



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 1 563 485 B1

(12)

FASCICULE DE BREVET EUROPEEN

(45) Date de publication et mention
de la délivrance du brevet:
29.03.2006 Bulletin 2006/13

(21) Numéro de dépôt: **03782553.6**

(22) Date de dépôt: **13.11.2003**

(51) Int Cl.:
G10H 1/00 (2006.01)

(86) Numéro de dépôt international:
PCT/FR2003/003367

(87) Numéro de publication internationale:
WO 2004/049299 (10.06.2004 Gazette 2004/24)

(54) **PROCEDE DE TRAITEMENT DE DONNEES SONORES ET DISPOSITIF D'ACQUISITION SONORE
METTANT EN OEUVRE CE PROCEDE**

VERFAHREN ZUR VERARBEITUNG VON AUDIODATEIEN UND ERFASSUNGSVORRICHTUNG
ZUR ANWENDUNG DAVON

METHOD FOR PROCESSING AUDIO DATA AND SOUND ACQUISITION DEVICE THEREFOR

(84) Etats contractants désignés:
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR
HU IE IT LI LU MC NL PT RO SE SI SK TR**

(30) Priorité: **19.11.2002 FR 0214444**

(43) Date de publication de la demande:
17.08.2005 Bulletin 2005/33

(73) Titulaire: **FRANCE TELECOM
75015 Paris (FR)**

(72) Inventeur: **DANIEL, Jérôme
F-22710 Penvenan (FR)**

(74) Mandataire: **Fréchède, Michel
Cabinet Plasseraud
65/67 rue de la Victoire
75440 Paris Cedex 09 (FR)**

(56) Documents cités:
**US-A- 4 219 696 US-A- 4 731 848
US-A- 5 452 360 US-A- 5 771 294
US-A- 6 154 553**

• **CHEN J ET AL: "SYNTHESIS OF 3D VIRTUAL
AUDITORY SPACE VIA A SPATIAL FEATURE
EXTRACTION AND REGULARIZATION MODEL"
PROCEEDINGS OF THE VIRTUAL REALITY
ANNUAL INTERNATIONAL SYMPOSIUM.
SEATTLE, SEPT. 18 - 22, 1993, NEW YORK, IEEE,
US, vol. SYMP. 1, 18 septembre 1993 (1993-09-18),
pages 188-193, XP000457687**

Il est rappelé que: Dans un délai de neuf mois à compter de la date de publication de la mention de la délivrance du brevet européen, toute personne peut faire opposition au brevet européen délivré, auprès de l'Office européen des brevets. L'opposition doit être formée par écrit et motivée. Elle n'est réputée formée qu'après paiement de la taxe d'opposition. (Art. 99(1) Convention sur le brevet européen).

EP 1 563 485 B1

Description

[0001] La présente invention concerne le traitement de données sonores.

[0002] Des techniques relatives à la propagation d'une onde sonore dans l'espace tridimensionnel, impliquant notamment une simulation et/ou une restitution sonores spécialisées, mettent en oeuvre des procédés de traitement du signal audio appliqués à la simulation de phénomènes acoustiques et psycho-acoustiques. De tels procédés de traitement prévoient un encodage spatial du champ acoustique, sa transmission et sa reproduction spatialisée sur un ensemble de hauts-parleurs ou sur des écouteurs d'un casque stéréophonique.

[0003] Parmi les techniques de son spatialisé, on distingue deux catégories de traitements complémentaires l'une de l'autre mais qui sont généralement mise en oeuvre, l'une et l'autre, au sein d'un même système.

[0004] D'une part, une première catégorie de traitements concerne les procédés de synthèse d'effet de salle, ou plus généralement d'effets environnementaux. A partir d'une description d'une ou plusieurs sources sonores (signal émis, position, orientation, directivité, ou autre) et en se basant sur un modèle d'effet de salle (impliquant une géométrie de salle, ou encore une perception acoustique souhaitée), on calcule et l'on décrit un ensemble de phénomènes acoustiques élémentaires (ondes directes, réfléchies ou diffractées), ou encore un phénomène acoustique macroscopique (champ réverbéré et diffus), permettant de traduire l'effet spatial au niveau d'un auditeur situé à un point choisi de perception auditive, dans l'espace tridimensionnel. On calcule alors un ensemble de signaux associés typiquement aux réflexions (sources "secondaires", actives par ré-émission d'une onde principale reçue, ayant un attribut de position spatiale) et/ou associés à une réverbération tardive (signaux décorrélés pour un champ diffus).

[0005] D'autre part, une seconde catégorie de procédés concerne le rendu positionnel ou directionnel de sources sonores. Ces procédés sont appliqués à des signaux déterminés par un procédé de la première catégorie décrite ci-avant (impliquant des sources primaires et secondaires) en fonction de la description spatiale (position de la source) qui leur est associée. En particulier, de tels procédés selon cette seconde catégorie permettent d'obtenir des signaux à diffuser sur des hauts-parleurs ou écouteurs, pour finalement donner à un auditeur l'impression auditive de sources sonores placées à des positions respectives prédéterminées, autour de l'auditeur. Les procédés selon cette seconde catégorie sont qualifiés de "*créateurs d'images sonores tridimensionnelles*", du fait de la répartition dans l'espace tridimensionnel du ressenti de la position des sources par un auditeur. Des procédés selon la seconde catégorie comportent généralement une première étape d'encodage spatial des événements acoustiques élémentaires qui produit une représentation du champ sonore dans l'espace tridimensionnel. Dans une seconde étape, cette représentation est transmise ou stockée pour un usage différé. Dans une troisième étape, de décodage, les signaux décodés sont délivrés sur des hauts-parleurs ou des écouteurs d'un dispositif de restitution.

[0006] La présente invention s'inscrit plutôt dans la seconde catégorie précitée. Elle concerne en particulier l'encodage spatial de sources sonores et une spécification de la représentation sonore tridimensionnelle de ces sources. Elle s'applique aussi bien à un encodage de sources sonores "*virtuelles*" (applications où des sources sonores sont simulées telles que des jeux, une conférence spatialisée, ou autres), qu'un encodage "*acoustique*" d'un champ sonore naturel, lors d'une prise de son par un ou plusieurs réseaux tridimensionnels de microphones. Une méthode d'encodage acoustique similaire est présentée par J. Chen et al: "Synthesis of 3D virtual auditory space via a spatial feature extraction and regularisation model", Proceedings of the virtual reality annual international symposium, Seattle, Sept. 18-22, 1993, IEEE, New-York, US, pages 188-193.

[0007] Parmi les techniques envisageables de spatialisation du son, l'approche "*ambisonique*" est préférée. L'encodage ambisonique, qui sera décrit en détail plus loin, consiste à représenter des signaux relatifs à une ou plusieurs ondes sonores dans une base d'harmoniques sphériques (en coordonnées sphériques impliquant notamment un angle d'élévation et un angle azimutal, caractérisant une direction du ou des sons). Les composantes représentant ces signaux et exprimées dans cette base d'harmoniques sphériques sont aussi fonction, pour les ondes émises en champ proche, d'une distance entre la source sonore émettant ce champ et un point correspondant à l'origine de la base des harmoniques sphériques. Plus particulièrement, cette dépendance de la distance s'exprime en fonction de la fréquence sonore, comme on le verra plus loin.

[0008] Cette approche ambisonique offre un grand nombre de fonctionnalités possibles, notamment en terme de simulation de sources virtuelles, et, de manière générale, présente les avantages suivants :

- elle traduit, de façon rationnelle, la réalité des phénomènes acoustiques et apporte un rendu auditif spatial réaliste, convaincant et immersif ;
- la représentation des phénomènes acoustiques est scalable : elle offre une résolution spatiale qui peut être adaptée à différentes situations. En effet, cette représentation peut être transmise et exploitée en fonction de contraintes de débit lors de la transmission des signaux encodés et/ou de limitations du dispositif de restitution ;
- la représentation ambisonique est flexible et il est possible simuler une rotation du champ sonore, ou encore, à la restitution, d'adapter le décodage des signaux ambisoniques à tout dispositif de restitution, de géométries diverses.

[0009] Dans l'approche ambisonique connue, l'encodage des sources virtuelles est essentiellement directionnel. Les fonctions d'encodage reviennent à calculer des gains qui dépendent de l'incidence de l'onde sonore exprimée par les fonctions harmoniques sphériques qui dépendent de l'angle d'élévation et de l'angle azimutal en coordonnées sphériques. En particulier, au décodage, on suppose que les hauts-parleurs, à la restitution, sont lointains. Il en résulte une distorsion (ou une incurvation) de la forme des fronts d'onde reconstruits. En effet, comme indiqué ci-avant, les composantes du signal sonore dans la base des harmoniques sphériques, pour un champ proche, dépendent en fait aussi de la distance de la source et de la fréquence sonore. Plus précisément, ces composantes peuvent s'exprimer mathématiquement sous la forme d'un polynôme dont la variable est inversement proportionnelle à la distance précitée et à la fréquence sonore. Ainsi, les composantes ambisoniques, au sens de leur expression théorique, sont divergentes dans les basses fréquences et, en particulier, tendent vers l'infini quand la fréquence sonore décroît vers zéro, lorsqu'elles représentent un son en champ proche émis par une source située à une distance finie. Ce phénomène mathématique est connu, dans le domaine de la représentation ambisonique, déjà pour l'ordre 1, par le terme de "bass boost", notamment par :

- M.A.GERZON, "General Metatheory of Auditory Localisation", preprint 3306 of the 92nd AES Convention, 1992, page 52.

Ce phénomène devient particulièrement critique pour des ordres d'harmoniques sphériques élevées impliquant des polynômes de puissance élevée.

[0010] On connaît par :

SONTACCHI et HÖLDRICH, "Further Investigations on 3D Sound Fields using Distance Coding" (Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-01), Limerick, Irlande, 6-8 Décembre 2001), une technique pour prendre en compte une incurvation des fronts d'ondes au sein d'une représentation proche d'une représentation ambisonique, dont le principe consiste à :

- appliquer un encodage ambisonique (d'ordre élevé) aux signaux issus d'une prise de son virtuelle (simulée), de type WFS (pour "Wave Field Synthesis") ;
- et reconstruire le champ acoustique sur une zone d'après ses valeurs sur une frontière de zone, se fondant ainsi sur le principe de HUYGENS-FRESNEL.

[0011] Cependant, la technique présentée dans ce document, bien que prometteuse du fait qu'elle utilise une représentation ambisonique à un ordre élevé, pose un certain nombre de problèmes :

- les ressources informatiques nécessaires pour le calcul de toutes les surfaces permettant d'appliquer le principe de HUYGENS-FRESNEL, ainsi que les temps de calcul nécessaires, sont excessifs ;
- des artefacts de traitement dits d'"aliasing spatial" apparaissent à cause de la distance entre les microphones, à moins de choisir un maillage de microphone virtuels serré dans l'espace, ce qui alourdit les traitements ;
- cette technique est difficilement transposable à un cas réel de capteurs à disposer en réseau, en présence d'une source réelle, à l'acquisition ;
- à la restitution, la représentation sonore tridimensionnelle est implicitement assujettie à un rayon figé du dispositif de restitution car le décodage ambisonique doit se faire, ici, sur un réseau de hauts-parleurs de mêmes dimensions que le réseau de microphones initial, ce document ne proposant aucun moyen d'adapter l'encodage ou le décodage à d'autres tailles de dispositifs de restitution.

[0012] Surtout, ce document présente un réseau horizontal de capteurs, ce qui suppose que les phénomènes acoustiques dont on tient compte, ici, ne se propagent que dans des directions horizontales, ce qui exclut toute autre direction de propagation et qui, donc, ne représente pas la réalité physique d'un champ acoustique ordinaire.

[0013] De façon plus générale, les techniques actuelles ne permettent pas de traiter de façon satisfaisante tout type de sources sonores, notamment en champ proche, mais plutôt des sources sonores lointaines (ondes planes), ce qui correspond à une situation restrictive et artificielle dans nombreuses applications.

[0014] Un objet de la présente invention est de fournir un procédé pour traiter, par encodage, transmission et restitution, un type quelconque de champ sonore, en particulier l'effet d'une source sonore en champ proche.

[0015] Un autre objet de la présente invention est de fournir un procédé permettant l'encodage de sources virtuelles, non seulement en direction, mais aussi en distance, et de définir un décodage adaptable à un dispositif de restitution quelconque.

[0016] Un autre objet de la présente invention est de fournir un procédé de traitement robuste pour des sons de toutes fréquences sonores (y compris les basses fréquences), notamment pour la prise de son de champs acoustiques naturels

à l'aide de réseaux tridimensionnels de microphones.

[0017] A cet effet, la présente invention propose un procédé de traitement de données sonores, dans lequel :

- a) on code des signaux représentatifs d'au moins un son se propageant dans l'espace tridimensionnel et issu d'une source située à une première distance d'un point de référence, pour obtenir une représentation du son par des composantes exprimées dans une base d'harmoniques sphériques, d'origine correspondant audit point de référence, et
- b) on applique auxdites composantes une compensation d'un effet de champ proche par un filtrage qui est fonction d'une seconde distance définissant sensiblement, pour une restitution du son par un dispositif de restitution, une distance entre un point de restitution et un point de perception auditive.

[0018] Dans un premier mode de réalisation, ladite source étant lointaine du point de référence,

- on obtient des composantes d'ordres successifs m pour la représentation du son dans ladite base d'harmoniques sphériques, et
- on applique un filtre dont les coefficients, appliqués chacun à une composante d'ordre m , s'expriment analytiquement sous la forme de l'inverse d'un polynôme de puissance m , dont la variable est inversement proportionnelle à la fréquence sonore et à ladite seconde distance, pour compenser un effet de champ, proche au niveau du dispositif de restitution.

[0019] Dans un second mode de réalisation, ladite source étant une source virtuelle prévue à ladite première distance,

- on obtient des composantes d'ordres successifs m pour la représentation du son dans ladite base d'harmoniques sphériques, et
- on applique un filtre global dont les coefficients, appliqués chacun à une composante d'ordre m , s'expriment analytiquement sous la forme d'une fraction, dont :
 - le numérateur est un polynôme de puissance m , dont la variable est inversement proportionnelle à la fréquence sonore et à ladite première distance, pour simuler un effet de champ proche de la source virtuelle, et
 - le dénominateur est un polynôme de puissance m , dont la variable est inversement proportionnelle à la fréquence sonore et à ladite seconde distance, pour compenser l'effet du champ proche de la source virtuelle dans les basses fréquences sonores.

[0020] Préférentiellement, on transmet au dispositif de restitution les données codées et filtrées aux étapes a) et b) avec un paramètre représentatif de ladite seconde distance.

[0021] En complément ou en variante, le dispositif de restitution comportant des moyens de lecture d'un support mémoire, on mémorise sur un support mémoire destiné à être lu par le dispositif de restitution les données codées et filtrées aux étapes a) et b) avec un paramètre représentatif de ladite seconde distance.

[0022] Avantageusement, préalablement à une restitution sonore par un dispositif de restitution comportant une pluralité de hauts-parleurs disposés à une troisième distance dudit point de perception auditive, on applique aux données codées et filtrées un filtre d'adaptation dont les coefficients sont fonction desdites seconde et troisième distances.

[0023] Dans une réalisation particulière, les coefficients de ce filtre d'adaptation, appliqués chacun à une composante d'ordre m , s'expriment analytiquement sous la forme d'une fraction, dont :

- le numérateur est un polynôme de puissance m , dont la variable est inversement proportionnelle à la fréquence sonore et à ladite seconde distance,
- et le dénominateur est un polynôme de puissance m , dont la variable est inversement proportionnelle à la fréquence sonore et à ladite troisième distance.

[0024] Avantageusement, pour la mise en oeuvre de l'étape b), on prévoit :

- pour des composantes d'ordre m pair, des filtres audionumériques sous la forme d'une cascade de cellules d'ordre deux ; et
- pour des composantes d'ordre m impair, des filtres audionumériques sous la forme d'une cascade de cellules d'ordre deux et une cellule supplémentaire d'ordre un.

[0025] Dans cette réalisation, les coefficients d'un filtre audionumérique, pour une composante d'ordre m , sont définis à partir des valeurs numériques des racines desdits polynômes de puissance m .

[0026] Dans une réalisation particulière, les polynômes précités sont des polynômes de Bessel.

[0027] A l'acquisition des signaux sonores, on prévoit avantageusement un microphone comportant un réseau de transducteurs acoustiques agencés sensiblement sur la surface d'une sphère dont le centre correspond sensiblement audit point de référence, pour obtenir lesdits signaux représentatifs d'au moins un son se propageant dans l'espace tridimensionnel.

[0028] Dans cette réalisation, on applique à l'étape b) un filtre global pour, d'une part, compenser un effet de champ proche en fonction de ladite seconde distance et, d'autre part, égaliser les signaux issus des transducteurs pour compenser une pondération de directivité desdits transducteurs.

[0029] Préférentiellement, on prévoit un nombre de transducteurs fonction d'un nombre total choisi de composantes pour représenter le son dans ladite base d'harmoniques sphériques.

[0030] Selon une caractéristique avantageuse, on choisit à l'étape a) un nombre total de composantes dans la base des harmoniques sphériques pour obtenir, à la restitution, une région de l'espace autour du point de perception dans laquelle la restitution du son est fidèle et dont les dimensions sont croissantes avec le nombre total de composantes.

[0031] Préférentiellement, on prévoit en outre un dispositif de restitution comportant un nombre de haut-parleurs au moins égal audit nombre total de composantes.

[0032] En variante, dans le cadre d'une restitution avec synthèse binaurale ou transaurale :

- on prévoit un dispositif de restitution comportant au moins un premier et un second haut-parleur disposés à une distance choisie d'un auditeur,
- on obtient, pour cet auditeur, une information de ressenti attendu de la position dans l'espace de sources sonores situées à une distance de référence prédéterminée de l'auditeur pour l'application d'une technique dite de "*synthèse binaurale*" ou "*transaurale*", et
- on applique la compensation de l'étape b) avec ladite distance de référence sensiblement en tant que seconde distance.

[0033] Dans une variante où l'on introduit une adaptation au dispositif de restitution à deux écouteurs :

- on prévoit un dispositif de restitution comportant au moins un premier et un second haut-parleur disposés à une distance choisie d'un auditeur,
- on obtient, pour cet auditeur, une information de ressenti de la position dans l'espace de sources sonores situées à une distance de référence prédéterminée de l'auditeur, et
- préalablement à une restitution sonore par le dispositif de restitution, on applique aux données codées et filtrées aux étapes a) et b) un filtre d'adaptation dont les coefficients sont fonction de la seconde distance et sensiblement de la distance de référence.

[0034] En particulier, dans le cadre d'une restitution avec synthèse binaurale :

- le dispositif de restitution comporte un casque à deux écouteurs pour les oreilles respectives de l'auditeur,
- et préférentiellement, séparément pour chaque écouteur, on applique le codage et le filtrage des étapes a) et b) pour des signaux respectifs destinés à alimenter chaque écouteur, avec, en tant que première distance, respectivement une distance séparant chaque oreille d'une position d'une source à restituer dans l'espace de restitution.

[0035] Préférentiellement, on met en forme, aux étapes a) et b), un système matriciel comportant au moins :

- une matrice comportant lesdites composantes dans la base des harmoniques sphériques, et
- une matrice diagonale dont les coefficients correspondent à des coefficients de filtrage de l'étape b),

et on multiplie lesdites matrices pour obtenir une matrice résultat de composantes compensées.

[0036] De préférence, à la restitution :

- le dispositif de restitution comporte une pluralité de haut-parleurs disposés sensiblement à une même distance du point de perception auditive, et
- pour décoder lesdites données codées et filtrées aux étapes a) et b) et former des signaux adaptés pour alimenter lesdits haut-parleurs :

* on forme un système matriciel comportant ladite matrice résultat de composantes compensées et une matrice de décodage prédéterminée, propre au dispositif de restitution, et

* on obtient une matrice comportant des coefficients représentatifs des signaux d'alimentation des haut-parleurs

par multiplication de la matrice résultat par ladite matrice de décodage.

[0037] La présente invention vise aussi un dispositif d'acquisition sonore, comportant un microphone muni d'un réseau de transducteurs acoustiques disposés sensiblement sur la surface d'une sphère. Selon l'invention le dispositif comporte en outre une unité de traitement agencée pour :

- recevoir des signaux émanant chacun d'un transducteur,
- appliquer auxdits signaux un codage pour obtenir une représentation du son par des composantes exprimées dans une base d'harmoniques sphériques, d'origine correspondant au centre de ladite sphère,
- et appliquer auxdites composantes un filtrage qui est fonction, d'une part, d'une distance correspondant au rayon de la sphère et, d'autre part, d'une distance de référence.

[0038] Préférentiellement, le filtrage effectué par l'unité de traitement consiste, d'une part, à égaliser, en fonction du rayon de la sphère, les signaux issus des transducteurs pour compenser une pondération de directivité desdits transducteurs et, d'autre part, à compenser un effet de champ proche en fonction de ladite distance de référence.

[0039] D'autres avantages et caractéristiques de l'invention apparaîtront à la lecture de la description détaillée ci-après et à l'examen des figures qui l'accompagnent, sur lesquelles :

- la figure 1 illustre schématiquement un système d'acquisition et création, par simulation de sources virtuelles, de signaux sonores, avec encodage, transmission, décodage et restitution par un dispositif de restitution spatialisé,
- la figure 2 représente plus précisément un encodage de signaux définis à la fois en intensité et par rapport à la position d'une source dont ils sont issus,
- la figure 3 illustre les paramètres en jeu dans la représentation ambisonique, en coordonnées sphériques ;
- la figure 4 illustre une représentation par une métrique tridimensionnelle dans un repère de coordonnées sphériques,

d'harmoniques sphériques Y_{mn}^{σ} de différents ordres ;

- la figure 5 est un diagramme des variations du module de fonctions radiales $j_m(kr)$, qui sont des fonctions de Bessel sphériques, pour des valeurs d'ordre m successives, ces fonctions radiales intervenant dans la représentation ambisonique d'un champ de pression acoustique ;
- la figure 6 représente l'amplification due à l'effet de champ proche pour différents ordres successifs m , en particulier dans les basses fréquences ;
- la figure 7 représente schématiquement un dispositif de restitution comportant une pluralité de hauts-parleurs HP_i , avec le point (référéncé P) de perception auditive précité, la première distance précitée (référéncée p) et la seconde distance précitée (référéncée R) ;
- la figure 8 représente schématiquement les paramètres mis en jeu dans l'encodage ambisonique, avec un encodage directionnel, ainsi qu'un encodage de distance selon l'invention ;
- la figure 9 représente des spectres d'énergie des filtres de compensation et de champ proche simulés pour une première distance d'une source virtuelle $p = 1$ m et une pré-compensation de hauts-parleurs situés à une seconde distance $R = 1,5$ m ;
- la figure 10 représente des spectres d'énergie des filtres de compensation et de champ proche simulés pour une première distance de la source virtuelle $p = 3$ m et une pré-compensation de hauts-parleurs situés à une distance $R = 1,5$ m ;
- la figure 11A représente une reconstruction du champ proche avec compensation, au sens de la présente invention, pour une onde sphérique dans le plan horizontal ;
- la figure 11B, à comparer avec la figure 11A, représente le front d'onde initial, issu d'une source S ;
- la figure 12 représente schématiquement un module de filtrage pour adapter les composantes ambisoniques reçues et pré-compensées à l'encodage pour une distance de référence R en tant que seconde distance, à un dispositif de restitution comportant une pluralité de hauts-parleurs disposés à une troisième distance R_2 d'un point de perception auditive ;
- la figure 13A représente schématiquement la disposition d'une source sonore M , à la restitution, pour un auditeur utilisant un dispositif de restitution appliquant une synthèse binaurale, avec une source émettant en champ proche ;
- la figure 13B représente schématiquement les étapes d'encodage et de décodage avec effet de champ proche dans le cadre de la synthèse binaurale de la figure 13A à laquelle est combiné un encodage/décodage ambisonique ;
- la figure 14 représente schématiquement le traitement des signaux issus d'un microphone comportant une pluralité de capteurs de pression agencés sur une sphère, à titre illustratif, par encodage ambisonique, égalisation et compensation de champ proche au sens de l'invention.

[0040] On se réfère tout d'abord à la figure 1 qui représente à titre illustratif un système global de spatialisation sonore. Un module de simulation d'une scène virtuelle définit un objet sonore comme une source virtuelle d'un signal, par exemple monophonique, de position choisie dans l'espace tridimensionnel et qui définit une direction du son. Il peut être prévu en outre des spécifications de la géométrie d'une salle virtuelle, pour simuler une réverbération du son. Un module de traitement 11 applique une gestion d'une ou plusieurs de ces sources par rapport à un auditeur (définition d'une position virtuelle des sources par rapport à cet auditeur). Il met en oeuvre un processeur d'effet de salle pour simuler des réverbérations ou autres en appliquant des retards et/ou des filtrages usuels. Les signaux ainsi construits sont transmis à un module 2a d'encodage spatial des contributions élémentaires des sources.

[0041] Parallèlement, une prise de son naturelle peut être effectuée dans le cadre d'un enregistrement sonore par un ou plusieurs microphones disposés de façon choisie par rapport aux sources réelles (module 1b). Les signaux captés par les microphones sont encodés par un module 2b. Les signaux acquis et encodés peuvent être transformés selon un format de représentation intermédiaire (module 3b), avant d'être mixés par le module 3 aux signaux générés par le module 1a et encodés par le module 2a (issu des sources virtuelles). Les signaux mixés sont ensuite transmis, ou encore mémorisés sur un support, en vue d'une restitution ultérieure (flèche TR). Ils sont ensuite appliqués à un module de décodage 5, en vue de la restitution sur un dispositif de restitution 6 comportant des hauts-parleurs. Le cas échéant, l'étape de décodage 5 peut être précédée d'une étape de manipulation du champ sonore, par exemple par rotation, grâce à un module de traitement 4 prévu en amont du module de décodage 5.

[0042] Le dispositif de restitution peut se présenter sous la forme d'une multiplicité de hauts-parleurs, agencés par exemple à la surface d'une sphère dans une configuration tridimensionnelle (périphonique) pour assurer, à la restitution, notamment un ressenti d'une direction du son dans l'espace tridimensionnel. A cet effet, un auditeur se place généralement au centre de la sphère formée par le réseau de haut-parleurs, ce centre correspondant au point de perception auditive cité ci-avant. En variante, les hauts-parleurs du dispositif de restitution peuvent être agencés dans un plan (configuration panoramique bidimensionnelle), les hauts-parleurs étant disposés en particulier sur un cercle et l'auditeur se plaçant habituellement au centre de ce cercle. Dans une autre variante, le dispositif de restitution peut se présenter sous la forme d'un dispositif de type "surround" (5.1). Enfin, dans une variante avantageuse, le dispositif de restitution peut se présenter sous la forme d'un casque à deux écouteurs pour une synthèse binaurale du son restitué, qui permet à l'auditeur de ressentir une direction des sources dans l'espace tridimensionnel, comme on le verra plus loin de façon détaillée. Un tel dispositif de restitution à deux hauts-parleurs, pour un ressenti dans l'espace tridimensionnel, peut se présenter aussi sous la forme d'un dispositif de restitution transaurale, à deux hauts-parleurs disposés à une distance choisie d'un auditeur.

[0043] On se réfère maintenant à la figure 2 pour décrire un encodage spatial et un décodage pour une restitution sonore tridimensionnelle, de sources sonores élémentaires. On transmet à un module d'encodage spatial 2 le signal issu d'une source 1 à N, ainsi que sa position (réelle ou virtuelle). Sa position peut être aussi bien définie en terme d'incidence (direction de la source vue de l'auditeur) qu'en terme de distance entre cette source et un auditeur. La pluralité des signaux ainsi encodés permet d'obtenir une représentation multi-canal d'un champ sonore global. Les signaux encodés sont transmis (flèche TR) à un dispositif de restitution sonore 6, pour une restitution sonore dans l'espace tridimensionnel, comme indiqué ci-avant en référence à la figure 1.

[0044] On se réfère maintenant à la figure 3 pour décrire ci-après la représentation ambisonique par des harmoniques sphériques dans l'espace tridimensionnel, d'un champ acoustique. On considère une zone autour d'une origine O (sphère de rayon R) exempte de source acoustique. On adopte un système de coordonnées sphériques dans lequel chaque vecteur $\vec{r}$ dès l'origine O à un point de la sphère est décrit par un azimuth θ_r , une élévation δ_r et un rayon r (correspondant à la distance à l'origine O).

[0045] Le champ de pression $p(\vec{r})$ à l'intérieur de cette sphère ($r < R$ où R est le rayon de la sphère) peut s'écrire dans le domaine fréquentiel comme une série dont les termes sont les produits pondérés de fonctions angulaires $y_{mn}^\sigma(\theta, \delta)$ et de fonction radiale $j_m(kr)$ qui dépendent ainsi d'un terme de propagation où $k=2\pi f/c$, où f est la fréquence sonore et c est la vitesse du son dans le milieu de propagation.

[0046] Le champ de pression s'exprime alors par :

$$p(\vec{r}) = \sum_{m=0}^{\infty} j_m^m j_m(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{mn}^\sigma Y_{mn}^{\sigma(N3D)}(\theta_r, \delta_r) \quad [A1]$$

[0047] L'ensemble des facteurs de pondération B_{mn}^σ , qui sont implicitement fonction de la fréquence, décrivent ainsi le champ de pression dans la zone considérée. Pour cette raison, ces facteurs sont appelés "*composantes harmoniques sphériques*" et représentent une expression fréquentielle du son (ou du champ de pression) dans la base des harmoniques sphériques Y_{mn}^σ .

[0048] Les fonctions angulaires sont appelées "*harmoniques sphériques*" et sont définies par :

[A2]

$$Y_{mn}^\sigma(\theta, \delta) = \sqrt{2m+1} \sqrt{\frac{(2-\delta_{0,n})}{(m+n)!}} P_{mn}(\sin \delta) \times \begin{cases} \cos n\theta & \text{si } \sigma = +1 \\ \sin n\theta & \text{si } \sigma = -1 \end{cases}$$

où $P_{mn}(\sin \delta)$ sont des fonctions de Legendre de degré m et d'ordre n ;

$\delta_{p,q}$ est le symbole de Kröneckner (égal à 1 si $p=q$ et 0, sinon)

[0049] Les harmoniques sphériques forment une base orthonormée où les produits scalaires entre composantes harmoniques et, de façon générale entre deux fonctions F et G , sont respectivement définies par :

$$\left\langle Y_{mn}^\sigma \middle| Y_{m'n'}^{\sigma'} \right\rangle_{4\pi} = \delta_{mm'} \delta_{nn'} \delta_{\sigma\sigma'}$$

[A'2]

$$\langle F|G \rangle_{4\pi} = \frac{1}{4\pi} \oint\!\!\!\oint F(\theta, \delta) G(\theta, \delta), d\Omega(\theta, \delta)$$

[0050] Les harmoniques sphériques sont des fonctions réelles bornées, comme représenté sur la figure 4, en fonction de l'ordre m et des indices n et σ . Les parties sombres et claires correspondent respectivement aux valeurs positives et négatives des fonctions harmoniques sphériques. Plus l'ordre m est élevé et plus la fréquence angulaire (et donc la discrimination entre fonctions) est élevée. Les fonctions radiales $j_m(kr)$ sont des fonctions de Bessel sphériques, dont le module est illustré pour quelques valeurs de l'ordre m dans la figure 5.

[0051] On peut donner une interprétation de la représentation ambisonique par une base d'harmoniques sphériques comme suit. Les composantes ambisoniques de même ordre m expriment finalement des "dérivées" ou des "moments" d'ordre m du champ de pression au voisinage de l'origine O (centre de la sphère représentée sur la figure 3).

[0052] En particulier, $B_{00}^{+1} = W$ décrit la grandeur scalaire de la pression, tandis que

$B_{11}^{+1} = X, B_{11}^{-1} = Y, B_{10}^{+1} = Z$ sont liés aux gradients de pression (ou encore à la vitesse particulière), à l'origine

O . Ces quatre premières composantes W, X, Y et Z sont obtenues lors d'une prise de son naturelle à l'aide de microphones omnidirectionnels (pour la composante W d'ordre 0) et bidirectionnels (pour les trois autres composantes suivantes). En utilisant un plus grand nombre de transducteurs acoustiques, un traitement approprié, notamment par égalisation, permet d'obtenir d'avantage de composantes ambisoniques (ordres m plus élevés supérieurs à 1).

[0053] En prenant en compte des composantes supplémentaires d'ordre plus élevé (supérieur à 1), donc en augmentant la résolution angulaire de la description ambisonique, on accède à une approximation du champ de pression sur un voisinage plus large au regard de la longueur d'onde de l'onde sonore, autour de l'origine O . On comprendra ainsi qu'il existe une relation étroite entre la résolution angulaire (ordre des harmoniques sphériques) et la portée radiale (rayon r) qui peut être représentée. En bref, lorsque l'on s'écarte spatialement du point d'origine O de la figure 3, plus le nombre de composantes ambisoniques est élevé (ordre M élevé) et meilleure est la représentation du son par l'ensemble de ces composantes ambisoniques. On comprendra aussi que la représentation ambisonique du son est toutefois moins satisfaisante au fur et à mesure que l'on s'éloigne de l'origine O . Cet effet devient critique en particulier

pour des fréquences sonores élevées (de longueur d'onde courte). On a donc intérêt à obtenir un nombre de composantes ambisoniques qui soit le plus grand possible, ce qui permet de créer une région de l'espace autour du point de perception, dans laquelle la restitution du son est fidèle et dont les dimensions sont croissantes avec le nombre total de composantes.

[0054] On décrit ci-après une application à système d'encodage/transmission/restitution d'un son spatialisé.

[0055] En pratique, un système ambisonique prend en compte un sous-ensemble de composantes harmoniques sphériques, comme décrit ci-avant. On parle d'un système d'ordre M lorsque celui-ci prend en compte des composantes ambisoniques d'indice $m < M$. Lorsqu'il s'agit d'une restitution par un dispositif de restitution à hauts-parleurs, on comprendra que si ces hauts-parleurs sont disposés dans un plan horizontal, seules les harmoniques d'indice $m=n$ sont exploitées. En revanche, lorsque le dispositif de restitution comporte des hauts-parleurs disposés sur la surface d'une sphère ("*périphonie*"), on peut en principe exploiter autant d'harmoniques qu'il existe de haut-parleurs.

[0056] On désigne par la référence S le signal de pression porté par une onde plane et capté au point O correspondant au centre de la sphère de la figure 3 (origine de la base en coordonnées sphériques). L'incidence de l'onde est décrite par l'azimut θ et l'élévation δ . L'expression des composantes du champ associé à cette onde plane est donnée par la relation :

$$B_{mn}^{\sigma} = S Y_{mn}^{\sigma}(\theta, \delta) \quad [\text{A3}]$$

[0057] Pour encoder (simuler) une source en champ proche à une distance p de l'origine O , on applique un filtre

$F_m^{(\rho/c)}$ pour "incurver" la forme des fronts d'onde, en considérant qu'un champ proche émet, en première approximation, une onde sphérique. Les composantes encodées du champ deviennent :

$$B_{mn}^{\sigma} = S F_m^{(\rho/c)}(\omega) Y_{mn}^{\sigma}(\theta, \delta) \quad [\text{A4}]$$

et l'expression du filtre précité $F_m^{(\rho/c)}$ est donnée par la relation :

$$F_m^{(\rho/c)}(\omega) = \sum_{n=0}^m \frac{(m+n)!}{(m-n)!n!} (2j\omega\rho/c)^{-n} \quad [\text{A5}]$$

où $\omega = 2\pi f$ est la pulsation de l'onde, f étant la fréquence du son.

[0058] Ces deux dernières relations [A4] et [A5] montrent finalement que, aussi bien pour une source virtuelle (simulée) que pour une source réelle en champ proche, les composantes du son dans la représentation ambisonique s'expriment mathématiquement (en particulier analytiquement) sous la forme d'un polynôme, ici de Bessel, de puissance m et dont la variable $(c/2j\omega\rho)$ est inversement proportionnelle à la fréquence sonore.

[0059] Ainsi, on comprendra que :

- dans le cas d'une onde plane, l'encodage produit des signaux qui ne diffèrent du signal d'origine que d'un gain réel, fini, ce qui correspond à un encodage purement directionnel (relation [A3]) ;
- dans le cas d'une onde sphérique (source en champ proche), le filtre supplémentaire $F_m^{(\rho/c)}(\omega)$ encode l'information de distance en introduisant, dans l'expression des composantes ambisoniques, des rapports d'amplitudes complexes qui dépendent de la fréquence, comme exprimé dans la relation [A5].

[0060] Il est à noter que ce filtre supplémentaire est de type "intégrateur", avec un effet d'amplification croissant et divergent (non-borné) au fur et à mesure que les fréquences sonores décroissent vers zéro. La figure 6 montre, pour chaque ordre m , une augmentation du gain en basses fréquences (ici la première distance $\rho = 1\text{m}$). Il s'agit donc de filtres instables et divergents lorsque l'on cherche à les appliquer à des signaux audio quelconques. Cette divergence est d'autant plus critique pour les ordres m de valeur élevée.

[0061] On comprendra en particulier, à partir des relations [A3], [A4], et [A5], que la modélisation d'une source virtuelle en champ proche présente des composantes ambisoniques divergentes en basses fréquences, de façon particulièrement critique pour des ordres m élevés, comme représenté sur la figure 6. Cette divergence, dans les basses fréquences, correspond au phénomène de "bass boost" énoncé ci-avant. Il se manifeste aussi en acquisition sonore, pour des sources réelles.

[0062] Pour cette raison notamment, l'approche ambisonique, en particulier pour des ordres m élevés, n'a pas connu, dans l'état de la technique, une application concrète (autre que théorique) dans le traitement du son.

[0063] On comprend en particulier qu'une compensation du champ proche est nécessaire pour respecter, à la restitution, la forme des fronts d'ondes encodés dans la représentation ambisonique. En se référant à la figure 7, un dispositif de restitution comporte une pluralité de hauts-parleurs HP_i , disposés à une même distance R , dans l'exemple décrit, d'un point de perception auditive P . Sur cette figure 7 :

- chaque point où se situe un haut-parleur HP_i correspond à un point de restitution énoncé ci-avant,
- le point P est le point de perception auditive énoncé ci-avant,
- ces points sont séparés de la seconde distance R énoncée ci-avant,

tandis que sur la figure 3 décrite ci-avant :

- le point O correspond au point de référence, énoncé ci-avant, qui forme l'origine de la base des harmoniques sphériques,
- le point M correspond à la position d'une source (réelle ou virtuelle) située à la première distance p , énoncée ci-avant, du point de référence O .

[0064] Selon l'invention, on introduit une pré-compensation du champ proche au stade même de l'encodage, cette

compensation mettant en jeu des filtres de la forme analytique $\frac{1}{F_m^{(R/c)}(\omega)}$ et qui s'appliquent aux composantes

ambisoniques B_{mn}^σ précitées.

[0065] Selon l'un des avantages que procure l'invention, l'amplification $F_m^{(p/c)}(\omega)$ dont l'effet apparaît sur la figure

6 est compensée par l'atténuation du filtre appliqué dès l'encodage $\frac{1}{F_m^{(R/c)}(\omega)}$. En particulier, les coefficients de

ce filtre de compensation $\frac{1}{F_m^{(R/c)}(\omega)}$ sont croissants avec la fréquence du son et, en particulier, tendent vers zéro,

pour les basses fréquences. Avantageusement, cette pré-compensation, effectuée dès l'encodage, assure que les données transmises ne sont pas divergentes pour les basses fréquences.

[0066] Pour indiquer la signification physique de la distance R qui intervient dans le filtre de compensation, on considère, à titre illustratif, une onde plane réelle, initiale, à l'acquisition des signaux sonores. Pour simuler un effet de champ proche de cette source lointaine, on applique le premier filtre de la relation [A5], comme indiqué dans la relation [A4]. La distance p représente alors une distance entre une source virtuelle proche M et le point O représentant l'origine de la base sphérique de la figure 3. On applique ainsi un premier filtre de simulation de champ proche pour simuler la présence d'une source virtuelle à la distance p décrite ci-avant. Néanmoins, d'une part, comme indiqué ci-avant, les termes du coefficient de ce filtre divergent dans les basses fréquences (figure 6) et, d'autre part, la distance p précitée ne représentera pas forcément la distance entre les hauts-parleurs d'un dispositif de restitution et un point P de perception (figure 7). Selon l'invention, on applique une pré-compensation, à l'encodage, mettant en jeu un filtre de type

$\frac{1}{F_m^{(R/c)}(\omega)}$ comme indiqué ci-avant, ce qui permet, d'une part, de transmettre des signaux bornés, et, d'autre part,

de choisir la distance R , dès l'encodage, pour la restitution du son à partir des hauts-parleurs HP_i , tel que représenté sur la figure 7. En particulier, on comprendra que si l'on a simulé, à l'acquisition, une source virtuelle placée à la distance p de l'origine O , à la restitution (figure 7), un auditeur placé au point P de perception auditive (à une distance R des hauts-parleurs HP_i) ressentira, à l'audition, la présence d'une source sonore S , placée à la distance p du point de

perception P et qui correspond à la source virtuelle simulée lors de l'acquisition.

[0067] Ainsi, la pré-compensation du champ proche des hauts-parleurs (placés à la distance R), au stade de l'encodage, peut être combinée à un effet de champ proche simulé d'une source virtuelle placée à une distance p. A l'encodage, on met finalement en jeu un filtre total résultant, d'une part, de la simulation du champ proche, et, d'autre part, de la compensation du champ proche, les coefficients de ce filtre pouvant s'exprimer analytiquement par la relation :

$$H_m^{NFC(p/c, R/c)}(\omega) = \frac{F_m^{(p/c)}(\omega)}{F_m^{(R/c)}(\omega)} \quad [A11]$$

Le filtre total donné par la relation [A11] est stable et constitue la partie "encodage de distance" dans l'encodage ambisonique spatial selon l'invention, tel que représenté sur la figure 8. Les coefficients de ces filtres correspondent à des fonctions de transfert monotones de la fréquence, qui tendent vers la valeur 1 en hautes fréquences et vers la valeur

$(R/p)^m$ en basses fréquences. En se référant à la figure 9, les spectres d'énergie des filtres $H_m^{NFC(p/c, R/c)}(\omega)$

traduisent l'amplification des composantes encodées, dues à l'effet de champ de la source virtuelle (placée ici à une distance $p = 1$ m), avec une pré-compensation du champ des hauts-parleurs (placés à une distance $R = 1,5$ m). L'amplification en décibels est donc positive lorsque $p < R$ (cas de la figure 9) et négative quand $p > R$ (cas de la figure 10 où $p = 3$ m et $R = 1,5$ m). Dans un dispositif de restitution spatialisée, la distance R entre un point de perception auditive et les haut-parleurs HP_i est effectivement de l'ordre de un ou quelques mètres.

[0068] En se référant à nouveau à la figure 8, on comprendra que, outre les paramètres de direction θ et δ habituels, on transmettra une information sur les distances qui interviennent à l'encodage. Ainsi, les fonctions angulaires corres-

pondant aux harmoniques sphériques $Y_{mn}^{\sigma(\theta, \delta)}$ sont conservées pour l'encodage directionnel.

[0069] Toutefois, au sens de la présente invention, on prévoit en outre des filtres totaux (compensation de champ

proche et, le cas échéant, simulation d'un champ proche) $H_m^{NFC(p/c, R/c)}(\omega)$ qui sont appliqués aux compo-

santes ambisoniques, en fonction de leur ordre m, pour réaliser l'encodage de la distance, comme représenté sur la figure 8. Un mode de réalisation de ces filtres dans le domaine audionumérique sera décrit en détail plus loin.

[0070] On remarquera en particulier que ces filtres peuvent être appliqués dès même l'encodage de distance (r) et avant même l'encodage de direction (θ, δ). On comprendra ainsi que les étapes a) et b) ci-avant peuvent être rassemblées en une même étape globale, ou même être interverties (avec un encodage de distance et filtrage de compensation, suivis d'un encodage de direction). Le procédé selon l'invention ne se limite donc pas à une mise en oeuvre successive dans le temps des étapes a) et b).

[0071] La figure 11A représente une visualisation (vue de dessus) d'une reconstruction d'un champ proche avec compensation, d'une onde sphérique, dans le plan horizontal (avec les mêmes paramètres de distance que ceux de la figure 9), pour un système d'ordre total $M = 15$ et une restitution sur 32 hauts-parleurs. Sur la figure 11B, on a représenté la propagation de l'onde sonore initiale à partir d'une source en champ proche située à une distance p d'un point de l'espace d'acquisition qui correspond, dans l'espace de restitution, au point P de la figure 7 de perception auditive. On remarque sur la figure 11A que les auditeurs (symbolisés par des têtes schématisées) peuvent localiser la source virtuelle en un même lieu géographique situé à la distance p du point de perception P sur la figure 11B.

[0072] On vérifie bien ainsi que la forme du front d'onde encodé est respectée après décodage et restitution. Toutefois, on constate sensiblement des interférences à droite du point P tel que représenté sur la figure 11A qui sont dues au fait que le nombre de hauts-parleurs (donc de composantes ambisoniques prises en compte) n'est pas suffisant pour restituer parfaitement le front d'ondes en jeu sur toute la surface délimitée par les haut-parleurs.

[0073] Dans ce qui suit, on décrit, à titre d'exemple, l'obtention d'un filtre audionumérique pour la mise en oeuvre du procédé au sens de l'invention.

[0074] Comme indiqué ci-avant, si l'on cherche à simuler un effet de champ proche, compensé dès l'encodage, on applique aux composantes ambisoniques du son un filtre de la forme :

$$H_m^{NFC(\rho/c, R/c)}(\omega) = \frac{F_m^{(\rho/c)}(\omega)}{F_m^{(R/c)}(\omega)} \quad [\text{A11}]$$

[0075] De l'expression de la simulation d'un champ proche donné par la relation [A5], il apparaît que pour des sources lointaines ($p = \infty$), la relation [A11] devient simplement :

$$\frac{1}{F_m^{(R/c)}(\omega)} = H_m^{NFC(\infty, R/c)}(\omega) \quad [\text{A12}]$$

[0076] Il apparaît donc de cette dernière relation [A12] que le cas où la source à simuler émet en champ lointain (source lointaine) n'est qu'un cas particulier de l'expression générale du filtre formulée dans la relation [A11].

[0077] Dans le domaine des traitements audionumériques, un procédé avantageux pour définir un filtre numérique à partir de l'expression analytique de ce filtre dans le domaine analogique à temps continu consiste en une "transformée bilinéaire".

[0078] On exprime d'abord la relation [A5] sous la forme d'une transformée de Laplace, ce qui correspond à :

$$F_m^{(\tau)}(p) = \sum_{n=0}^m \frac{(m+n)!}{(m-n)!n!} (2\tau p)^{-n} \quad [\text{A13}]$$

où $\tau = \rho/c$ (c étant la vitesse acoustique dans le milieu, typiquement 340 m/s dans l'air).

[0079] La transformée bilinéaire consiste à présenter, pour une fréquence d'échantillonnage f_s , la relation [A11] sous la forme :

$$H_m(z) = \prod_{q=1}^{m/2} \frac{b_0^q + b_1^q z^{-1} + b_2^q z^{-2}}{a_0^q + a_1^q z^{-1} + a_2^q z^{-2}} \times \frac{b_0^{(m+1)/2} + b_1^{(m+1)/2} z^{-1}}{a_0^{(m+1)/2} + a_1^{(m+1)/2} z^{-1}} \quad [\text{A14}]$$

si m est impair et

$$H_m(z) = \prod_{q=1}^{m/2} \frac{b_0^q + b_1^q z^{-1} + b_2^q z^{-2}}{a_0^q + a_1^q z^{-1} + a_2^q z^{-2}}$$

si m est pair,

où z est défini par $p = 2f_s \frac{1-z^{-1}}{1+z^{-1}}$ par rapport à la relation [A13] précédente,

et avec :

$$x_0 = 1 - 2 \frac{\text{Re}(X_{m,q})}{\alpha} + \frac{|X_{m,q}|^2}{\alpha^2}, \quad x_1 = -2 \left(1 - \frac{|X_{m,q}|^2}{\alpha^2} \right)$$

et

$$x_2 = 1 + 2 \frac{\operatorname{Re}(X_{m,q})}{\alpha} + \frac{|X_{m,q}|^2}{\alpha^2}$$

$$x_0^{(m+1)/2} = 1 - \frac{X_{m,q}}{\alpha} \quad \text{et} \quad x_1^{(m+1)/2} = -\left(1 + \frac{X_{m,q}}{\alpha}\right)$$

où $\alpha = 4f_s R/c$ pour $x=a$ et $\alpha = 4f_s \rho/c$ pour $x=b$ **[0080]** $X_{m,q}$ sont les q racines successives du polynôme de Bessel :

$$F_m(x) = \sum_{n=0}^m \frac{(m+n)!}{(m-n)!n!} X^{m-n}$$

$$= \prod_{q=1}^m (X - X_{m,q})$$

et sont exprimées dans le tableau 1 ci-après, pour différents ordres m , sous les formes respectives de leur partie réelle, leur module (séparés par une virgule) et leur valeur (réelle) lorsque m est impair.

Tableau 1 : valeurs $R_e[X_{m,q}]$, $|X_{m,q}|$ (et $R_e[X_{m,m}]$ lorsque m est impair) d'un polynôme de Bessel calculées à l'aide du logiciel de calcul MATLAB®.

m=1	-2.0000000000
m=2	-3.0000000000, 3.4641016151
m=3	-3.6778146454, 5.0830828022 ; -4.6443707093
m=4	-4.2075787944, 6.7787315854 ; -5.7924212056, 6.0465298776
m=5	-4.6493486064, 8.5220456027 ; -6.7039127983, 7.5557873219 ; -7.2934771907
m=6	-5.0318644956, 10.2983543043 -7.4714167127, 9.1329783045 -8.4967187917, 8.6720541026
m=7	-5.3713537579, 12.0990553610; -8.1402783273, 10.7585400670 ; -9.5165810563, 10.1324122997 ; -9.9435737171
m=8	-5.6779678978, 13.9186233016 ; -8.7365784344, 12.4208298072 ; -10.4096815813, 11.6507064310 ; -11.1757720865, 11.3096817388
m=9	-5.9585215964, 15.7532774523 ; -9.2768797744, 14.1121936859 ; -11.2088436390, 13.2131216226 ; -12.2587358086, 12.7419414392 ; -12.5940383634

EP 1 563 485 B1

Suite de tableau

5	m=10	-6.2178324673, 17.6003068759 ; -9.7724391337, 15.8272658299 ; -11.9350566572, 14.8106929213 ; -13.2305819310, 14.2242555605 ; -13.8440898109, 13.9524261065
10	m=11	-6.4594441798, 19.4576958063 ; -10.2312965678, 17.5621095176 ; -12.6026749098, 16.4371594915 ; -14.1157847751, 15.7463731900 ; -14.9684597220, 15.3663558234 ; -15.2446796908
15	m=12	-6.6860466156, 21.3239012076 ; -10.6594171817, 19.3137363168 ; -13.2220085001, 18.0879209819 ; -14.9311424804, 17.3012295772 ; -15.9945411996, 16.8242165032 ; -16.5068440226, 16.5978151615
20	m=13	-6.8997344413, 23.1977134580 ; -11.0613619668, 21.0798161546 ; -13.8007456514, 19.7594692366 ; -15.6887605582, 18.8836767359 ; -16.9411835315, 18.3181073534 ; -17.6605041890, 17.9988179873 ; -17.8954193236
25	m=14	-7.1021737668, 25.0781652657 ; -11.4407047669, 22.8584924996 ; -14.3447919297, 21.4490520815 ; -16.3976939224, 20.4898067617 ; -17.8220011429, 19.8423306934 ; -18.7262916698, 19.4389130000 ; -19.1663428016, 19.2447495545
30	m=15	-7.2947137247, 26.9644699653 ; -11.8003034312, 24.6482592959 ; -14.8587939669, 23.1544615283 ; -17.0649181370, 22.1165594535 ; -18.6471986915, 21.3925954403 ; -19.7191341042, 20.9118275261 ; -20.3418287818, 20.6361378957 ; -20.5462183256
35	m=16	-7.4784635949, 28.8559784487 ; -12.1424827551, 26.4478760957 ; -15.3464816324, 24.8738935490 ; -17.6959363478, 23.7614799683 ; -19.4246523327, 22.9655586516 ; -20.6502404436, 22.4128776078 ; -21.4379698156, 22.0627133056 ; -21.8237730778, 21.8926662470
40	m=17	-7.6543475694, 30.7521483222 ; -12.4691619784, 28.2563077987 ; -15.8108990691, 26.6058519104 ; -18.2951775164, 25.4225585034 ; -20.1605894729, 24.5585534450 ; -21.5282660840, 23.9384287933 ; -22.4668764601, 23.5193877036 ; -23.0161527444, 23.2766166711 ; -23.1970582109
45	m=18	-7.8231445835, 32.6525213363 ; -12.7819455282, 30.0726807554 ; -16.2545681590, 28.3490792784 ; -18.8662638563, 27.0981271991 ; -20.8600257104, 26.1693913642 ; -22.3600808236, 25.4856138632 ; -23.4378933084, 25.0022244227 ; -24.1362741870, 24.6925542646 ; -24.4798038436, 24.5412441597
50		

55

Suite de tableau

m=19	-7.9855178345, 34.5567065132 ; -13.0821901901, 31.8962504142 ; -16.6796008200, 30.1025072510 ; -19.4122071436, 28.7867778706 ; -21.5270719955, 27.7962699865 ; -23.1512112785, 27.0520753105 ; -24.3584393996, 26.5081174988 ; -25.1941793616, 26.1363057951 ; -25.6855663388, 25.9191817486 ; -25.8480312755
------	---

[0081] On implémente ainsi les filtres numériques, à partir des valeurs du tableau 1, en prévoyant des cascades de cellules d'ordre 2 (pour m pair), et une cellule supplémentaire (pour m impair), à partir des relations [A14] données ci-avant.

[0082] On réalise ainsi des filtres numériques sous une forme de réponse impulsionnelle infinie, aisément paramétrable comme montré ci-avant. Il est à noter qu'une implémentation sous une forme de réponse impulsionnelle finie peut être envisagée et consiste à calculer le spectre complexe de la fonction de transfert à partir de la formule analytique, puis à en déduire une réponse impulsionnelle finie par transformée de Fourier inverse. On applique ensuite une opération de convolution pour le filtrage.

[0083] Ainsi, en introduisant cette pré-compensation du champ proche à l'encodage, on définit une représentation ambisonique modifiée (figure 8), en adoptant comme représentation transmissible des signaux exprimés dans le domaine fréquentiel, sous la forme :

$$\tilde{B}_{mn}^{\sigma(R/c)} = \frac{1}{F_m^{R/c}(\omega)} B_{mn}^{\sigma} \quad [\text{A15}]$$

[0084] Comme indiqué ci-avant, R est une distance de référence à laquelle est associé un effet de champ proche compensé et c est la vitesse du son (typiquement 340 m/s dans l'air). Cette représentation ambisonique modifiée possède les mêmes propriétés de scalabilité (schématiquement représentée par des données transmises "entourées" près de la flèche TR de la figure 1) et obéit aux mêmes transformations de rotation du champ (module 4 de la figure 1) que la représentation ambisonique habituelle.

[0085] On indique ci-après les opérations à mettre en oeuvre pour le décodage des signaux ambisoniques reçus.

[0086] On indique tout d'abord que l'opération de décodage est adaptable à un dispositif de restitution quelconque, de rayon R_2 , différent de la distance de référence R ci-avant. A cet effet, on applique des filtres de ty-

pe $H_m^{NFC(\rho/c, R/c)}(\omega)$, tels que décrits plus haut, mais avec des paramètres de distance R et R_2 , au lieu de p et R. En particulier, il est à noter que seul le paramètre R/c est à mémoriser (et/ou transmettre) entre l'encodage et le décodage.

[0087] En se référant à la figure 12, le module de filtrage qui y est représenté est prévu par exemple dans une unité de traitement d'un dispositif de restitution. Les composantes ambisoniques reçues ont été pré-compensées à l'encodage pour une distance de référence R_1 en tant que seconde distance. Toutefois, le dispositif de restitution comporte une pluralité de hauts-parleurs disposés à une troisième distance R_2 d'un point de perception auditive P, cette troisième distance R_2 étant différente de la seconde distance précitée R_1 . Le module de filtrage de la figure 12, sous la for-

me $H_m^{NFC(R_1/c, R_2/c)}(\omega)$, adapte alors, à la réception des données, la pré-compensation à la distance R_1 pour une restitution à la distance R_2 . Bien entendu, comme indiqué ci-avant, le dispositif de restitution reçoit aussi le paramètre R_1/c .

[0088] Il est à noter que l'invention permet en outre de mixer plusieurs représentations ambisoniques de champs sonores (sources réelles et/ou virtuelles), dont les distances de référence R sont différentes (le cas échéant avec des distances de référence infinies et correspondant à des sources lointaines). Préférentiellement, on filtrera une pré-compensation de toutes ces sources à une distance de référence la plus petite, avant de mélanger les signaux ambisoniques, ce qui permet à la restitution d'obtenir une définition correcte du relief sonore.

[0089] Dans le cadre d'un traitement dit de "focalisation sonore" avec, à la restitution, un effet d'enrichissement sonore pour une direction choisie de l'espace (à la manière d'un projecteur lumineux éclairant dans une direction choisie en optique), impliquant un traitement matriciel de focalisation sonore (avec pondération des composantes ambisoniques),

on applique avantageusement l'encodage de distance avec pré-compensation de champ proche de façon combinée au traitement de focalisation.

[0090] Dans ce qui suit, on décrit un procédé de décodage ambisonique, avec compensation du champ proche des hauts-parleurs, à la restitution.

[0091] Pour reconstruire un champ acoustique encodé suivant le formalisme ambisonique, à partir des composan-

tes B_{mn}^{σ} et en utilisant des hauts-parleurs d'un dispositif de restitution qui prévoit un emplacement "idéal" d'un auditeur qui correspond au point de restitution P de la figure 7, l'onde émise par chaque haut-parleur est définie par un traitement préalable de "ré-encodage" du champ ambisonique au centre du dispositif de restitution, comme suit.

[0092] Dans ce contexte de "ré-encodage", on considère dans un premier temps et pour simplification que les sources émettent en champ lointain.

[0093] En se référant à nouveau à la figure 7, l'onde émise par un haut-parleur d'indice i et d'incidence (θ_i et δ_i) est

alimenté par un signal S_i . Ce haut-parleur participe à la reconstruction de la composante B_{mn}' , par sa contribu-

tion $S_i \cdot Y_{mn}^{\sigma}(\theta_i, \delta_i)$.

[0094] Le vecteur c_i des coefficients d'encodage associés aux hauts-parleurs d'indice i s'exprime par la relation :

$$c_i = \begin{bmatrix} Y_{00}^{+1}(\theta_i, \delta_i) \\ Y_{11}^{+1}(\theta_i, \delta_i) \\ Y_{11}^{-1}(\theta_i, \delta_i) \\ \dots \\ Y_{mn}^{\delta}(\theta_i, \delta_i) \\ \dots \end{bmatrix} \quad [B1]$$

[0095] Le vecteur S des signaux émanant de l'ensemble des N hauts-parleurs est donné par l'expression :

$$S = \begin{bmatrix} S_1 \\ S_2 \\ \dots \\ S_N \end{bmatrix} \quad [B2]$$

[0096] La matrice d'encodage de ces N hauts-parleurs (qui correspond finalement à une matrice de "ré-encodage"), s'exprime par la relation :

$$C = [c_1 \ c_2 \ \dots c_N] \quad [B3]$$

où chaque terme c_i représente un vecteur selon la relation [B1] ci-avant.

[0097] Ainsi, la reconstruction du champ ambisonique B' est définie par la relation :

$$\tilde{B} = \begin{bmatrix} B_{00}^{+1} \\ B_{11}^{+1} \\ B_{11}^{-1} \\ \dots \\ B_{mn}^{\sigma} \\ \dots \end{bmatrix} = C.S \quad [B4]$$

[0098] La relation [B4] définit ainsi une opération de ré-encodage, préalable à la restitution. Finalement, le décodage, en tant que tel, consiste à comparer les signaux ambisoniques originaux et reçus par le dispositif de restitution, sous la forme :

$$B = \begin{bmatrix} B_{00}^{+1} \\ B_{11}^{+1} \\ B_{11}^{-1} \\ \dots \\ B_{mn}^{\sigma} \\ \dots \end{bmatrix} \quad [B5]$$

aux signaux ré-encodés $\tilde{B}$, pour définir la relation générale :

$$B' = B \quad [B6]$$

[0099] Il s'agit, en particulier, de déterminer les coefficients d'une matrice de décodage D, qui vérifie la relation :

$$S = D.B \quad [B7]$$

[0100] De préférence, le nombre de hauts-parleurs est supérieur ou égal au nombre de composantes ambisoniques à décoder et la matrice de décodage D s'exprime, en fonction de la matrice de ré-encodage C, sous la forme :

$$D = C^T . (C.C^T)^{-1} \quad [B8]$$

où la notation C^T correspond à la transposée de la matrice C.

[0101] Il est à noter que la définition d'un décodage vérifiant des critères différents par bandes de fréquences est possible, ce qui permet d'offrir une restitution optimisée en fonction des conditions d'écoute, notamment pour ce qui concerne la contrainte de positionnement au centre O de la sphère de la figure 3, lors de la restitution. A cet effet, on prévoit avantageusement un filtrage simple, en égalisation fréquentielle par paliers, à chaque composante ambisonique.

[0102] Toutefois, pour obtenir une reconstruction d'une onde originellement encodée, il faut corriger l'hypothèse de champ lointain pour les hauts-parleurs, c'est-à-dire exprimer l'effet de leur champ proche dans la matrice de ré-encodage C ci-avant et inverser ce nouveau système pour définir le décodeur. A cet effet, en supposant une concentricité des hauts-parleurs (disposés à une même distance R du point P de la figure 7), tous les hauts-parleurs ont un même effet

de champ proche $F_m^{(R/c)}(\omega)$, sur chaque composante ambisonique du type B_{mn}^σ . En introduisant les termes de champ proche sous la forme d'une matrice diagonale, la relation [B4] ci-avant devient :

$$B' = \text{Diag} \left(\left[1 \quad F_1^{R/c}(\omega) \quad F_1^{R/c}(\omega) \quad \dots F_m^{R/c}(\omega) \quad F_m^{R/c}(\omega) \quad \dots \right] \right) . C . S \quad [\text{B9}]$$

[0103] La relation [B7] ci-avant devient :

$$S = D . \text{Diag} \left(\left[1 \frac{1}{F_1^{R/c}(\omega)} \quad \frac{1}{F_1^{R/c}(\omega)} \quad \dots \frac{1}{F_m^{R/c}(\omega)} \quad \frac{1}{F_m^{R/c}(\omega)} \quad \dots \right] \right) . B \quad [\text{B10}]$$

[0104] Ainsi, l'opération de matricage est précédée par une opération de filtrage qui compense le champ proche sur chaque composante B_{mn}^σ , et qui peut être mise en oeuvre sous forme numérique, comme décrit ci-avant, en référence à la relation [A14].

[0105] On retiendra qu'en pratique, la matrice C de "ré-encodage" est propre au dispositif de restitution. Ses coefficients peuvent être déterminés initialement par paramétrage et caractérisation sonore du dispositif de restitution réagissant à une excitation prédéterminée. La matrice de décodage D est, elle aussi, propre au dispositif de restitution. Ses coefficients peuvent être déterminés par la relation [B8]. En reprenant la notation précédente où $\tilde{B}$ est la matrice des composantes ambisoniques pré-compensées, ces dernières peuvent être transmises au dispositif de restitution sous forme matricielle $\tilde{B}$ avec :

$$\tilde{B} = \text{Diag} \left(\left[1 \frac{1}{F_1^{R/c}(\omega)} \quad \frac{1}{F_1^{R/c}(\omega)} \quad \dots \frac{1}{F_m^{R/c}(\omega)} \quad \frac{1}{F_m^{R/c}(\omega)} \quad \dots \right] \right) . B$$

[0106] Le dispositif de restitution décode ensuite les données reçues sous forme matricielle $\tilde{B}$ (vecteur colonne des composantes transmises) en appliquant la matrice de décodage D aux composantes ambisoniques pré-compensées, pour former les signaux S_i destinés à alimenter les haut-parleurs HP_i , avec :

$$S = \begin{pmatrix} S_1 \\ S_i \\ S_N \end{pmatrix} = D . \tilde{B} \quad [\text{B11}]$$

[0107] En se référant à nouveau à la figure 12, si une opération de décodage doit être adaptée à un dispositif de restitution de rayon R_2 différent de la distance de référence R_1 , un module d'adaptation préalable au décodage propre-

ment dit et décrit ci-avant permet de filtrer chaque composante ambisonique $\tilde{B}_{mn}^\sigma$, pour l'adapter à un dispositif de restitution de rayon R_2 . L'opération de décodage proprement dite est effectuée ensuite, comme décrit ci-avant, en référence à la relation [B11].

[0108] On décrit ci-après une application de l'invention à la synthèse binaurale.

[0109] On se réfère à la figure 13A sur laquelle un auditeur disposant d'un casque à deux écouteurs d'un dispositif de synthèse binaurale est représenté. Les deux oreilles de l'auditeur sont disposées à des points respectifs O_L (oreille

gauche) et O_R (oreille droite) de l'espace. Le centre de la tête de l'auditeur est disposé au point O et le rayon de la tête de l'auditeur est de valeur a. Une source sonore doit être perçue auditivement à un point M de l'espace, situé à une distance r du centre de la tête de l'auditeur (et respectivement à des distances r_R de l'oreille droite et r_L de l'oreille

gauche). Par ailleurs, la direction de la source placée au point M est définie par les vecteurs $\vec{r}$, $\vec{r}_R$ et $\vec{r}_L$.

[0110] De façon générale, la synthèse binaurale se définit comme suit.

[0111] Chaque auditeur a une forme d'oreille qui lui est propre. La perception d'un son dans l'espace par cet auditeur se fait par apprentissage, depuis la naissance, en fonction de la forme des oreilles (notamment la forme des pavillons et les dimensions de la tête) propre à cet auditeur. La perception d'un son dans l'espace se manifeste entre autres par le fait que le son parvient à une oreille, avant l'autre oreille, ce qui se traduit par un retard τ entre les signaux à émettre par chaque écouteur du dispositif de restitution appliquant la synthèse binaurale.

[0112] Le dispositif de restitution est paramétré initialement, pour un même auditeur, en balayant une source sonore autour de sa tête, à une même distance R du centre de sa tête. On comprendra ainsi que cette distance R peut être considérée comme une distance entre un "point de restitution" comme énoncé ci-avant et un point de perception auditive (ici le centre O de la tête de l'auditeur).

[0113] Dans ce qui suit, l'indice L est associé au signal à restituer par l'écouteur accolé à l'oreille gauche et l'indice R est associé au signal à restituer par l'écouteur accolé à l'oreille droite. En se référant à la figure 13B, on applique au signal initial S un retard pour chaque voie destinée à produire un signal pour un écouteur distinct. Ces retards τ_L et τ_R sont fonction d'un retard maximum τ_{MAX} qui correspond ici au rapport a/c où a, comme indiqué précédemment, correspond au rayon de la tête de l'auditeur et c à la vitesse du son. En particulier, ces retards sont définis en fonction de la différence de distance du point O (centre de la tête) au point M (position de la source dont le son est à restituer, sur la figure 13A) et de chaque oreille à ce point M. Avantagusement, on applique en outre des gains respectifs g_L et g_R , à chaque voie, qui sont fonction d'un rapport des distances du point O au point M et de chaque oreille au point M. Des modules respectifs appliqués à chaque voie 2_L et 2_R encodent les signaux de chaque voie, dans une représentation ambisonique, avec pré-compensation de champ proche NFC (pour "Near Field Compensation") au sens de la présente invention. On comprendra ainsi que, par la mise en oeuvre du procédé au sens de la présente invention, on peut définir les signaux issus de la source M, non seulement par leur direction (angles azimutaux θ_L et θ_R et angles d'élévation δ_L et δ_R), mais aussi en fonction de la distance séparant chaque oreille r_L et r_R de la source M. Les signaux ainsi encodés sont transmis au dispositif de restitution comportant des modules de décodage ambisonique, pour chaque voie, 5_L et 5_R . Ainsi, on applique un encodage/décodage ambisonique, avec compensation de champ proche, pour chaque voie (écouteur gauche, écouteur droit) dans la restitution avec synthèse binaurale (ici de type "B-FORMAT"), sous forme dédoublée. La compensation de champ proche s'effectue, pour chaque voie, avec comme première distance p une distance r_L et r_R entre chaque oreille et la position M de la source sonore à restituer.

[0114] On décrit ci-après une application de la compensation au sens de l'invention, au contexte de l'acquisition sonore en représentation ambisonique.

[0115] On se réfère à la figure 14 sur laquelle un microphone 141 comporte une pluralité de capsules transductrices, capables de capter des pressions acoustiques et restituer des signaux électriques $S_1, \dots, S_N$. Les capsules CAP_i sont agencées sur une sphère de rayon r prédéterminé (ici, une sphère rigide, telle qu'une balle de ping-pong par exemple). Les capsules sont espacées d'un pas régulier sur la sphère. En pratique, on choisit le nombre N de capsules en fonction de l'ordre M désiré pour la représentation ambisonique.

[0116] On indique ci-après, dans le contexte d'un microphone comportant des capsules agencées sur une sphère rigide, comment compenser l'effet de champ proche, dès l'encodage dans le contexte ambisonique. On montrera ainsi que la pré-compensation du champ proche peut s'appliquer non seulement pour la simulation de source virtuelle, comme indiqué ci-avant, mais aussi à l'acquisition et, de façon plus générale, en combinant la pré-compensation de champ proche à tous types de traitements impliquant une représentation ambisonique.

[0117] En présence d'une sphère rigide (susceptible d'introduire une diffraction des ondes sonores reçues), la relation [A1] donnée ci-avant devient :

$$P_r(\vec{u}_i) = \sum_{m=0}^{\infty} \frac{j^{m-1}}{(kr)^2 h_m'(kr)} \sum_{\substack{0 \leq n \leq m \\ \sigma = \pm 1}} B_{mn}^{\sigma} Y_{mn}^{\sigma}(\vec{u}_i) \quad [C1]$$

[0118] Les dérivées des fonctions de Hankel sphériques h_m^- obéissent à la loi de récurrence :

$$(2m+1)h_m^-(x) = m h_{m-1}^-(x) - (m+1) h_{m+1}^-(x) \quad [C2]$$

[0119] On déduit les composantes ambisoniques B_{mn}^σ du champ initial à partir du champ de pression à la surface de la sphère, en mettant en oeuvre des opérations de projection et d'égalisation données par la relation :

$$B_{mn}^\sigma = EQ_m < p_r | Y_{mn}^\sigma >_{4\pi} \quad [C3]$$

[0120] Dans cette expression, EQ_m est un filtre égaliseur qui compense une pondération W_m qui est liée à la directivité des capsules et qui inclut en outre la diffraction par la sphère rigide.

[0121] L'expression de ce filtre EQ_m est donnée par la relation suivante :

$$EQ_m = \frac{1}{W_m} = (kr)^2 h_m^-(kr) j^{-m+1} \quad [C4]$$

[0122] Les coefficients de ce filtre d'égalisation ne sont pas stables et on obtient un gain infini en très basses fréquences. D'ailleurs, il convient de noter que les composantes harmoniques sphériques, elles-mêmes, ne sont pas d'amplitude finie lorsque le champ sonore n'est pas limité à une propagation d'ondes planes, c'est-à-dire issues de sources lointaines, comme on l'a vu précédemment.

[0123] Par ailleurs, si, plutôt que de prévoir des capsules encastrées dans une sphère solide, on prévoit des capsules de type cardioïdes, avec une directivité en champ lointain donnée par l'expression :

$$G(\theta) = \alpha + (1 - \alpha) \cos \theta \quad [C5]$$

[0124] En considérant ces capsules montées sur un support "transparent acoustiquement", le terme de pondération à compenser devient :

$$W_m = j^m (\alpha j m(kr) - j(1 - \alpha) j m'(kr)) \quad [C6]$$

[0125] Il apparaît encore que les coefficients d'un filtre d'égalisation correspondant à l'inverse analytique de cette pondération donnée par la relation [C6] sont divergents pour les très basses fréquences.

[0126] De façon générale, on indique que pour tout type de directivité de capteurs, le gain du filtre EQ_m pour compenser la pondération W_m liée à la directivité des capteurs est infini pour les basses fréquences sonores. En se référant à la figure 14, on applique avantageusement une pré-compensation de champ proche dans l'expression même du filtre d'égalisation EQ_m , donnée par la relation :

$$EQ_m^{NFC(R/c)}(\omega) = \frac{EQ_m(r, \omega)}{F_m^{(R/c)}(\omega)} \quad [C7]$$

[0127] Ainsi, les signaux S_1 à S_N sont récupérés du microphone 141. Le cas échéant, on applique une pré-égalisation de ces signaux par un module de traitement 142. Le module 143 permet d'exprimer ces signaux dans le contexte ambisonique, sous forme matricielle. Le module 144 applique le filtre de la relation [C7] aux composantes ambisoniques

exprimées en fonction du rayon r de la sphère du microphone 141. La compensation de champ proche s'effectue pour une distance de référence R en tant que seconde distance. Les signaux encodés et ainsi filtrés par le module 144 peuvent être transmis, le cas échéant, avec le paramètre représentatif de la distance de référence R/c .

[0128] Ainsi, il apparaît dans les différents modes de réalisation liés respectivement à la création d'une source virtuelle en champ proche, à l'acquisition de signaux sonores issues de sources réelles, ou même à la restitution (pour compenser un effet de champ proche des hauts-parleurs), que la compensation de champ proche au sens de la présente invention peut s'appliquer à tous types de traitements faisant intervenir une représentation ambisonique. Cette compensation de champ proche permet d'appliquer la représentation ambisonique à une multiplicité de contextes sonores où la direction d'une source et avantageusement sa distance doivent être prises en compte. De plus, la possibilité de la représentation de phénomènes sonores de tous types (champs proches ou lointains) dans le contexte ambisonique est assurée par cette pré-compensation, du fait de la limitation à des valeurs réelles finies des composantes ambisoniques.

[0129] Bien entendu, la présente invention ne se limite pas à la forme de réalisation décrite ci-avant à titre d'exemple ; elle s'étend à d'autres variantes.

[0130] Ainsi, on comprendra que la pré-compensation de champ proche peut être intégrée, à l'encodage, autant pour une source proche que pour une source lointaine. Dans ce dernier cas (source lointaine et réception d'ondes planes), la distance p exprimée ci-avant sera considérée comme infinie, sans modifier de façon substantielle l'expression des filtres H_m donnée ci-avant. Ainsi, le traitement utilisant des processeurs d'effet de salle qui fournissent en général des signaux décorrélés utilisables pour modéliser le champ diffus tardif (réverbération tardive) peut être combiné à une pré-compensation de champ proche. On peut considérer que ces signaux sont de même énergie et correspondent à une

part de champ diffus correspondant à la composante omnidirective $W = B_{00}^{+1}$ (figure 4). On peut alors construire

les diverses composantes harmoniques sphériques (avec un ordre M choisi) en appliquant une correction de gain pour chaque composante ambisonique et on applique une compensation de champ proche des hauts-parleurs (avec une distance de référence R séparant les haut-parleurs du point de perception auditive comme représenté sur la figure 7).

[0131] Bien entendu, le principe d'encodage au sens de la présente invention est généralisable à des modèles de rayonnement autres que des sources monopolaires (réelles ou virtuelles) et/ou des hauts-parleurs. En effet, toute forme de rayonnement (notamment une source étalée dans l'espace) peut être exprimée par intégration d'une distribution continue de sources élémentaires ponctuelles.

[0132] En outre, dans le contexte de la restitution, il est possible d'adapter la compensation de champ proche à tout contexte de restitution. A cet effet, il peut être prévu de calculer des fonctions de transfert (ré-encodage des composantes harmoniques sphériques de champ proche pour chaque haut-parleur, compte tenu d'une propagation réelle dans la salle où le son est restitué), ainsi qu'une inversion de ce ré-encodage pour redéfinir le décodage.

[0133] On a décrit ci-avant un procédé de décodage dans lequel on appliquait un système matriciel faisant intervenir les composantes ambisoniques. Dans une variante, il peut être prévu un traitement généralisé par transformées de Fourier rapides (circulaire ou sphérique) pour limiter les temps de calcul et les ressources informatiques (en terme de mémoire) nécessaires au traitement de décodage.

[0134] Comme indiqué ci-avant en référence aux figures 9 et 10, on constate que le choix d'une distance de référence R par rapport à la distance p de la source en champ proche introduit une différence de gain pour différentes valeurs de la fréquence sonore. On indique que le procédé d'encodage avec pré-compensation peut être couplé à une compression audionumérique permettant de quantifier et d'ajuster le gain pour chaque sous-bande fréquentielle.

[0135] Avantageusement, la présente invention s'applique à tous types de systèmes de spatialisation sonore, notamment pour des applications de type "réalité virtuelle" (navigation dans des scènes virtuelles dans l'espace tridimensionnel, conversations de type "chat" sonorisées sur le réseau Internet), à des sonifications d'interfaces, à des logiciels d'édition audio pour enregistrer, mixer et restituer de la musique, mais aussi à l'acquisition, à partir d'usage de microphones tridimensionnels, pour la prise de son musicale ou cinématographique, ou encore pour la transmission d'ambiance sonore sur Internet, par exemple pour des "Webcam" sonorisées.

Revendications

1. Procédé de traitement de données sonores, dans lequel :

a) on code des signaux représentatifs d'au moins un son se propageant dans l'espace tridimensionnel et issu d'une source située à une première distance (p) d'un point de référence (O), pour obtenir une représentation du son par des composantes (B_{mn}^σ) exprimées dans une base d'harmoniques sphériques, d'origine correspondant audit point de référence (O),

b) et on applique auxdites composantes (B_{mn}^σ) une compensation d'un effet de champ proche par un filtrage

qui est fonction d'une seconde distance (R) définissant sensiblement, pour une restitution du son par un dispositif de restitution, une distance entre un point de restitution (HP_i) et un point (P) de perception auditive.

2. Procédé selon la revendication 1, dans lequel, ladite source étant lointaine du point de référence (O),

- on obtient des composantes d'ordres successifs m pour la représentation du son dans ladite base d'harmomoniques sphériques, et
 - on applique un filtre ($1/F_m$) dont les coefficients, appliqués chacun à une composante d'ordre m, s'expriment analytiquement sous la forme de l'inverse d'un polynôme de puissance m, dont la variable est inversement proportionnelle à la fréquence sonore et à ladite seconde distance (R), pour compenser un effet de champ proche au niveau du dispositif de restitution.

3. Procédé selon la revendication 1, dans lequel, ladite source étant une source virtuelle prévue à ladite première distance (ρ),

- on obtient des composantes d'ordres successifs m pour la représentation du son dans ladite base d'harmomoniques sphériques, et
 - on applique un filtre global (H_m) dont les coefficients, appliqués chacun à une composante d'ordre m, s'expriment analytiquement sous la forme d'une fraction, dont :

- le numérateur est un polynôme de puissance m, dont la variable est inversement proportionnelle à la fréquence sonore et à ladite première distance (ρ), pour simuler un effet de champ proche de la source virtuelle, et
 - le dénominateur est un polynôme de puissance m, dont la variable est inversement proportionnelle à la fréquence sonore et à ladite seconde distance (R), pour compenser l'effet du champ proche de la source virtuelle dans les basses fréquences sonores.

4. Procédé selon l'une des revendications précédentes, dans lequel on transmet au dispositif de restitution les données codées et filtrées aux étapes a) et b) avec un paramètre représentatif de ladite seconde distance (R/c).

5. Procédé selon l'une des revendications 1 à 3, dans lequel, le dispositif de restitution comportant des moyens de lecture d'un support mémoire, on mémorise sur un support mémoire destiné à être lu par le dispositif de restitution les données codées et filtrées aux étapes a) et b) avec un paramètre représentatif de ladite seconde distance (R/c).

6. Procédé selon l'une des revendications 4 et 5, dans lequel, préalablement à une restitution sonore par un dispositif de restitution comportant une pluralité de hauts-parleurs disposés à une troisième distance (R₂) dudit point de perception auditive (P), on applique aux données codées et filtrées un filtre d'adaptation ($H_m^{(R1/C, R2/C)}$) dont les coefficients sont fonction desdites seconde (R₁) et troisième distances (R₂).

7. Procédé selon la revendication 6, dans lequel les coefficients dudit filtre d'adaptation ($H_m^{(R1/C, R2/C)}$), appliqués chacun à une composante d'ordre m, s'expriment analytiquement sous la forme d'une fraction, dont :

- le numérateur est un polynôme de puissance m, dont la variable est inversement proportionnelle à la fréquence sonore et à ladite seconde distance (R),
 - et le dénominateur est un polynôme de puissance m, dont la variable est inversement proportionnelle à la fréquence sonore et à ladite troisième distance (R₂).

8. Procédé selon l'une des revendications 2, 3 et 7, dans lequel, pour la mise en oeuvre de l'étape b), on prévoit :

- pour des composantes d'ordre m pair, des filtres audionumériques sous la forme d'une cascade de cellules d'ordre deux ; et
 - pour des composantes d'ordre m impair, des filtres audionumériques sous la forme d'une cascade de cellules d'ordre deux et une cellule supplémentaire d'ordre un.

9. Procédé selon la revendication 8, dans lequel les coefficients d'un filtre audionumérique, pour une composante d'ordre m, sont définis à partir des valeurs numériques des racines desdits polynômes de puissance m.

10. Procédé selon l'une des revendications 2, 3, 7, 8 et 9, dans lequel lesdits polynômes sont des polynômes de Bessel.

11. Procédé selon l'une des revendications 1, 2 et 4 à 10, dans lequel on prévoit un microphone comportant un réseau de transducteurs acoustiques agencés sensiblement sur la surface d'une sphère dont le centre correspond sensiblement audit point de référence (O), pour obtenir lesdits signaux représentatifs d'au moins un son se propageant dans l'espace tridimensionnel.

12. Procédé selon la revendication 11, dans lequel on applique à l'étape b) un filtre global pour, d'une part, compenser un effet de champ proche en fonction de ladite seconde distance (R) et, d'autre part, égaliser les signaux issus des transducteurs pour compenser une pondération de directivité desdits transducteurs.

13. Procédé selon l'une des revendications 11 et 12, dans lequel on prévoit un nombre de transducteurs fonction d'un nombre total choisi de composantes pour représenter le son dans ladite base d'harmoniques sphériques.

14. Procédé selon l'une des revendications précédentes, dans lequel on choisit à l'étape a) un nombre total de composantes dans la base des harmoniques sphériques pour obtenir, à la restitution, une région de l'espace autour du point de perception (P) dans laquelle la restitution du son est fidèle et dont les dimensions sont croissantes avec le nombre total de composantes.

15. Procédé selon la revendication 14, dans lequel on prévoit un dispositif de restitution comportant un nombre de haut-parleurs au moins égal audit nombre total de composantes.

16. Procédé selon l'une des revendications 1 à 5 et 8 à 13, dans lequel :

- on prévoit un dispositif de restitution comportant au moins un premier et un second haut-parleur disposés à une distance choisie d'un auditeur,

- on obtient, pour cet auditeur, une information de ressenti de la position dans l'espace de sources sonores situées à une distance de référence prédéterminée (R) de l'auditeur, et

- on applique la compensation de l'étape b) avec ladite distance de référence sensiblement en tant que seconde distance.

17. Procédé selon l'une des revendications 1 à 3 et 8 à 13, prises en combinaison avec l'une des revendications 4 et 5, dans lequel :

- on prévoit un dispositif de restitution comportant au moins un premier et un second haut-parleur disposés à une distance choisie d'un auditeur,

- on obtient, pour cet auditeur, une information de ressenti de la position dans l'espace de sources sonores situées à une distance de référence prédéterminée (R_2) de l'auditeur, et

- préalablement à une restitution sonore par le dispositif de restitution, on applique aux données codées et filtrées aux étapes a) et b) un filtre d'adaptation ($H_m^{(R/C, R_2/C)}$) dont les coefficients sont fonction de la seconde distance (R) et sensiblement de la distance de référence (R_2).

18. Procédé selon l'une des revendications 16 et 17, dans lequel :

- le dispositif de restitution comporte un casque à deux écouteurs pour les oreilles respectives de l'auditeur, et

- séparément pour chaque écouteur, on applique le codage et le filtrage des étapes a) et b) pour des signaux respectifs destinés à alimenter chaque écouteur, avec, en tant que première distance (p), respectivement une distance (r_R, r_L) séparant chaque oreille d'une position (M) d'une source à restituer.

19. Procédé selon l'une des revendications précédentes, dans lequel on met en forme, aux étapes a) et b), un système matriciel comportant au moins :

- une matrice (B) comportant lesdites composantes dans la base des harmoniques sphériques, et
- une matrice diagonale ($\text{Diag}(1/F_m)$) dont les coefficients correspondent à des coefficients de filtrage de l'étape b),

et on multiplie lesdites matrices pour obtenir une matrice résultat de composantes compensées ($\tilde{B}$).

20. Procédé selon la revendication 19, dans lequel :

- le dispositif de restitution comporte une pluralité de haut-parleurs disposés sensiblement à une même distance (R) du point de perception auditive (P), et
- pour décoder lesdites données codées et filtrées aux étapes a) et b) et former des signaux adaptés pour alimenter lesdits haut-parleurs :

- * on forme un système matriciel comportant ladite matrice résultat ($\tilde{B}$) et une matrice de décodage (D) prédéterminée, propre au dispositif de restitution, et
- * on obtient une matrice (S) comportant des coefficients représentatifs des signaux d'alimentation des hauts-parleurs par multiplication de la matrice des composantes compensées ($\tilde{B}$) par ladite matrice de décodage (D).

21. Dispositif d'acquisition sonore, comportant un microphone muni d'un réseau de transducteurs acoustiques disposés sensiblement sur la surface d'une sphère, **caractérisé en ce qu'il** comporte en outre une unité de traitement agencée pour :

- recevoir des signaux émanant chacun d'un transducteur,
- appliquer auxdits signaux un codage pour obtenir une représentation du son par des composantes (B_{mn}^σ) exprimées dans une base d'harmoniques sphériques, d'origine correspondant au centre de ladite sphère (O),
- et appliquer auxdites composantes (B_{mn}^σ) un filtrage qui est fonction, d'une part, d'une distance correspondant au rayon de la sphère (r) et, d'autre part, d'une distance de référence (R).

22. Dispositif selon la revendication 21, **caractérisé en ce que** ledit filtrage consiste, d'une part, à égaliser, en fonction du rayon de la sphère, les signaux issus des transducteurs pour compenser une pondération de directivité desdits transducteurs et, d'autre part, à compenser un effet de champ proche en fonction d'une distance de référence choisie (R), définissant sensiblement, pour une restitution du son, une distance entre un point de restitution (HP_i) et un point (P) de perception auditive.

Patentansprüche

1. Verfahren zur Verarbeitung von Audiodateien, bei dem:

- a) man Signale kodiert, die mindestens einen Ton darstellen, der sich im dreidimensionalen Raum ausbreitet und von einer Quelle kommt, die in einem ersten Abstand (p) von einem Bezugspunkt (O) gelegen ist, um eine Darstellung des Tons durch in einer Basis von sphärischen Harmonischen ausgedrückte Komponenten (B_{mn}^σ) mit einem diesem Bezugspunkt (O) entsprechenden Ursprung zu erhalten,
- b) und man an diese Komponenten (B_{mn}^σ) eine Kompensierung eines Nahfeldeffekts durch eine Filterung anlegt, die eine Funktion von einem zweiten Abstand (R) ist, der bei einer Wiedergabe des Tons durch eine Wiedergabevorrichtung im Wesentlichen einen Abstand zwischen einem Wiedergabepunkt (HP_i) und einem Hörwahrnehmungspunkt (P) definiert.

2. Verfahren nach Anspruch 1, bei dem, wenn diese Quelle vom Bezugspunkt (O) entfernt ist,

- man Komponenten von aufeinanderfolgenden Ordnungen m bei der Darstellung des Tons in dieser Basis von sphärischen Harmonischen erhält und
- man ein Filter ($1/F_m$) anlegt, dessen jeweils an eine Komponente der Ordnung m angelegte Koeffizienten sich analytisch in der Form des Kehrwerts eines Polynoms der Potenz m ausdrückt, dessen Variable umgekehrt proportional zur Tonfrequenz und zum zweiten Abstand (R) ist, um einen Nahfeldeffekt auf Höhe der Wiedergabevorrichtung zu kompensieren.

3. Verfahren nach Anspruch 1, bei dem, wenn diese Quelle eine in diesem ersten Abstand (p) vorgesehene virtuelle Quelle ist,

- man Komponenten von aufeinanderfolgenden Ordnungen m bei der Darstellung des Tons in dieser Basis von sphärischen Harmonischen erhält und
- man ein Globalfilter (H_m) anlegt, dessen jeweils an eine Komponente der Ordnung m angelegte Koeffizienten sich analytisch in der Form eines Bruchs ausdrücken, dessen

- Zähler ein Polynom der Potenz m ist, dessen Variable umgekehrt proportional zur Tonfrequenz und zum ersten Abstand (p) ist, um einen Nahfeldeffekt der virtuellen Quelle zu simulieren, und
- dessen Nenner ein Polynom der Potenz m ist, dessen Variable umgekehrt proportional zu der Tonfrequenz und zu dem zweiten Abstand (R) ist, um den Nahfeldeffekt der virtuellen Quelle in den niedrigen Tonfrequenzen zu kompensieren.

4. Verfahren nach einem der vorhergehenden Ansprüche, bei dem man die in den Schritten a) und b) kodierten und gefilterten Daten mit einem diesen zweiten Abstand darstellenden Parameter (R/c) zur Wiedergabevorrichtung überträgt.

5. Verfahren nach einem der Ansprüche 1 bis 3, bei dem man, wenn die Wiedergabevorrichtung Mittel zum Lesen eines Speicherträgers umfasst, auf einem Speicherträger, der dazu bestimmt ist, von der Wiedergabevorrichtung gelesen zu werden, die in den Schritten a) und b) kodierten und gefilterten Daten mit einem diesen zweiten Abstand darstellenden Parameter (R/c) speichert.

6. Verfahren nach einem der Ansprüche 4 und 5, bei dem man vor einer Tonwiedergabe durch eine Wiedergabevorrichtung, die eine Vielzahl von Lautsprechern umfasst, die in einem dritten Abstand (R_2) von diesem Hörwahrnehmungspunkt (P) angeordnet sind, an die kodierten und gefilterten Daten ein Anpassungsfilter ($H_m^{(R1/c, R2/c)}$) anlegt, dessen Koeffizienten eine Funktion von dem zweiten Abstand (R_1) und dem dritten Abstand (R_2) sind.

7. Verfahren nach Anspruch 6, bei dem die jeweils an eine Komponente der Ordnung m angelegten Koeffizienten des Anpassungsfilters ($H_m^{(R1/c, R2/c)}$) sich analytisch in der Form eines Bruchs ausdrücken, dessen

- Zähler ein Polynom der Potenz m ist, dessen Variable umgekehrt proportional zur Tonfrequenz und zum zweiten Abstand (R) ist, und
- dessen Nenner ein Polynom der Potenz m ist, dessen Variable umgekehrt proportional zu der Tonfrequenz und zu dem dritten Abstand (R_2) ist.

8. Verfahren nach einem der Ansprüche 2, 3 und 7, bei dem man für die Durchführung des Schritts b)

- für Komponenten geradzahlgiger Ordnung m audionumerische Filter in der Form einer Kaskade von Zellen der Ordnung zwei und
- für Komponenten ungeradzahlgiger Ordnung m audionumerische Filter in der Form einer Kaskade von Zellen der Ordnung zwei und eine zusätzliche Zelle der Ordnung eins vorsieht.

9. Verfahren nach Anspruch 8, bei dem die Koeffizienten eines audionumerischen Filters bei einer Komponente der Ordnung m ausgehend von den numerischen Werten der Wurzeln dieser Polynome der Potenz m definiert sind.

10. Verfahren nach einem der Ansprüche 2, 3, 7, 8 und 9, bei dem diese Polynome Bessel-Polynome sind.

11. Verfahren nach einem der Ansprüche 1, 2 und 4 bis 10, bei dem man ein Mikrophon vorsieht, das ein Netz von akustischen Wandlern aufweist, die im Wesentlichen auf der Oberfläche einer Kugel angeordnet sind, deren Mittelpunkt im Wesentlichen dem Bezugspunkt (O) entspricht, um die Signale zu erhalten, die mindestens einen sich im dreidimensionalen Raum ausbreitenden Ton darstellen.

12. Verfahren nach Anspruch 11, bei dem man im Schritt b) ein Globalfilter anlegt, um einerseits einen Nahfeldeffekt in Abhängigkeit von diesem zweiten Abstand (R) zu kompensieren und andererseits die von den Wandlern kommenden Signale zu equalisieren, um eine Richtwirkungsgewichtung der Wandler zu kompensieren.

13. Verfahren nach einem der Ansprüche 11 und 12, bei dem man eine Anzahl von Wandlern vorsieht, die von einer gewählten Gesamtzahl von Komponenten abhängt, um den Ton in der Basis der sphärischen Harmonischen darzustellen.

14. Verfahren nach einem der vorhergehenden Ansprüche, bei dem man im Schritt a) eine Gesamtzahl von Komponenten in der Basis der sphärischen Harmonischen wählt, um bei der Wiedergabe einen Bereich des Raums um den Wahrnehmungspunkt (P) herum zu erhalten, in dem die Wiedergabe des Tons getreu ist und dessen Abmessungen mit der Gesamtzahl von Komponenten zunehmen.

15. Verfahren nach Anspruch 14, bei dem man eine Wiedergabevorrichtung vorsieht, die eine Anzahl von Lautsprechern von mindestens gleich der Gesamtzahl der Komponenten vorsieht.

16. Verfahren nach einem der Ansprüche 1 bis 5 und 8 bis 13, bei dem

- man eine Wiedergabevorrichtung vorsieht, die mindestens einen ersten und einen zweiten Lautsprecher vorsieht, die in einem gewählten Abstand von einem Hörer angeordnet sind,
- für diesen Hörer eine Information der Empfindung der Lage von Tonquellen, die in einem vorbestimmten Bezugsabstand (R) vom Hörer gelegen sind, im Raum erhält und
- die Kompensierung des Schritts b) mit dem Bezugsabstand im Wesentlichen als zweiten Abstand anlegt.

17. Verfahren nach einem der Ansprüche 1 bis 3 und 8 bis 13 in Kombination mit einem der Ansprüche 4 und 5, bei dem

- man eine Wiedergabevorrichtung vorsieht, die mindestens einen ersten und einen zweiten Lautsprecher umfasst, die in einem gewählten Abstand von einem Hörer angeordnet sind,
- man für diesen Hörer eine Information der Empfindung der Lage von Tonquellen, die in einem vorbestimmten Bezugsabstand (R₂) vom Hörer gelegen sind, im Raum erhält und
- man vor einer Tonwiedergabe durch die Wiedergabevorrichtung an die in den Schritten a) und b) kodierten und gefilterten Daten ein Anpassungsfilter ($H_m^{(R/c, R_2/c)}$) anlegt, dessen Koeffizienten eine Funktion des zweiten Abstands (R) und im Wesentlichen des Bezugsabstands (R₂) sind.

18. Verfahren nach einem der Ansprüche 16 und 17, bei dem

- die Wiedergabevorrichtung einen Kopfhörer mit zwei Hörern für die Ohren des Hörers umfasst und
- man getrennt für jeden Hörer die Kodierung und Filterung der Schritte a) und b) für Signale, die jeweils zur Versorgung jedes Hörers bestimmt sind, anlegt, mit als erstem Abstand (p) jeweils einen Abstand (r_R, r_L) der jedes Ohr von einem Standort (M) einer wiederzugebenden Quelle trennt.

19. Verfahren nach einem der vorhergehenden Ansprüche, bei dem man in den Schritten a) und b) ein Matrixsystem in Form bringt, das mindestens umfasst:

- eine Matrix (B), die die Komponenten in der Basis der sphärischen Harmonischen umfasst, und
- eine diagonale Matrix ($\text{Diag}(1/F_m)$), deren Koeffizienten Filterkoeffizienten des Schritts b) entsprechen, und man die Matrizen multipliziert, um eine resultierende Matrix von kompensierten Komponenten ($\tilde{B}$) zu erhalten.

20. Verfahren nach Anspruch 19, bei dem

- die Wiedergabevorrichtung eine Vielzahl von Lautsprechern umfasst, die in einem gleichen Abstand (R) vom dem Hörwahrnehmungspunkt (P) angeordnet sind, und,
- um diese in den Schritten a) und b) kodierten und gefilterten Daten zu dekodieren und Signale zu formen, die dafür ausgelegt sind, die Lautsprecher zu speisen,

* man ein Matrixsystem bildet, das die resultierende Matrix ($\tilde{B}$) und eine vorbestimmte, der Wiedergabevorrichtung zugeordnete Dekodierungsmatrix (D) umfasst, und

* man eine Matrix (S), die die Signale zur Versorgung der Lautsprecher darstellende Koeffizienten aufweist, durch Multiplikation der Matrix ($\tilde{B}$) der kompensierten Komponenten mit der Dekodierungsmatrix (D) erhält.

21. Tonerfassungsvorrichtung, umfassend ein Mikrophon, das mit einem Netz von akustischen Wandlern versehen ist, die im Wesentlichen auf der Oberfläche einer Kugel angeordnet sind, **dadurch gekennzeichnet, dass** sie außerdem eine Verarbeitungseinheit aufweist, die dafür ausgelegt ist,

- jeweils von einem Wandler ausgehende Signale zu empfangen,
- an diese Signale eine Kodierung anzulegen, um eine Darstellung des Tons durch Komponenten (B_{mn}^σ), die in einer Basis von sphärischen Harmonischen ausgedrückt sind, mit einem dem Mittelpunkt dieser Kugel (O) entsprechenden Ursprung zu erhalten,
- und an diese Komponenten (B_{mn}^σ) eine Filterung anzulegen, die eine Funktion einerseits von einem dem Radius der Kugel (r) entsprechenden Abstand und andererseits von einem Bezugsabstand (R) ist.

22. Vorrichtung nach Anspruch 21, **dadurch gekennzeichnet, dass** die Filterung darin besteht, dass einerseits in Abhängigkeit von dem Radius der Kugel die von den Wandlern kommenden Signale egalisiert werden, um eine Richtwirkungsgewichtung der Wandler zu kompensieren, und andererseits ein Nahfeldeffekt in Abhängigkeit von einem gewählten Bezugsabstand (R) kompensiert wird, der bei einer Wiedergabe des Tons im Wesentlichen einen Abstand zwischen einem Wiedergabepunkt (HP_i) und einem Hörwahrnehmungspunkt (P) definiert.

Claims

1. A method of processing sound data, in which:

a) signals representative of at least one sound propagating in a three-dimensional space and arising from a source situated at a first distance (ρ) from a reference point (O) are coded so as to obtain a representation of the sound by components (B_{mn}^σ) expressed in a base of spherical harmonics, of origin corresponding to said reference point (O),

b) and a compensation of a near field effect is applied to said components (B_{mn}^σ) by a filtering which is dependent on a second distance (R) defining substantially, for a playback of the sound by a playback device, a distance between a playback point (HP_i) and a point (P) of auditory perception.

2. The method as claimed in claim 1, in which, said source being far removed from the reference point (O),

- components of successive orders m are obtained for the representation of the sound in said base of spherical harmonics, and

- a filter ($1/F_m$) is applied, the coefficients of which, each applied to a component of order m, are expressed analytically in the form of the inverse of a polynomial of power m, whose variable is inversely proportional to the sound frequency and to said second distance (R), so as to compensate for a near field effect at the level of the playback device.

3. The method as claimed in claim 1, in which, said source being a virtual source envisaged at said first distance (ρ),

- components of successive orders m are obtained for the representation of the sound in said base of spherical harmonics, and

- a global filter (H_m) is applied, the coefficients of which, each applied to a component of order m, are expressed analytically in the form of a fraction; in which:

- the numerator is a polynomial of power m, whose variable is inversely proportional to the sound frequency and to said first distance (ρ), so as to simulate a near field effect of the virtual source, and

- the denominator is a polynomial of power m, whose variable is inversely proportional to the sound frequency and to said second distance (R), so as to compensate for the effect of the near field of the virtual source in the low sound frequencies.

4. The method as claimed in one of the preceding claims, in which the data coded and filtered in steps a) and b) are transmitted to the playback device with a parameter representative of said second distance (R/c).

5. The method as claimed in one of claims 1 to 3, in which, the playback device comprising means for reading a memory medium, the data coded and filtered in steps a) and b) are stored with a parameter representative of said second distance (R/c) on a memory medium intended to be read by the playback device.

6. The method as claimed in one of claims 4 and 5, in which, prior to a sound playback by a playback device comprising a plurality of loudspeakers disposed at a third distance (R_2) from said point of auditory perception (P), an adaptation filter ($H_m^{(R1/c, R2/c)}$) whose coefficients are dependent on said second (R_1) and third distances (R_2) is applied to the coded and filtered data.

7. The method as claimed in claim 6, in which the coefficients of said adaptation filter ($H_m^{(R1/c, R2/c)}$), each applied to a component of order m, are expressed analytically in the form of a fraction, in which:

- the numerator is a polynomial of power m, whose variable is inversely proportional to the sound frequency and to said second distance (R),

- and the denominator is a polynomial of power m , whose variable is inversely proportional to the sound frequency and to said third distance (R_2),

8. The method as claimed in one of claims 2, 3 and 7, in which, for the implementation of step b), there is provided:

- in respect of the components of even order m , audiodigital filters in the form of a cascade of cells of order two; and
- in respect of the components of odd order m , audiodigital filters in the form of a cascade of cells of order two and an additional cell of order one.

9. The method as claimed in claim 8, in which the coefficients of an audiodigital filter, for a component of order m , are defined from the numerical values of the roots of said polynomials of power m .

10. The method as claimed in one of claims 2, 3, 7, 8 and 9, in which said polynomials are Bessel polynomials.

11. The method as claimed in one of claims 1, 2 and 4 to 10, in which there is provided a microphone comprising an array of acoustic transducers arranged substantially on the surface of a sphere whose center corresponds substantially to said reference point (O), so as to obtain said signals representative of at least one sound propagating in the three-dimensional space.

12. The method as claimed in claim 11, in which a global filter is applied in step b) so as, on the one hand, to compensate for a near field effect as a function of said second distance (R) and, on the other hand, to equalize the signals arising from the transducers so as to compensate for a weighting of directivity of said transducers.

13. The method as claimed in one of claims 11 and 12, in which there is provided a number of transducers that depends on a total number of components chosen to represent the sound in said base of spherical harmonics.

14. The method as claimed in one of the preceding claims, in which in step a) a total number of components is chosen from the base of spherical harmonics so as to obtain, on playback, a region of the space around the point of perception (P) in which the playback of the sound is faithful and whose dimensions are increasing with the total number of components.

15. The method as claimed in claim 14, in which there is provided a playback device comprising a number of loudspeakers at least equal to said total number of components.

16. The method as claimed in one of claims 1 to 5 and 8 to 13, in which:

- there is provided a playback device comprising at least a first and a second loudspeaker disposed at a chosen distance from a listener,
- a cue of awareness of the position in space of sound sources situated at a predetermined reference distance (R) from the listener is obtained for this listener, and
- the compensation of step b) is applied with said reference distance substantially as second distance.

17. The method as claimed in one of claims 1 to 3 and 8 to 13, taken in combination with one of claims 4 and 5, in which:

- there is provided a playback device comprising at least a first and a second loudspeaker disposed at a chosen distance from a listener,
- a cue of awareness of the position in space of sound sources situated at a predetermined reference distance (R_2) from the listener is obtained for this listener, and
- prior to a sound playback by the playback device, an adaptation filter ($H_m^{(R/c, R_2/c)}$), whose coefficients are dependent on the second distance (R) and substantially on the reference distance (R_2), is applied to the data coded and filtered in steps a) and b).

18. The method as claimed in one of claims 16 and 17, in which:

- the playback device comprises a headset with two headphones for the respective ears of the listener, and
- separately for each headphone, the coding and the filtering of steps a) and b) are applied with regard to respective signals intended to be fed to each headphone, with, as first distance (p), respectively a distance (r_R , r_L) separating each ear from a position (M) of a source to be played back.

19. The method as claimed in one of the preceding claims, in which a matrix system is fashioned, in steps a) and b), said system comprising at least:

- a matrix (B) comprising said components in the base of spherical harmonics, and
- a diagonal matrix ($\text{Diag} (1/F_m)$) whose coefficients correspond to filtering coefficients of step b), and said matrices are multiplied to obtain a result matrix of compensated components ($\tilde{B}$).

20. The method as claimed in claim 19, in which:

- the playback device comprises a plurality of loudspeakers disposed substantially at one and the same distance (R) from the point of auditory perception (P), and
- to decode said data coded and filtered in steps a) and b) and to form signals suitable for feeding said loudspeakers:

- * a matrix system is formed comprising said result matrix ($\tilde{B}$) and a predetermined decoding matrix (D), specific to the playback device, and
- * a matrix (S) is obtained comprising coefficients representative of the loudspeakers feed signals by multiplication of the matrix of the compensated components (B) by said decoding matrix (D).

21. A. sound acquisition device, comprising a microphone furnished with an array of acoustic transducers disposed substantially on the surface of a sphere, **characterized in that** it furthermore comprises a processing unit arranged so as to:

- receive signals each emanating from a transducer,
- apply a coding to said signals so as to obtain a representation of the sound by components (B_{mn}^o) expressed in a base of spherical harmonics, of origin corresponding to the center of said sphere (O),
- and apply a filtering to said components (B_{mn}^o), which filtering is dependent, on the one hand, on a distance corresponding to the radius of the sphere (r) and, on the other hand, on a reference distance (R).

22. The device as claimed in claim 21, **characterized in that** said filtering consists, on the one hand, in equalizing, as a function of the radius of the sphere, the signals arising from the transducers so as to compensate for a weighting of directivity of said transducers and, on the other hand, in compensating for a near field effect as a function of a chosen reference distance (R), defining substantially, for a playback of the sound, a distance between a playback point (HP_i) and a point (P) of auditory perception.

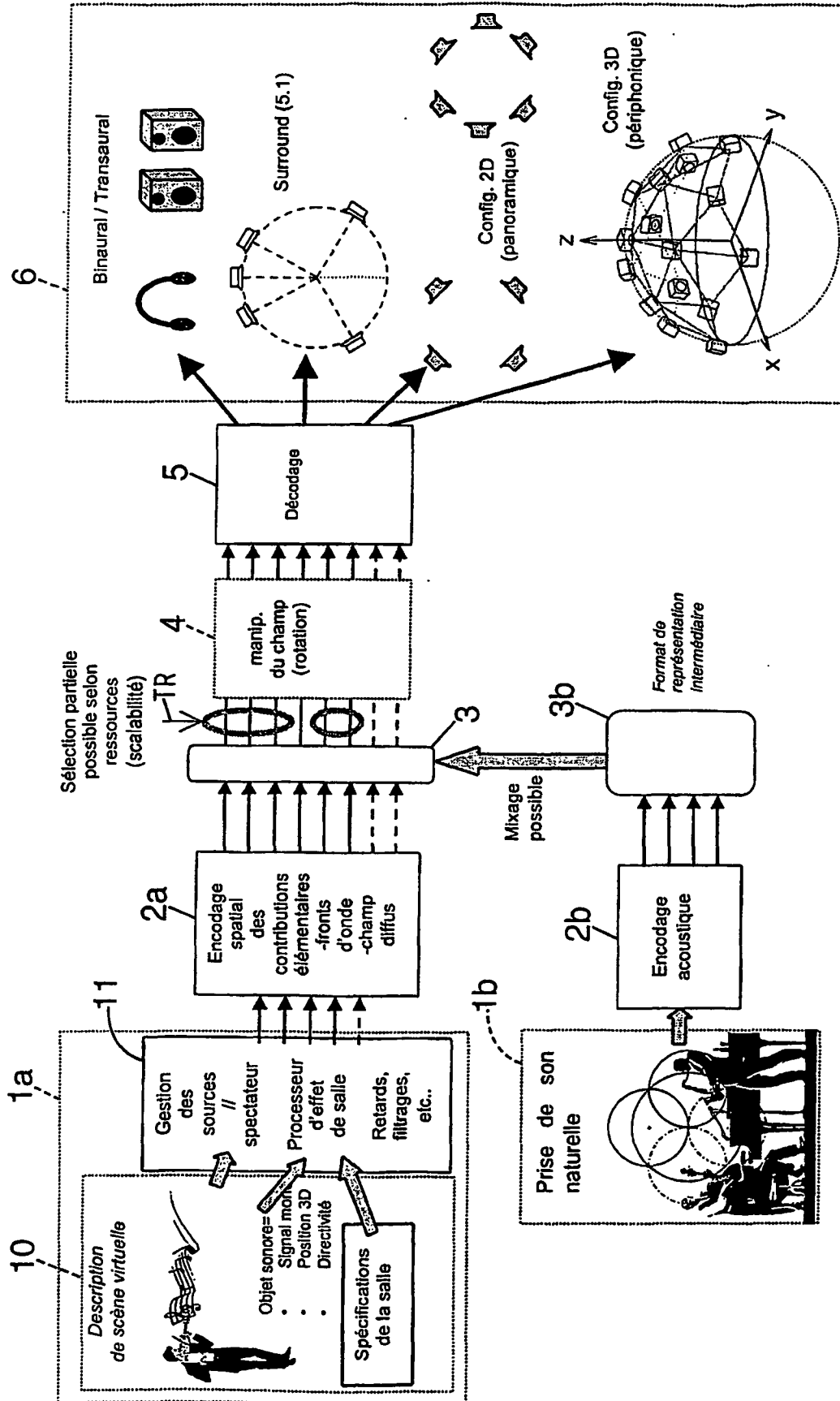


FIG. 1

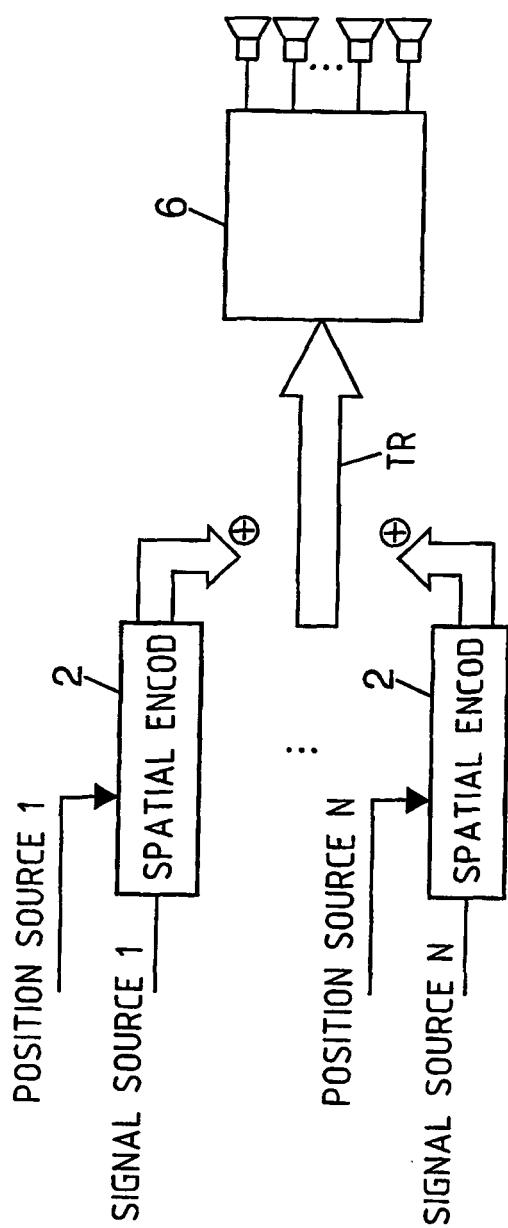


FIG. 2

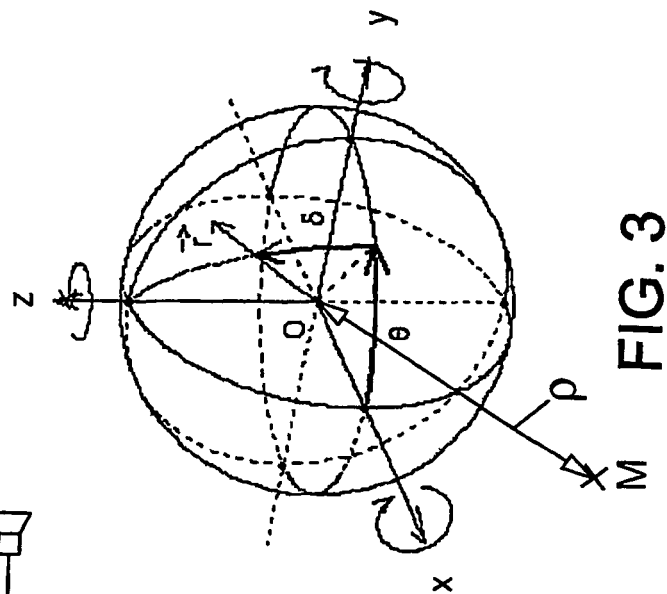
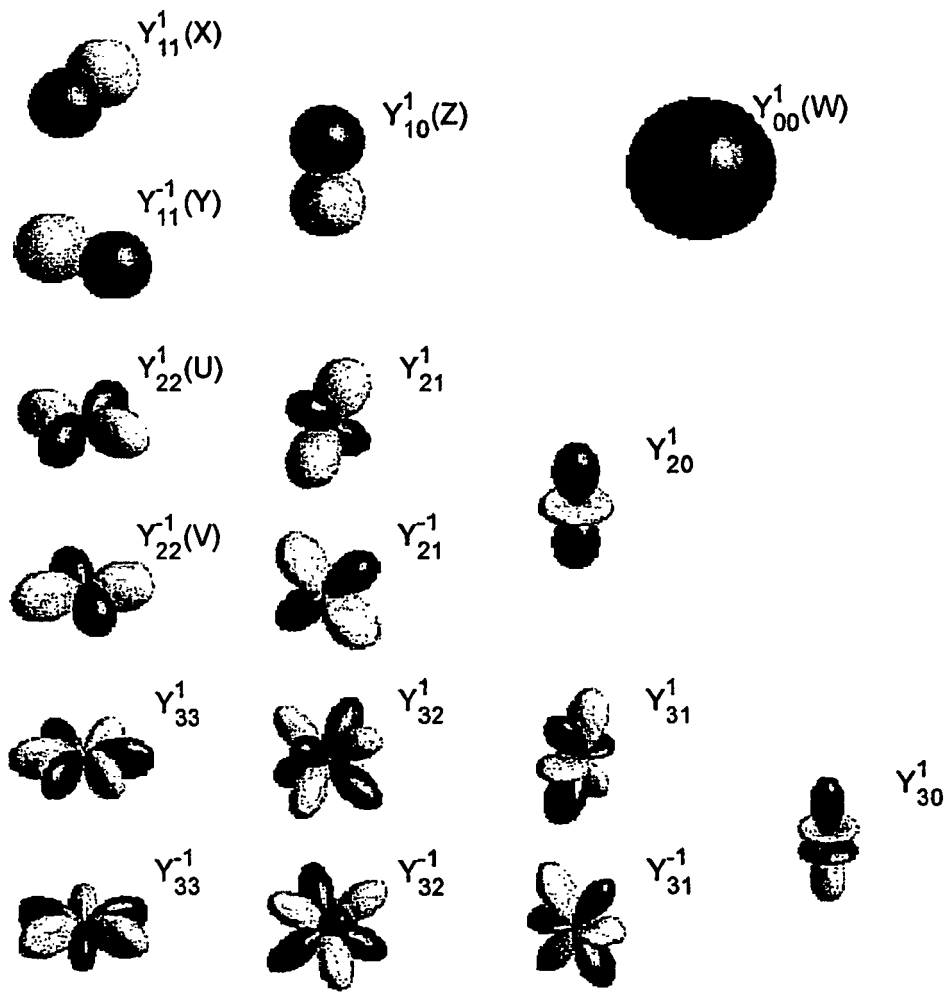


FIG. 3

FIG. 4



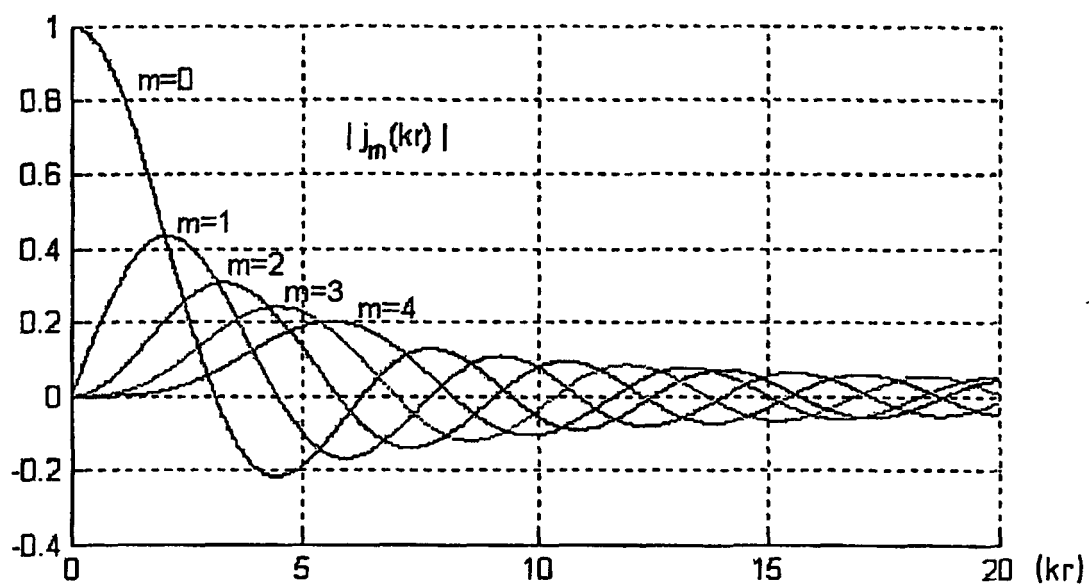


FIG. 5

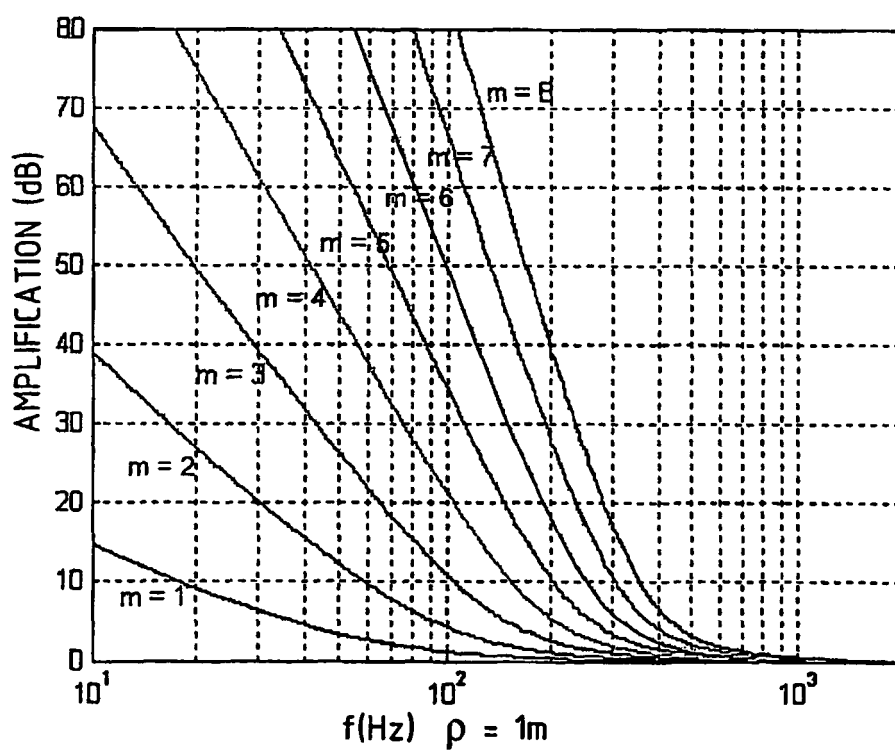


FIG. 6

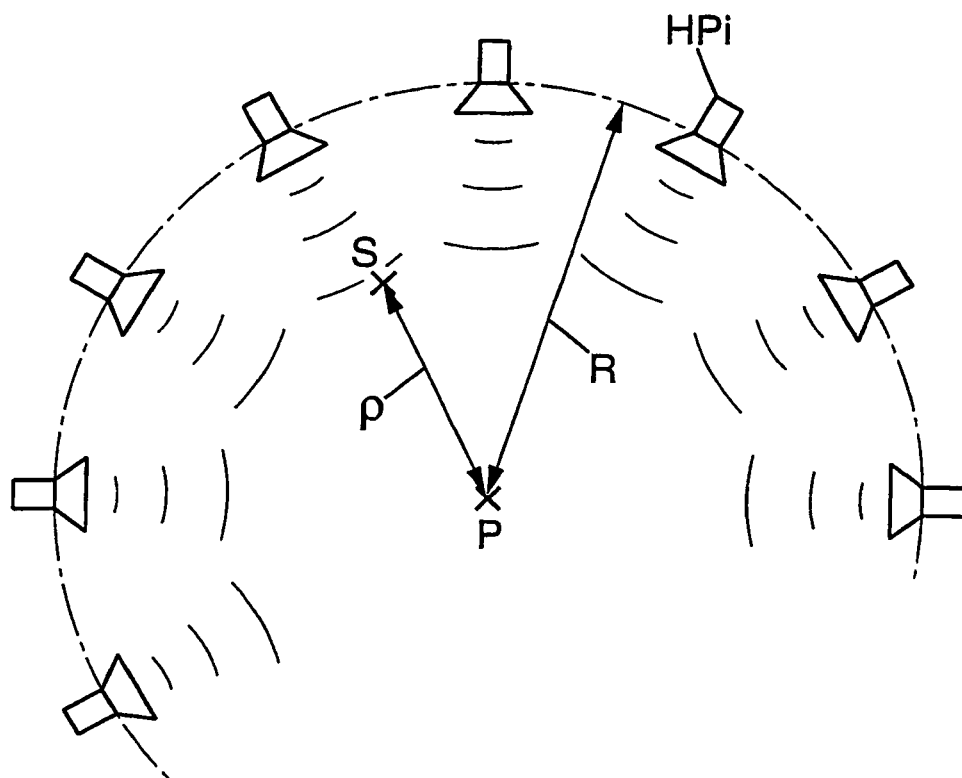
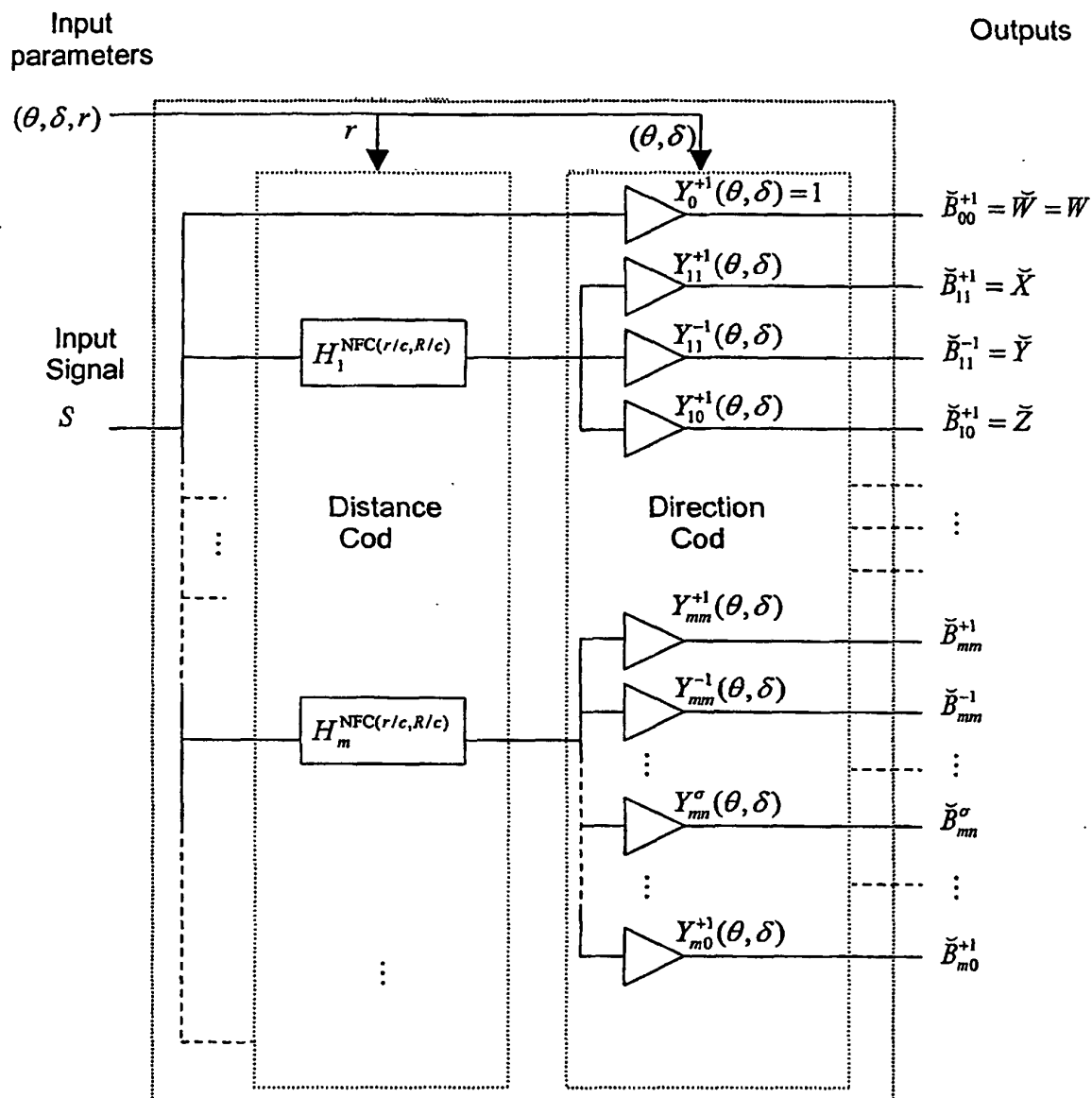


FIG. 7

FIG. 8



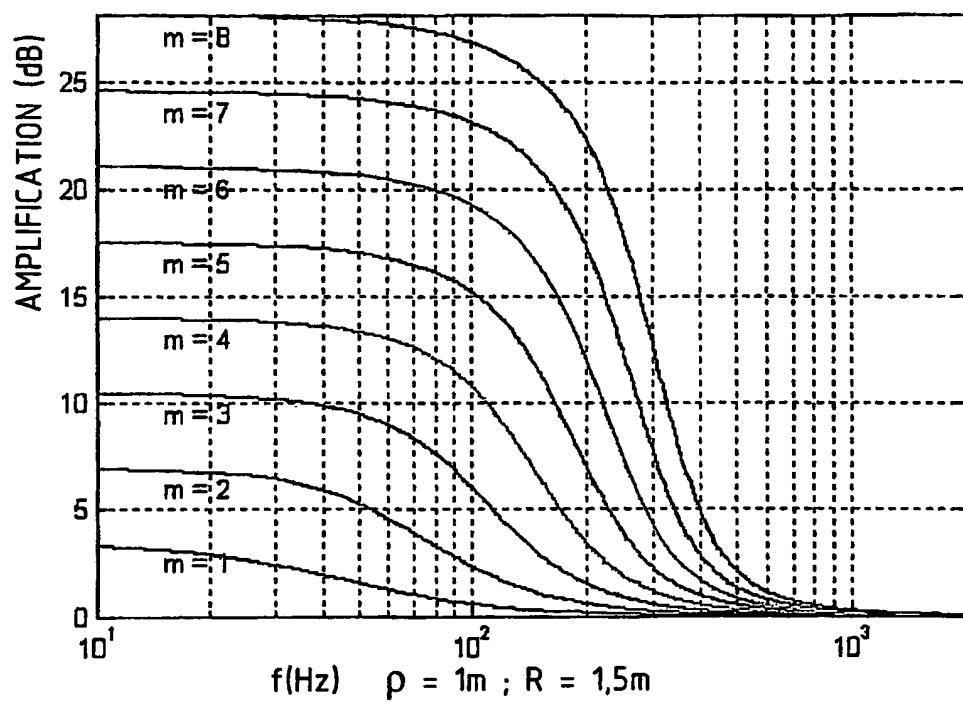


FIG. 9

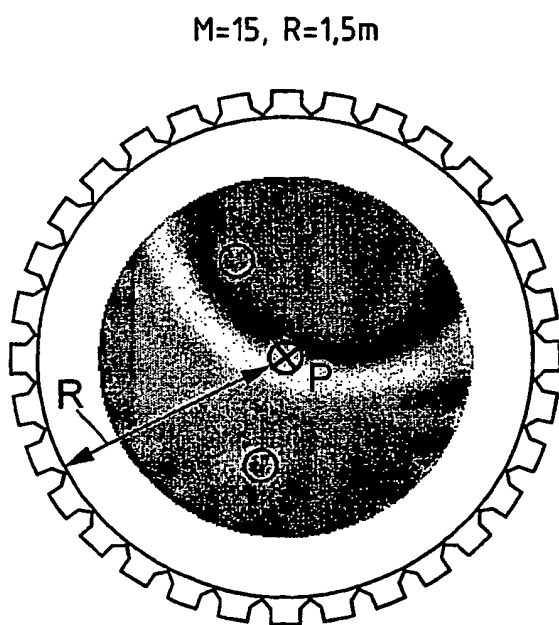


FIG. 11A

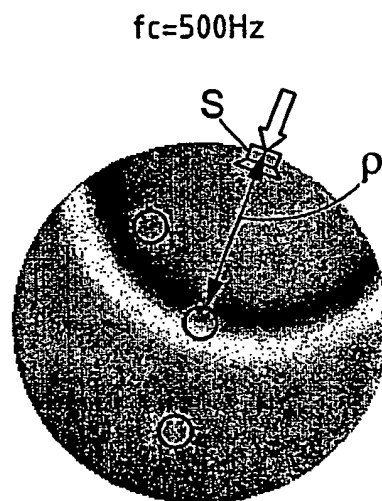


FIG. 11B

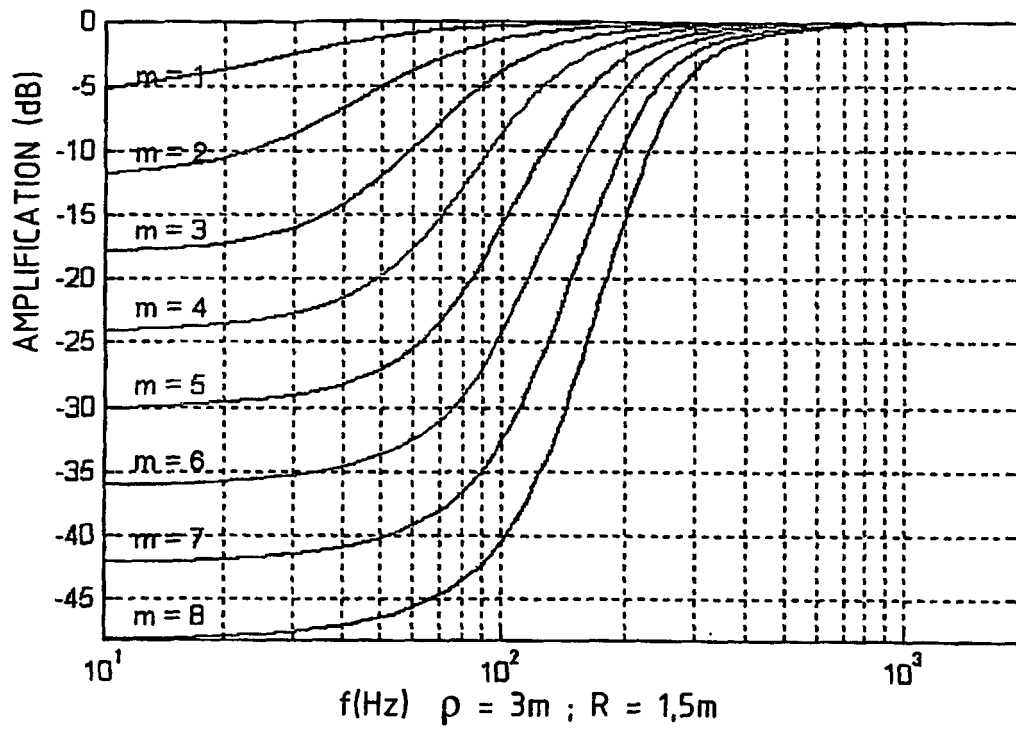


FIG. 10

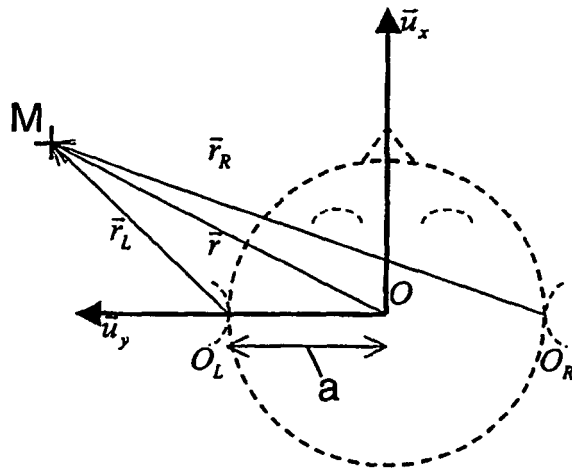
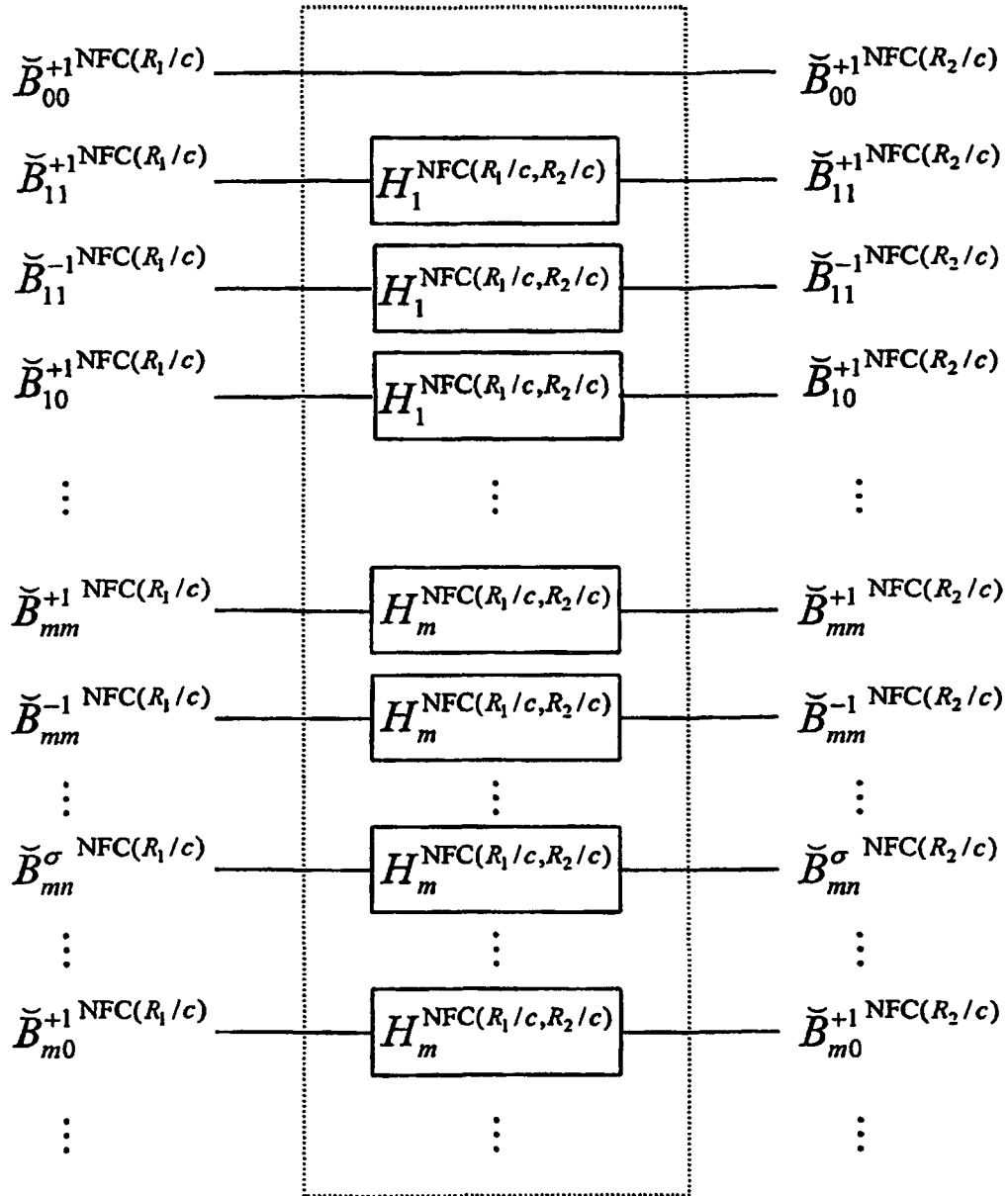


FIG. 13A

FIG. 12



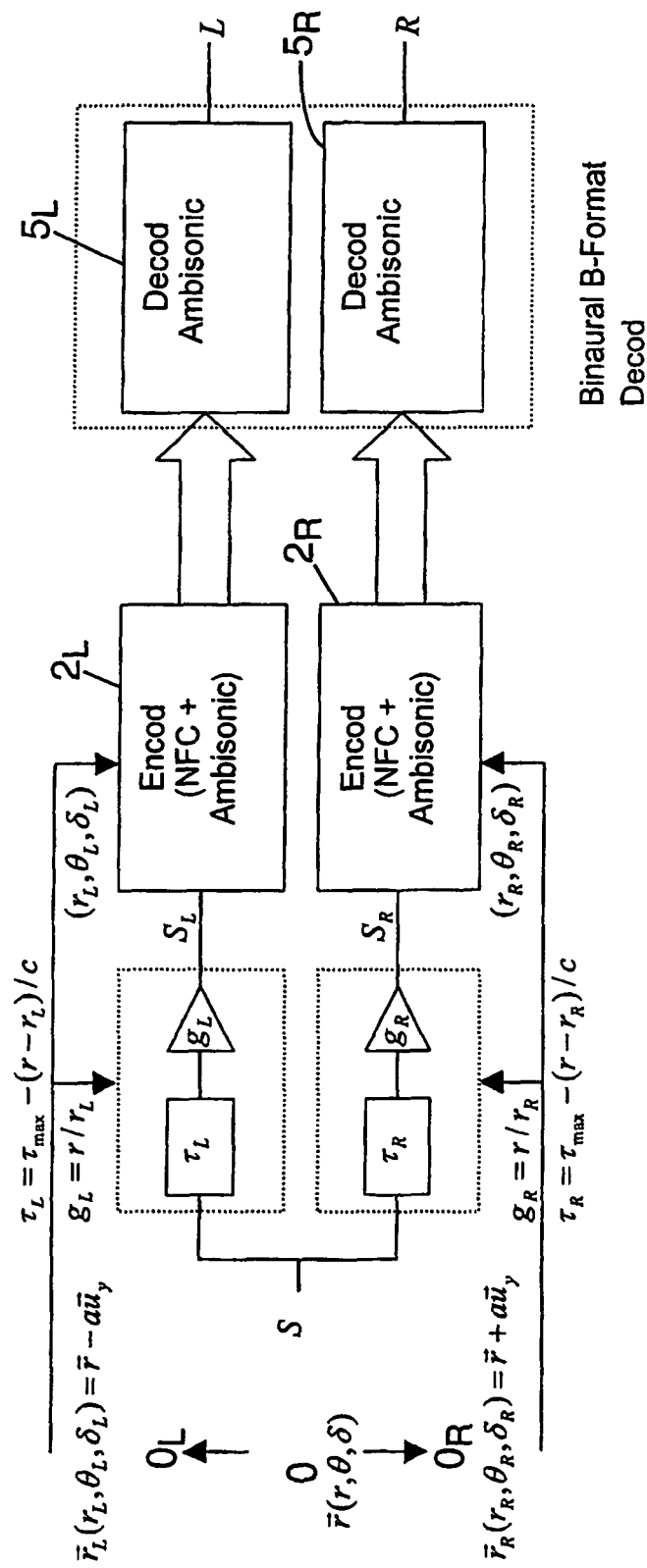


FIG. 13B

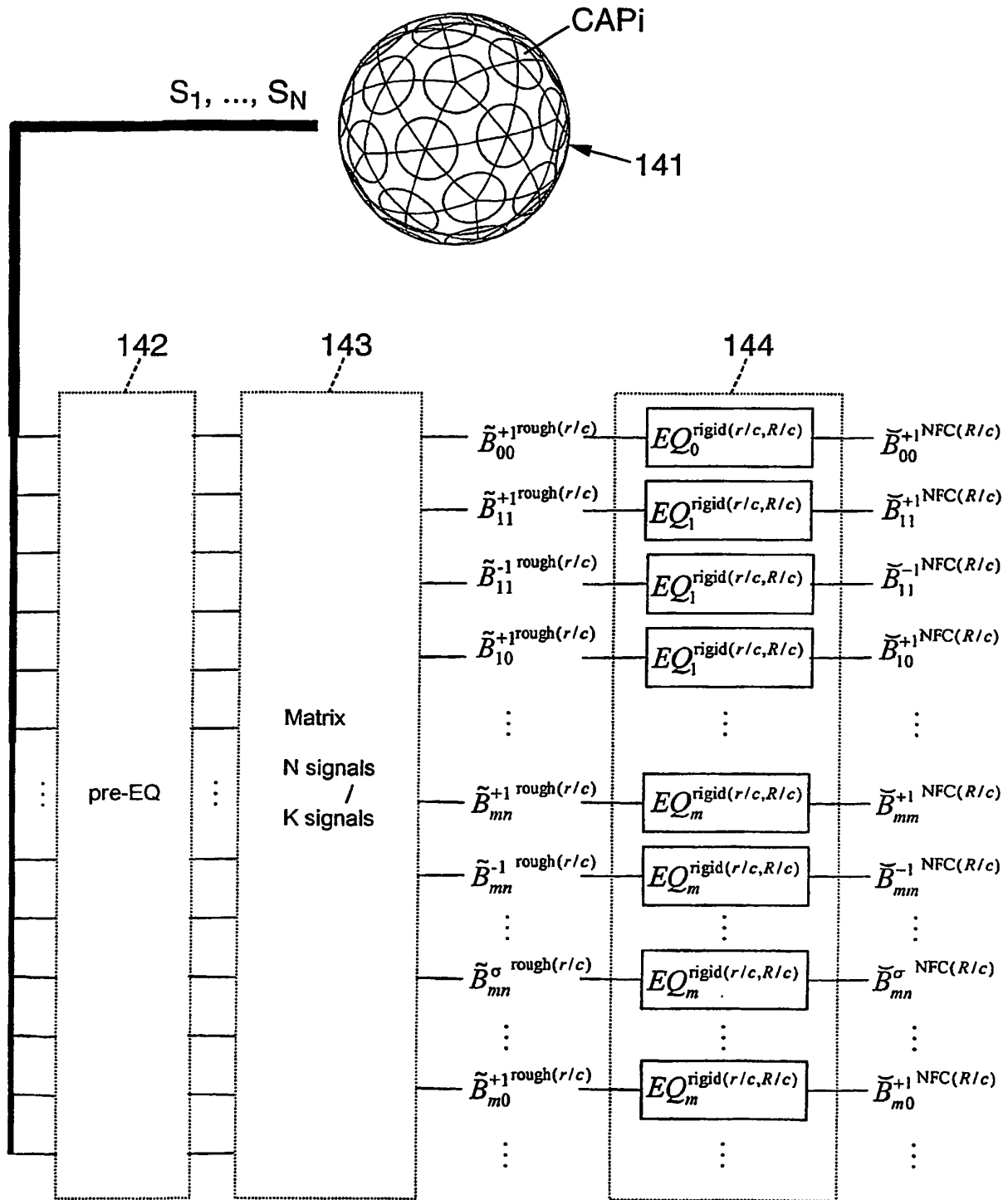


FIG. 14