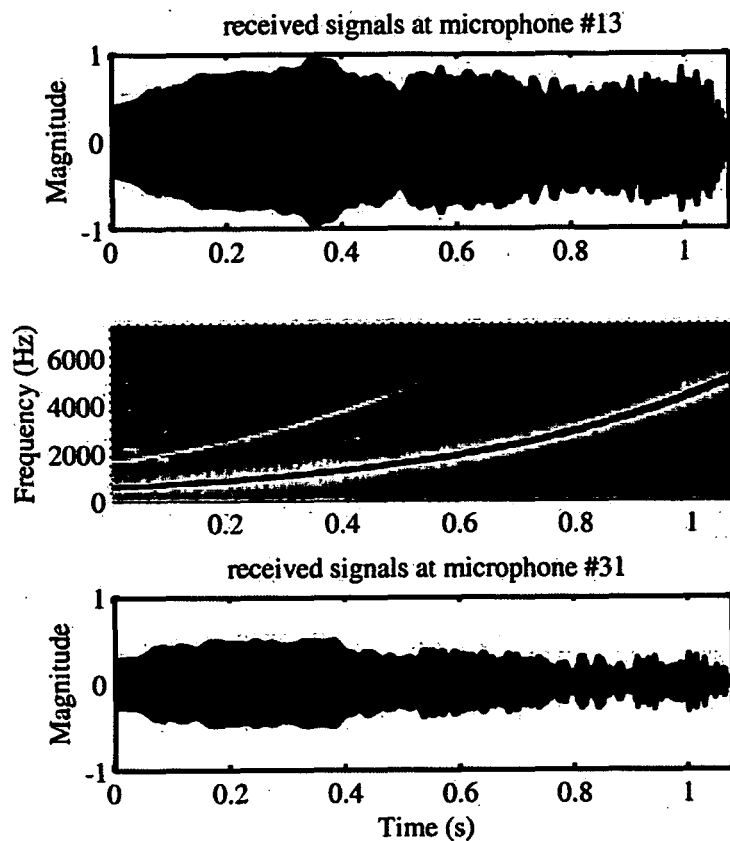


FIG. 27A

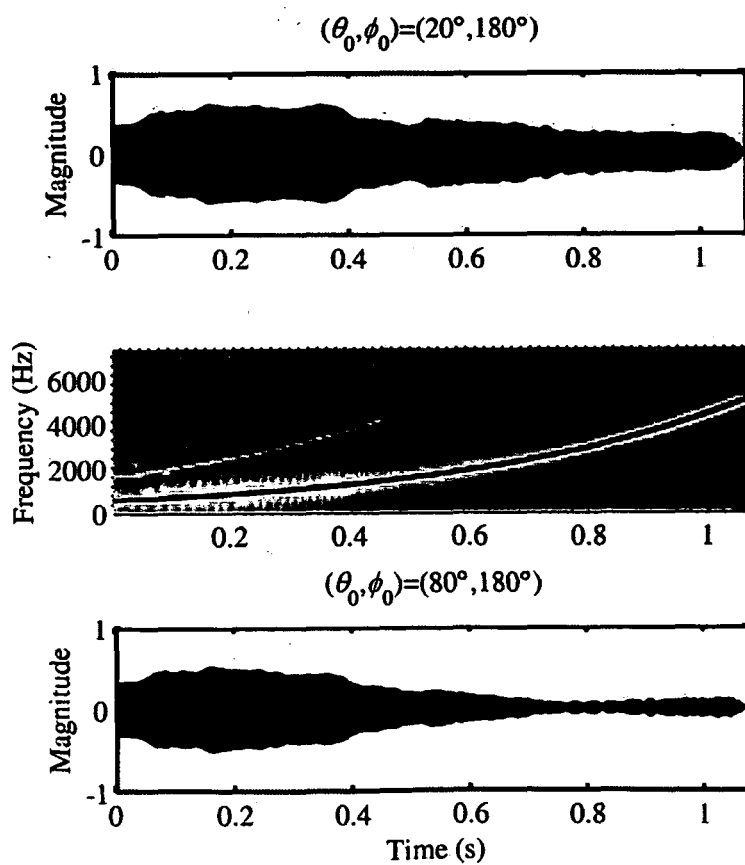


FIG. 27B

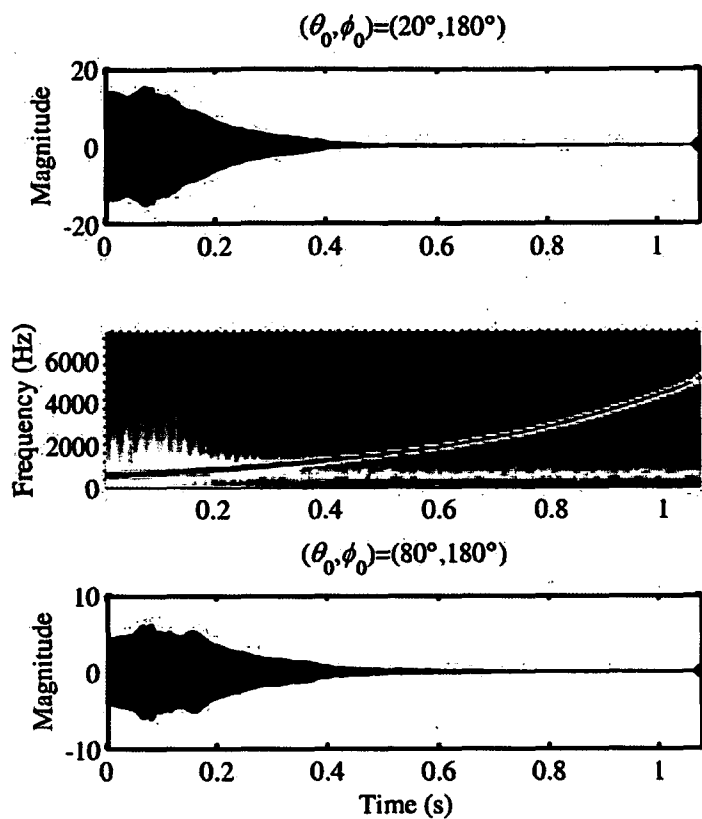


FIG. 27C

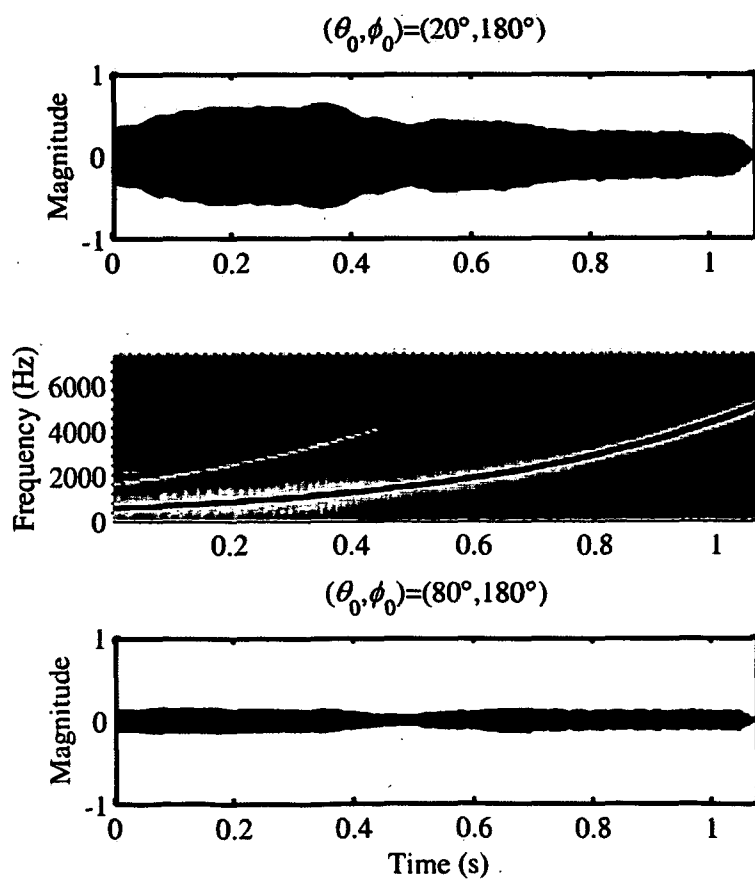


FIG. 27D

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2010/000730

A. CLASSIFICATION OF SUBJECT MATTER

INV. H04R3/00
ADD. H04R3/02 H04R1/40

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)
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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	YAN, SHEFENG; MA, YUANLIANG, HOU, CHAOHUAN: "Optimal array pattern synthesis for broadband arrays" JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, vol. 122, no. 5, November 2007 (2007-11), pages 2686-2696, XP012105274 New York, NY, US ISSN: 0001-4966	30-35
A	abstract section "I. Introduction"; page 2686 section "II. Background"; page 2687 - page 2689 section "III. Proposed Beamformers"; page 2689 - page 2690 section "IV. Convex optimisation based implementation"; page 2690 - page 2691 ----- -/--	1-29

☒ Further documents are listed in the continuation of Box C.

☒ See patent family 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

11 June 2010

Date of mailing of the international search report

28/06/2010

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Guillaume, Mathieu

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2010/000730

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 03/061336 A1 (MH ACOUSTICS LLC [US]; ELKO GARY W [US]; KUBLI ROBERT A [US]; MEYER JE) 24 July 2003 (2003-07-24) cited in the application	30-35
A	figure 2 figure 17 page 6 page 18 - page 26 -----	2-5,9,10
A	ZHIYUN LI ET AL: "Flexible and Optimal Design of Spherical Microphone Arrays for Beamforming" IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, US LNKD- DOI:10.1109/TASL.2006.876764, vol. 15, no. 2, 1 February 2007 (2007-02-01), pages 702-714, XP011157525 ISSN: 1558-7916 cited in the application abstract figures 6,9 section "II. Background"; page 703 - page 704 section "IV. Simplified optimization of desired beampattern for discrete array"; page 707 - page 708 section "IV-A Discrete Spherical Beamformer as A Constrained Linear System"; page 708 - page 709 -----	2-6,8, 13,20,24
A	RAFAELY B: "Spatial Sampling and Beamforming for Spherical Microphone Arrays" HANDS-FREE SPEECH COMMUNICATION AND MICROPHONE ARRAYS, 2008. HSCMA 2008, IEEE, PISCATAWAY, NJ, USA, 6 May 2008 (2008-05-06), pages 5-8, XP031269733 ISBN: 978-1-4244-2337-8 section "3. Spatial sampling"; page 6 - page 7 section "4.5 Beam Pattern with Desired Multiple Nulls "; page 7 ----- -/--	2,3,5, 14-16

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2010/000730

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CHAPMAN D J: "Partial Adaptivity for the Large Array" IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, IEEE SERVICE CENTER, PISCATAWAY, NJ, US LNKD- DOI:10.1109/TAP.1976.1141408, vol. AP-24, no. 5, 1 September 1976 (1976-09-01), pages 685-696, XP002501537 ISSN: 0018-926X abstract figure 2	24
X,P	----- HAOHAI SUN ET AL: "Robust spherical microphone array beamforming with multi-beam-multi-null steering, and sidelobe control" APPLICATIONS OF SIGNAL PROCESSING TO AUDIO AND ACOUSTICS, 2009. WASPAA '09. IEEE WORKSHOP ON, IEEE, PISCATAWAY, NJ, USA, 18 October 2009 (2009-10-18), pages 113-116, XP031575160 ISBN: 978-1-4244-3678-1 the whole document -----	1-23,25, 26,28, 31-35

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2010/000730

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 03061336	A1	24-07-2003	
		AU 2003202945 A1	30-07-2003
		EP 1466498 A1	13-10-2004
		US 2003147539 A1	07-08-2003
		US 2005123149 A1	09-06-2005
		US 2010008517 A1	14-01-2010