(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) **EP 2 898 707 B1**

(12) **FASCICULE DE BREVET EUROPEEN**

(45) Date de publication et mention de la délivrance du brevet:
22.04.2020 Bulletin 2020/17

(21) Numéro de dépôt: **13774728.3**

(22) Date de dépôt: **05.09.2013**

(51) Int Cl.:
H04S 7/00 *(2006.01)*

(86) Numéro de dépôt international:
PCT/FR2013/052047

(87) Numéro de publication internationale:
WO 2014/044948 (27.03.2014 Gazette 2014/13)

(54) **CALIBRATION OPTIMISEE D'UN SYSTEME DE RESTITUTION SONORE MULTI HAUT-PARLEURS**

OPTIMIERTE KALIBRIERUNG EINES KLANGWIEDERGABESYSTEMS MIT MEHREREN LAUTSPRECHERN

OPTIMIZED CALIBRATION OF A MULTI-LOUDSPEAKER SOUND RESTITUTION SYSTEM

(84) Etats contractants désignés:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priorité: **18.09.2012 FR 1258760**

(43) Date de publication de la demande:
29.07.2015 Bulletin 2015/31

(73) Titulaire: **Orange**
75015 Paris (FR)

(72) Inventeurs:
- **DEPREZ, Romain**
  75020 Paris (FR)
- **NICOL, Rozenn**
  F-22450 La Roche Derrien (FR)

(56) Documents cités:
US-A1- 2006 262 939

- **Romain Deprez ET AL: "Validation théorique de la correction des réflexions sur la base d'une représentation en harmoniques sphériques", 10ème Congrès Français d'Acoustique,, 12 avril 2010 (2010-04-12), pages 1-6, XP055062229, Lyon Extrait de l'Internet: URL:http://hal.archives-ouvertes.fr/docs/00/55/09/06/PDF/000599.pdf [extrait le 2013-05-07]**
- **HACIHABIBOGLU H ET AL: "Perceptual simplification for model-based binaural room auralisation", APPLIED ACOUSTICS, ELSEVIER PUBLISHING, GB, vol. 69, no. 8, 1 août 2008 (2008-08-01), pages 715-727, XP022703192, ISSN: 0003-682X, DOI: 10.1016/J.APACOUST.2007.02.006 [extrait le 2008-06-03]**
- **RENE E. JENSEN, TODD S. WELTI: "The Importance of Reflections in a BinauralRoom Impulse Response", AES 114TH CONVENTION, 22 mars 2003 (2003-03-22), - 25 mars 2003 (2003-03-25), pages 1-18, XP040372098, Amsterdam, The Netherlands**
- **JÖRG M. BUCHHOLZ, JOHN MOURJOPOULOS, JENS BLAUERT: "Room Masking:Understanding and Modelling the Masking of Room Reflections", AES 110TH CONVENTION 2001, 12 mai 2001 (2001-05-12), - 15 mai 2001 (2001-05-15), pages 1-7, XP040371707, AMSTERDAM, THE NETHERLANDS**
- **ETIENNE CORTEEL, ROZENN NICOL: "Listening room compensation for Wave Field Synthesis. What can be done?", AES 23 RD INTERNATIONAL CONFERENCE, 23 mai 2003 (2003-05-23), - 25 mai 2003 (2003-05-25), pages 1-17, XP040374481, COPENHAGEN, DENMARK**

Il est rappelé que: Dans un délai de neuf mois à compter de la publication de la mention de la délivrance du brevet européen au Bulletin européen des brevets, toute personne peut faire opposition à ce brevet auprès de l'Office européen des brevets, conformément au règlement d'exécution. L'opposition n'est réputée formée qu'après le paiement de la taxe d'opposition. (Art. 99(1) Convention sur le brevet européen).

Printed by Jouve, 75001 PARIS (FR)

EP 2 898 707 B1

Description

[0001] La présente invention se rapporte à un procédé et un dispositif de calibration d'un système de restitution sonore comportant une pluralité de haut-parleurs ou d'éléments de restitution sonore. La calibration permet d'optimiser la qualité d'écoute du système de restitution que constitue l'ensemble des éléments de restitution, comprenant le dispositif des haut-parleurs et la salle d'écoute.

[0002] Les systèmes de restitution particulièrement concernés sont les systèmes de restitution sonore de type multi-canal (5.1, 7.1, 10.2, 22.2, etc...) ou encore de type ambisonique (Ambisonics en anglais ou Higher Order Ambisonics (HOA)).

[0003] Pour permettre une restitution de bonne qualité des signaux multicanaux, les dispositifs actuels de calibration de l'acoustique du lieu d'écoute sont basés sur une méthode générale de type "égalisation multicanale" dans laquelle les réponses impulsionnelles de chaque haut-parleur du système de restitution sont mesurées à l'aide d'un ou plusieurs microphones en un ou plusieurs points du lieu d'écoute et un filtrage d'égalisation fréquentielle est effectué sur chaque haut-parleur, indépendamment, en inversant tout ou partie de la réponse impulsionnelle mesurée pour le haut-parleur concerné.

[0004] L'inversion vise à corriger la réponse du haut-parleur de façon à ce qu'elle se rapproche au mieux d'une courbe "cible" généralement définie dans le domaine fréquentiel pour améliorer le rendu du timbre des sources sonores.

[0005] Une telle méthode est par exemple décrite dans le document intitulé "Digital Filter Design for Inversion Problems in Sound Reproduction", des auteurs Kirkeby et Nelson, dans JAES 7/8, pp.583-595, 1999.

[0006] Ce type de calibration ou correction se focalise sur la correction de l'aspect fréquentiel de la réponse du système de restitution du lieu d'écoute sans exploiter les informations temporelles comme les phénomènes de réflexions et notamment les premières réflexions des signaux sonores.

[0007] Or les premières réflexions de signaux sonores ont un impact non négligeable sur la perception auditive du signal sonore restitué.

[0008] De plus, l'analyse des réponses impulsionnelles effectuée dans les méthodes de calibration existantes est de type monophonique, c'est-à-dire qu'elle ne prend pas non plus en compte l'information spatiale des réflexions comme la direction d'incidence.

[0009] L'absence de données temporelles et spatiales des réflexions, ne permet pas de prendre en compte le rôle des ces réflexions sur la perception de l'onde directe du signal sonore par un auditeur, et ainsi d'ajuster la correction en fonction de leur effet spécifique. La qualité du signal sonore restitué et perçue par l'auditeur n'est alors pas optimale.

[0010] Les techniques de l'état de l'art sont basées sur l'application de filtres de correction sur chacun des canaux du signal multi canal, c'est-à-dire que chaque haut-parleur du système de restitution est corrigé individuellement sans tenir compte de l'ensemble du réseau de haut-parleurs.

[0011] Le document intitulé «Validation de la correction des réflexions sur la base d'une représentation en harmoniques sphériques » de Romain Deprez ET AL publié au 10ième congrès Français d'Acoustique le 12 avril 2010 (2010-04-12), pages 1-6 XP055062229, décrit une méthode de correction de signaux alimentant les haut-parleurs pour minimiser les effets des réflexions engendrées par ces haut-parleurs. Cette méthode ne prend pas en compte la perception auditive des réflexions sonores pour adapter l'ensemble de restitution sonore.

[0012] Il existe donc un besoin d'optimisation du calibrage effectué sur les systèmes de restitution de signaux audio multicanaux pour d'une part prendre en compte les propriétés temporelles et spatiales des réflexions sonores qui impactent la perception auditive des ondes directes, afin d'ajuster l'effort de traitement selon la perceptibilité des dégradations et ainsi limiter les artefacts audibles susceptibles d'être générés par les traitements trop contraints effectués dans les méthodes de calibration existantes ; et d'autre part utiliser de façon conjointe les différents haut-parleurs, afin de répartir l'effort de traitement sur l'ensemble des haut-parleurs.

[0013] La présente invention vient améliorer la situation.

[0014] Elle propose à cet effet, un procédé de calibration d'un ensemble de restitution sonore d'un signal sonore multi canal comportant une pluralité de haut-parleurs. Le procédé est tel qu'il comporte les étapes suivantes:

- obtention de réponses impulsionnelles multidirectionnelles des haut-parleurs de l'ensemble de restitution à la reproduction d'un signal audio de référence;
- analyse des réponses impulsionnelles multidirectionnelles obtenues, dans un domaine de représentation spatio-temporelle, sur au moins une fenêtre temporelle englobant les instants d'arrivée des premières réflexions du signal audio de référence reproduit pour déterminer un ensemble de caractéristiques des ondes directes et des premières réflexions comprenant au moins l'amplitude;
- comparaison de l'amplitude de chacune des réflexions à un seuil de perceptibilité déterminé en fonction de caractéristiques de l'onde directe et des premières réflexions du signal audio de référence et identification des réflexions non perceptibles pour lesquelles l'amplitude est inférieure au seuil déterminé;
- modification des réponses impulsionnelles obtenues pour obtenir des réponses impulsionnelles perceptives, par

suppression des réflexions identifiées comme non perceptibles;

- détermination d'une matrice de filtrage par les étapes de :

  - détermination d'un signal d'erreur défini par la différence entre un signal de réponse cible d'un ensemble de restitution et un signal de réponse reconstruit à partir des réponses impulsionnelles perceptives;
  - inversion multicanale par minimisation du signal d'erreur ainsi déterminé pour obtenir les filtres de la matrice de filtrage,

pour une application de cette matrice de filtrage au signal audio multi canal avant restitution sonore.

[0015] Ainsi, dans la mise en œuvre de la correction du système de restitution audio multi canal, l'effet des premières réflexions des ondes sonores diffusées par le système de restitution sur la perception auditive des ondes directes est évalué et pris en compte pour adapter le traitement appliqué aux canaux du signal multi canal selon l'effet perceptif spécifique associé à chaque réflexion. Le filtrage des canaux du signal multi canal prend ainsi en compte exclusivement les réflexions qui ont un impact sur la perception auditive des ondes directes.

[0016] Ceci permet donc d'augmenter la qualité du signal audio restitué.

[0017] De plus, comme il n'est pas nécessaire de prendre en compte les réflexions qui ne sont pas perceptibles, au sens ou leur amplitude est inférieure à un seuil de perceptibilité, les contraintes de la correction sont allégées du fait qu'elles prennent en compte les réponses impulsionnelles perceptives au lieu des réponses impulsionnelles brutes. De plus, certaines des réflexions non perceptibles qui sont éliminées des réponses impulsionnelles obtenues correspondent à des composantes de la réponse impulsionnelle qui sont justement à l'origine d'instabilités du traitement (notamment des composantes à phase non minimale). Avec les réponses impulsionnelles perceptives, on diminue ainsi les risques d'instabilités et d'artefacts qui peuvent être générés lors de traitements prenant en compte la totalité des réflexions.

[0018] Le signal d'erreur ainsi déterminé permet de prendre en compte dans le calcul de la matrice de filtrage, uniquement les réflexions qui ont un impact sur la perception auditive de l'onde directe. En effet, seules les réflexions qui ne sont pas perceptibles sont enlevées pour la détermination du signal d'erreur.

[0019] L'influence des réflexions sur la perception de l'onde directe dépend en effet de plusieurs caractéristiques des réflexions. Avantageusement, le seuil de perceptibilité peut être obtenu à partir de caractéristiques déterminées par l'étape d'analyse des réponses impulsionnelles multidirectionnelles des haut-parleurs.

[0020] Les différents modes particuliers de réalisation mentionnés ci-après peuvent être ajoutés indépendamment ou en combinaison les uns avec les autres, aux étapes du procédé défini ci-dessus.

[0021] Plus particulièrement, le seuil de perceptibilité est déterminé en fonction de la direction d'incidence de l'onde directe et/ou de son amplitude, et des directions d'incidences des premières réflexions et/ou de leurs délais d'arrivée par rapport à l'onde directe.

[0022] L'effet d'une réflexion sur la perception de l'onde directe dépend généralement de cinq paramètres au total ; d'une part il dépend de deux caractéristiques de l'onde directe : son amplitude et sa direction; d'autre part il dépend de trois caractéristiques de la réflexion : son amplitude, son instant d'arrivée et son incidence.

[0023] Cependant, si l'une des caractéristiques de l'onde directe n'est pas connue, il est possible d'estimer la caractéristique manquante en fixant à une valeur arbitraire l'autre caractéristique.

[0024] De même, si l'une des informations concernant les réflexions n'est pas connue, on peut par exemple estimer l'effet perceptif de la réflexion en fixant à une valeur arbitraire la caractéristique manquante, en prenant par exemple la valeur correspondant au cas le plus défavorable afin de majorer la perceptibilité. Ainsi, dans le cas où seule l'information de direction des réflexions est connue, il est possible de fixer une valeur à la caractéristique d'instant d'arrivée de la réflexion pour déterminer une valeur du seuil de perceptibilité uniquement par rapport à la valeur de la direction, de même si seule l'information d'instant d'arrivée de la réflexion est connue, on peut fixer la valeur de direction et déterminer le seuil de perceptibilité uniquement selon la valeur de l'instant d'arrivée. Enfin, dans le cas où les deux caractéristiques sont connues, la valeur du seuil peut être déterminée en fonction de ces deux caractéristiques.

[0025] Dans un mode de réalisation possible, le signal de réponse cible correspond à la réponse de l'onde directe seule sans aucune réflexion.

[0026] Ceci permet de prendre en compte comme signal de référence un signal dépourvu de tout effet de salle.

[0027] Dans une première variante de réalisation, le signal de réponse cible correspond à la réponse d'une onde directe associée à des réflexions représentatives d'un lieu d'écoute prédéterminé.

[0028] La réponse de référence peut alors être volontairement choisie comme un lieu d'écoute voulue dans lequel le son est à une qualité souhaitée.

[0029] Dans une seconde variante de réalisation, le signal de réponse cible correspond à la réponse d'une onde directe associée à des réflexions représentatives d'un ensemble de restitution différent.

[0030] La réponse de référence est ici choisie en fonction d'un système de restitution de référence choisi, dans lequel le nombre et la position des haut-parleurs peuvent être différents du système de restitution faisant l'objet de la correction.

[0031] La présente invention vise également un dispositif de calibration d'un ensemble de restitution sonore d'un signal sonore multi canal comportant une pluralité de haut-parleurs. Ce dispositif est tel qu'il comporte:

- un module d'obtention de réponses impulsionnelles multidirectionnelles des haut-parleurs de l'ensemble de restitution à la reproduction d'un signal audio de référence;
- un module d'analyse des réponses impulsionnelles multidirectionnelles obtenues, dans un domaine de représentation spatio-temporelle, sur au moins une fenêtre temporelle englobant les instants d'arrivée des premières réflexions du signal audio de référence reproduit pour déterminer un ensemble de caractéristiques des ondes directe et des premières réflexions associées comprenant au moins l'amplitude;
- un module de comparaison de l'amplitude de chacune des réflexions à un seuil de perceptibilité déterminé en fonction de caractéristiques de l'onde directe et des premières réflexions du signal audio de référence et d'identification des réflexions non perceptibles pour lesquelles l'amplitude est inférieure au seuil déterminé;
- un module de modification des réponses impulsionnelles obtenues pour obtenir des réponses impulsionnelles perceptives, par suppression des réflexions identifiées comme non perceptibles par le module d'identification;
- un module de calcul d'une matrice de filtrage apte à mettre en œuvre les étapes de:

    - détermination d'un signal d'erreur défini par la différence entre un signal de réponse cible d'un ensemble de restitution et un signal de réponse reconstruit à partir des réponses impulsionnelles perceptives;
    - inversion multicanale par minimisation du signal d'erreur ainsi déterminé pour obtenir les filtres de la matrice de filtrage ;

    pour une application de cette matrice de filtrage au signal audio multi canal avant restitution sonore.

[0032] Ce dispositif présente les mêmes avantages que le procédé décrit précédemment, qu'il met en œuvre.

[0033] L'invention vise également un décodeur audio comportant un dispositif de calibration tel que décrit.

[0034] Elle vise un programme informatique comportant des instructions de code pour la mise en œuvre des étapes du procédé de calibration tel que décrit, lorsque ces instructions sont exécutées par un processeur.

[0035] Enfin l'invention se rapporte à un support de stockage, lisible par un processeur, intégré ou non au dispositif de calibration, éventuellement amovible, mémorisant un programme informatique mettant en œuvre un procédé de calibration tel que décrit précédemment.

[0036] D'autres caractéristiques et avantages de l'invention apparaîtront plus clairement à la lecture de la description suivante, donnée uniquement à titre d'exemple non limitatif, et faite en référence aux dessins annexés, sur lesquels :

- la figure 1 représente un système de restitution sonore et un dispositif de calibration du système de restitution selon un mode de réalisation de l'invention;
- la figure 2 représente sous forme d'organigramme les étapes principales d'un procédé de calibration selon un mode de réalisation de l'invention;
- la figure 3a est une représentation d'un repère sphérique;
- la figure 3b, illustre les composantes harmoniques sphériques dans le cas d'une représentation spatiale ambisonique d'ordre 3;
- la figure 4 représente un exemple de tableau de valeurs en dB que peut prendre le seuil de perceptibilité utilisé dans le procédé de calibration selon un mode de réalisation de l'invention, pour un son direct d'angle d'incidence de 60°, en fonction de l'angle d'incidence (exprimé en degrés) de la réflexion et du temps d'arrivée (exprimé en ms) de cette réflexion par rapport à l'instant t0 d'arrivée de l'onde directe; le seuil de perceptibilité est défini comme le niveau (en dB) de la réflexion auquel est soustrait le niveau (en dB) de l'onde directe ;
- la figure 5 propose une autre illustration des valeurs prises par le seuil de perceptibilité : le seuil est cette fois représenté en fonction de l'incidence de la réflexion, et ceci pour différentes directions de l'onde directe ; dans tous les cas, le retard de la réflexion par rapport à l'onde directe est fixe et vaut 15 ms;
- la figure 6 représente un exemple d'une réponse impulsionnelle d'un haut-parleur d'un système de restitution ; le seuil de perceptibilité associé à chaque réflexion est également reproduit par une courbe pointillée;
- la figure 7 représente un exemple de réalisation matérielle d'un dispositif de calibration selon un mode de réalisation de l'invention.

[0037] La **figure 1** illustre donc un exemple de système de restitution sonore dans lequel le procédé de calibration selon un mode de réalisation de l'invention est mis en œuvre. Ce système comporte un dispositif de traitement 100 comportant un dispositif de calibration E selon un mode de réalisation de l'invention pilotant un ensemble de restitution 180 qui comporte une pluralité d'éléments de restitutions (haut-parleurs, enceintes acoustiques, ...) représentés ici par des haut-parleurs HP_1, HP_2, HP_3, HP_i et HP_N.

[0038] Ces haut-parleurs sont agencés dans un lieu d'écoute dans lequel un microphone ou ensemble de microphones MA est aussi prévu.

[0039] Ces haut-parleurs et microphones sont pilotés par un dispositif de traitement 100 qui peut être un décodeur tel qu'un décodeur de salon de type "set top box" pour lire ou diffuser des contenus audio ou vidéo, un serveur de traitement apte à traiter des contenus audio et vidéo et à les retransmettre à l'ensemble de restitution, un pont de conférence apte à traiter les signaux audio de différents lieux de conférence ou tout dispositif de traitement audio de signal multi canal.

[0040] Le dispositif de traitement 100 comporte un dispositif de calibration E selon un mode de réalisation de l'invention et une matrice de filtrage 170 composée d'une pluralité de filtres de traitement qui sont déterminés par le dispositif de calibration selon un procédé de calibration tel qu'illustré ultérieurement en référence à la figure 2.

[0041] Cette matrice de filtrage reçoit en entrée un signal multi canal Si et transmet en sortie les signaux SC_1, SC_2, SC_i, SC_N aptes à être restitués par l'ensemble de restitution 180.

[0042] Le dispositif de calibration E comporte un module de réception et d'émission 110 apte à transmettre d'une part des signaux audio de référence (Sref) aux différents haut-parleurs de l'ensemble de restitution 180 et à recevoir par le microphone ou l'ensemble de microphones MA, les réponses impulsionnelles multidirectionnelles (RIs) de ces différents haut-parleurs correspondant à la diffusion de ces signaux de référence.

[0043] Une réponse impulsionnelle multidirectionnelle contient l'information temporelle et l'information spatiale relatives à l'ensemble des ondes sonores induites par le haut-parleur considéré dans la salle de reproduction.

[0044] Les signaux de référence sont par exemple des signaux dont la fréquence augmente de façon logarithmique avec le temps, ces signaux étant appelés en anglais "chirps" ou "sweeps" logarithmiques.

[0045] La convolution du signal mesuré à la sortie du haut-parleur avec un signal de référence inverse permet d'obtenir directement la réponse impulsionnelle du haut-parleur.

[0046] Dans un mode de réalisation particulier adapté au domaine de représentation des harmoniques sphériques lié au format ambisonique ou HOA, le microphone apte à mesurer les réponses impulsionnelles multidirectionnelles des haut-parleurs est un microphone de type HOA placé en un point du lieu d'écoute, par exemple au centre des haut-parleurs de l'ensemble de restitution.

[0047] Ce microphone va recevoir, pour chaque haut-parleur restituant un signal audio de référence, le son restitué dans plusieurs directions. En effet, le microphone HOA est constitué d'une pluralité de microphones. Par un traitement approprié, l'information spatiale des différents sons captés peut être extraite. Pour plus de détails sur ce type de microphone, on peut se référer au document intitulé "Etude et réalisation d'outils avancés d'encodage spatial pour la technique de spatialisation sonore Higher Order Ambisonics : microphone 3D et contrôle de la distance" de S. Moreau cité à Univ. du Maine, PhD thesis, 2006.

[0048] Le microphone HOA récupère alors les réponses impulsionnelles multidirectionnelles de chacun des haut-parleurs pour les transmettre au dispositif de calibration ou pour les stocker en mémoire dans un espace mémoire local ou distant.

[0049] Lorsque que ces informations sont stockées en mémoire, l'obtention de ces réponses impulsionnelles multidirectionnelles par le dispositif de calibration selon l'invention, s'effectue alors par une simple lecture en mémoire.

[0050] Ces réponses impulsionnelles multidirectionnelles permettent d'obtenir des informations sur les directions d'arrivée des ondes directes et des réflexions du signal restitué ainsi que des informations de temps d'arrivée à la fois des ondes directes et des réflexions.

[0051] Le module d'analyse 120 du dispositif E effectue une analyse conjointe des réponses impulsionnelles obtenues, ce qui permet d'obtenir ces caractéristiques et notamment les caractéristiques des premières réflexions des signaux restitués. Dans le mode de réalisation particulier adapté au domaine de représentation des harmoniques sphériques, les réponses impulsionnelles multidirectionnelles sont obtenues dans une représentation spatio-temporelle où l'information spatiale est décrite sur la base des harmoniques sphériques et permet d'identifier les directions d'incidence des différentes composantes sonores. Ainsi, on obtient au final l'ensemble des informations sur l'amplitude des réflexions, leurs directions d'arrivée et leurs temps d'arrivée en comparaison au temps d'arrivée de l'onde directe. Cette étape sera décrite ultérieurement en référence à la figure 2.

[0052] L'analyse des réponses impulsionnelles est faite sur une échelle temporelle prédéterminée, englobant les instants des premières réflexions.

[0053] Dans un exemple de réalisation cette fenêtre temporelle est de longueur comprise entre 50 et 100 ms, ce qui correspond à l'échelle temporelle des instants d'arrivée des premières réflexions.

[0054] Bien entendu, le mode de réalisation ainsi décrit est adapté au domaine de représentation des harmoniques sphériques mais il est tout à fait envisageable d'effectuer ces même étapes dans un domaine de représentation WFS (pour "Wave Field Synthesis" en anglais) ou dans le domaine des ondes planes. Dans ces cas de figures, les moyens de captation des signaux restitués par les haut-parleurs seront à adapter à ces domaines de représentation pour obtenir des réponses impulsionnelles multidirectionnelles, sans que cela s'éloigne du cadre de l'invention.

[0055] Le dispositif de calibration E comporte également un module 130 de comparaison et d'identification des ré-

flexions non perceptibles. Ce module met en œuvre une étape de comparaison des amplitudes des réflexions, obtenues par le module d'analyse 120, à un seuil de perceptibilité Se prédéterminé. Ce seuil de perceptibilité est déterminé par le module 140 à partir d'une table de valeurs prédéfinie et stockée dans un espace mémoire.

[0056] La détermination de ce seuil de perceptibilité sera explicitée ultérieurement en référence aux figures 4 et 5.

[0057] Dans le cas où l'amplitude d'une réflexion est inférieure au seuil de perceptibilité tel que défini, cela veut dire que cette réflexion n'a pas d'impact significatif sur la perception auditive de l'onde directe du signal restitué.

[0058] Une étape d'identification des ces réflexions "non perceptibles "est alors mise en œuvre par le module 130. Ces réflexions identifiées permettent de mettre en œuvre par le module 150 une étape de détermination de réponses impulsionnelles perceptives qui sont déduites des réponses impulsionnelles obtenues par le module 110 par suppression des réflexions jugées comme non perceptibles.

[0059] Ainsi, seules les réflexions qui ont un impact sur la perception des ondes directes sont prises en compte pour calculer dans le module 160, la matrice de filtrage Filt. du module de filtrage matriciel 170.

[0060] La **figure 2** illustre sous forme d'organigramme, les étapes principales mises en œuvre dans un mode de réalisation du procédé de calibration selon l'invention.

[0061] A l'étape E201, les réponses impulsionnelles multidirectionnelles des différents haut-parleurs de l'ensemble de restitution tel que décrit en référence à la figure 1, sont obtenues. Elles sont obtenues par le dispositif de calibration, soit par simple lecture en mémoire si celles-ci ont été sauvegardées au préalable, soit par réception du microphone ou d'un ensemble de microphones ayant effectué la mesure.

[0062] Ces réponses impulsionnelles multidirectionnelles sont les réponses de chaque haut-parleur suite à la reproduction d'un signal de référence tel que décrit en référence à la figure 1.

[0063] Une étape E202 d'analyse des réponses impulsionnelles multidirectionnelles ainsi obtenues est alors mise en œuvre. Cette analyse s'effectue dans un domaine de représentation spatio-temporelle. L'information spatiale peut par exemple être décrite dans le domaine de représentation des harmoniques sphériques. Dans cette représentation illustrée à la **figure 3a,** chaque point a pour coordonnées sphériques, une distance r par rapport à l'origine 0, un angle θ d'azimut ou d'orientation dans le plan horizontal et un angle δ d'élévation ou d'orientation dans le plan vertical. Préférentiellement, la direction définie par (θ=0°, δ=0°) correspond à la direction en face de l'auditeur. Dans un tel repère, une onde acoustique est parfaitement décrite si l'on définit en tout point à chaque instant t, la pression acoustique notée p(r, θ, δ, t) dont la transformée de Fourier temporelle est notée P(r, θ, δ, f) où f désigne la fréquence temporelle.

[0064] Dans le contexte de spatialisation ambisonique d'ordre supérieur (HOA), les composantes spatiales sont des composantes ambisoniques B_{mn}^{σ} qui correspondent à la décomposition de l'onde de pression acoustique *p* sur la base des harmoniques sphériques. Par exemple, pour une source sonore en champ lointain, c'est-à-dire une onde plane d'incidence (θ_S, δ_S) portant un signal S(t), les composantes ambisoniques B_{mn}^{σ} sont données par: $B_{mn}^{\sigma} = S(t).Y_{mn}^{\sigma}(\theta_s, \delta_s)$ où les fonctions harmoniques sphériques $Y_{mn}^{\sigma}(\theta, \delta)$ décrivent une base orthonormée:

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

Les P_{mn} (sin δ) sont les fonctions de Legendre associées.

[0065] Une illustration des fonctions harmoniques sphériques est représentée en **figure 3b.** On peut ainsi voir la composante omnidirective Y_{00}^{1} (désignée comme la « composante W » dans la terminologie ambisonique) correspondant à l'ordre 0, les composantes bidirectives Y_{10}^{1}, Y_{11}^{1}, Y_{11}^{-1} (désignée respectivement comme les «composantes Z, X et Y » dans la terminologie ambisonique) correspondant à l'ordre 1, et les composantes des ordres supérieurs.

[0066] Une représentation spatiale tridimensionnelle ou "3D" dite "d'ordre *M'* comprend $K = (M+1)^2$ composantes dont les triplets d'indices $\{m,n,\sigma\}$ sont tels que $0 \le m \le M$, $0 \le n \le m$, $\sigma = \pm 1$. Une représentation bidimensionnelle ou "2D" d'ordre *M* comprend un sous-ensemble de ces composantes en ne retenant que les indices *m*=*n,* soit $K=2M+1$ composantes.

[0067] La décomposition sur la base des harmoniques sphériques peut être considérée comme la transformée duale entre coordonnées spatiales et les fréquences spatiales. Les composantes B_{mn}^{σ} définissent donc un spectre spatial.

[0068] Pour chaque haut-parleur, on obtient à l'issue de l'étape E201, une réponse impulsionnelle multidirectionnelle qui est constituée de K réponses impulsionnelles correspondant aux K composantes de la représentation spatiale choisie.

Dans le cas de la représentation des harmoniques sphériques, il s'agit des K composantes sur les K=2M+1 harmoniques sphériques considérés. Pour le jième haut-parleur, la réponse impulsionnelle multidirectionnelle qui lui est associée se compose ainsi de K réponses élémentaires $H_{jl}(t)$ où l'indice l repère l'indice de la composante spatiale et t correspond à l'échantillon temporel. Par la suite, on désigne par $h_j(t)$ le vecteur des K composantes spatiales mesurées pour le jième haut-parleur :

$$h_j(t) = [H_{j1}(t) \dots H_{jl}(t) \dots H_{jK}(t)].$$

[0069] Si le système de reproduction comprend au total N haut-parleurs, l'ensemble des réponses impulsionnelles multidirectionnelles mesurées pour les N haut-parleurs et les K composantes spatiales définit une matrice H de taille KxN, dans laquelle la jième colonne correspond à la réponse impulsionnelle multidirectionnelle associée au jième haut-parleur.

[0070] Pour chaque haut-parleur, les K composantes spatiales contenues dans le vecteur $h_j(t)$ représentent le spectre spatial des sons captés par le microphone. Pour accéder à l'information de direction des sons, il convient donc d'effectuer une transformation inverse pour repasser d'une représentation en fonction des fréquences spatiales à une représentation en fonction des coordonnées spatiales.

[0071] Cette transformation inverse est réalisée en reconstruisant l'onde de pression p(r, θ, δ, t) par combinaison linéaire des harmoniques sphériques, chaque harmonique étant pondéré par l'amplitude de la composante qui lui est associée. On retrouve ces éléments dans la thèse de S. Moreau citée ci-dessus.

[0072] On peut alors évaluer l'onde de pression p(r, θ, δ, t) en tout point d'une sphère centrée sur le point de mesure des réponses impulsionnelles multidirectionnelles en reconstruisant l'onde de pression point par point par combinaison linéaire des harmoniques sphériques. On peut par exemple évaluer cette pression sur un réseau de P points définissant un « échantillonnage régulier » de la sphère au sens défini dans le mémoire de thèse de S. Moreau. Cette opération s'apparente alors au décodage spatial des composantes ambisoniques pour une restitution par un réseau sphérique régulier de P haut-parleurs virtuels. Cette étape de décodage spatial est par exemple décrite dans le document intitulé "Ambisonics encoding of other audio formats for multiple listening conditions" des auteurs Jérôme Daniel, Jean-Bernard Rault et Jean-Dominique Polack dans AES 105th Convention, September 1998.

[0073] En pratique, cette transformation des fréquences spatiales (composantes ambisoniques) vers les coordonnées spatiales s'effectue en multipliant, pour chaque haut-parleur et chaque échantillon temporel t, le vecteur $h_j(t)$ par une matrice de décodage D. Par exemple, la matrice D peut être obtenue comme $D=Y^T$, où la matrice Y est calculée en évaluant les K harmoniques sphériques $Y_{mn}^{\sigma}(\theta, \delta)$ pour les P directions des haut-parleurs virtuels, en regroupant les azimuths θ_q et élévations δ_q dans un unique doublet $C = (\theta_q, \delta_q)$ associé à un haut-parleur (q désigne l'indice du haut-parleur). Dans la matrice Y, chaque colonne est constituée des valeurs des K harmoniques sphériques pour un haut-parleur donné. Au final, on obtient, pour chaque haut-parleur et chaque échantillon temporel t, un vecteur $G_j(t)$ de longueur P décrivant la distribution spatiale des composantes sonores captées sur un réseau de P points définissant un échantillonnage régulier de la sphère:

$$G_j(t) = Y^T h_j(t)$$

[0074] Le maximum de cette fonction $G_j(t)$ identifie une réflexion. Si $G_j(t)$ présente plusieurs maxima, ces différents maxima identifient chacun une réflexion. Ainsi, pour chaque réflexion identifiée, ses caractéristiques sont déterminées selon la procédure suivante : son instant d'arrivée correspond à l'échantillon t_{Ri} = t pour lequel elle est identifiée, son incidence correspond aux coordonnées spatiales

$$C_{Ri} = (\theta_{Ri}, \delta_{Ri}) = (\theta_q, \delta_q)$$

du point pour lequel le maximum de $G_j(t)$ est observé, et son amplitude correspond à l'amplitude de ce maximum $A_{Ri}=G_j(t_i)$. Dans ce qui précède, l'indice i repère l'indice de la réflexion considérée. La précision d'estimation de ces caractéristiques dépend donc du nombre P de haut-parleurs virtuels utilisés pour cette analyse. Le premier échantillon temporel pour lequel on observe un maximum définit l'instant d'arrivée de l'onde directe. On a soin de relever aussi l'amplitude (A_D) et l'incidence de cette dernière ($C_D = (\theta_D, \delta_D)$ où θ_D et δ_D définissent respectivement l'angle d'azimut et l'angle d'élévation repérant la direction de l'onde directe).

[0075] Ainsi, à partir des réponses impulsionnelles multidirectionnelles obtenues, considérées sur une fenêtre d'ana-

lyse temporelle englobant les instants des premières réflexions du signal audio reproduit par les haut-parleurs, il est possible de déterminer, et ce pour chaque haut-parleur, les caractéristiques de l'onde directe et les caractéristiques des réflexions qui lui sont associées. Ainsi, pour le jième haut-parleur, sont déterminées d'une part les caractéristiques de l'onde directe comme son amplitude $A_D(j)$, son instant d'arrivée sur le microphone $T_D(j)$ ou sa direction d'incidence $C_D(j)$; et d'autre part les caractéristiques des réflexions comme leurs amplitudes $A_{Ri}(j)$, leurs instants d'arrivée sur le microphone $T_{Ri}(j)$ ou leurs directions d'incidences $C_{Ri}(j)$. Dans la suite, on utilisera plutôt l'amplitude normalisée par l'amplitude de l'onde directe :

$$AN_{Ri}(j) = \frac{A_{Ri}(j)}{A_D(j)},$$

et le retard entre l'onde directe et la réflexion :

$$\tau_{Ri}(j) = T_{Ri}(j) - T_D(j).$$

[0076] Les premières réflexions d'un signal audio restitué dépendent du lieu d'écoute dans lequel est placé l'ensemble de restitution. D'une façon générale, ces premières réflexions apparaissent dans un temps situé dans une plage allant de 50 à 100ms après l'onde directe.

[0077] De façon avantageuse, la fenêtre temporelle d'analyse de l'étape E202 sera, dans un mode de réalisation adapté, d'une taille comprise entre 50 et 100 ms.

[0078] L'étape E203 compare les amplitudes obtenues par l'étape d'analyse à un seuil de perceptibilité Se des réflexions qui a été défini au préalable et stocké en mémoire. L'étape E204 permet de retrouver la valeur de seuil prédéfinie en fonction de caractéristiques de chaque réflexion et de l'onde directe associée, obtenues à l'étape d'analyse E202.

[0079] En effet, plusieurs cas de figure peuvent se présenter. Dans un premier exemple de réalisation, seule l'information de direction des réflexions est connue et récupérée de l'étape d'analyse. Pour retrouver le seuil de perceptibilité correspondant, on fixe la valeur de la caractéristique d'instant d'arrivée de la réflexion, par exemple la valeur la plus critique (celle qui donne une perceptibilité maximale) et on détermine la valeur du seuil de perceptibilité uniquement par rapport à la valeur de la direction.

[0080] De même si seule l'information d'instant d'arrivée de la réflexion est connue, on peut fixer la valeur de direction, par exemple la valeur la plus critique (celle qui donne une perceptibilité maximale), et déterminer le seuil de perceptibilité selon la valeur de l'instant d'arrivée.

[0081] Enfin, dans le cas où les deux caractéristiques sont connues, la valeur du seuil peut être déterminée, avec une meilleure précision, en fonction de ces deux caractéristiques.

[0082] Pour cela, un tableau de valeurs de seuil de perceptibilité est stocké en mémoire. Un exemple d'un tel tableau est illustré en référence à la **figure 4.** Ce tableau montre, pour un son direct situé à un angle d'azimut à 60°, la valeur du seuil de perceptibilité d'une réflexion exprimée en dB, en fonction des caractéristiques d'angle d'incidence de la réflexion (i.e. son angle d'azimut θ_{Ri} dans le plan horizontal correspondant à l'élévation δ_{Ri}= 0°) et de temps d'arrivée de cette réflexion par rapport au temps d'arrivée de l'onde directe $\tau_{Ri}(j)$. Le seuil est défini comme le niveau relatif de la réflexion, c'est-à-dire qu'il représente la différence entre les valeurs d'amplitude (exprimées en dB) de la réflexion et de l'onde directe considérée.

[0083] Ce tableau de valeurs est un exemple de valeurs seuils définies à partir d'expériences psycho-acoustiques réalisées en considérant différents types de signal sonore (parole, clics, musique, etc...), différents angles d'incidences et différents temps d'arrivée des réflexions et de l'onde directe. Un seuil de perceptibilité de ces réflexions est défini en fonction de ces paramètres.

[0084] Pour compléter l'illustration des valeurs du seuil de perceptibilité de la figure 4, la **figure 5** montre différentes courbes de seuil de perceptibilité exprimé en dB (qui correspond toujours au seuil relatif correspondant à la différence entre le niveau de la réflexion et celui de l'onde directe). Ces différentes courbes correspondent à différentes positions de l'onde directe (azimut de 0° pour D1, 60° pour D2, 90° pour D3 et 150° pour D4) et représentent les seuils de perceptibilité en fonction de la direction de la réflexion, ceci pour un temps d'arrivée fixe (correspondant en l'occurrence à 15 ms).

[0085] Ainsi, à l'étape E204, la valeur de seuil correspondant aux caractéristiques obtenues à l'étape d'analyse est récupérée. On compare cette valeur de seuil à la valeur d'amplitude de chaque réflexion à l'étape E203. Pour être comparée au seuil de perceptibilité, la valeur de l'amplitude de la réflexion est référencée à celle de l'onde directe associée et exprimée en dB: $20\log(AN_{Ri}(j))$.

[0086] Dans le cas où la valeur d'amplitude de la réflexion est inférieure à la valeur de seuil de perceptibilité, cela veut dire que cette réflexion n'a pas d'impact sur la perception que peut avoir un auditeur de l'onde directe. Cette réflexion

n'est donc pas à prendre en compte pour le traitement d'un signal multi canal avant restitution. L'étape E203 permet ainsi d'identifier toutes les réflexions qui n'ont pas d'impact sur la perception de l'onde directe. L'étape E203 identifie donc toutes les réflexions pour lesquelles l'amplitude est inférieure au seuil de perceptibilité.

[0087] Pour illustrer cette étape E203, la **figure 6** représente un exemple de réponse impulsionnelle, pour une direction donnée, d'un des haut-parleurs de l'ensemble de restitution en comparaison avec la courbe en trait discontinu représentant le seuil de perceptibilité (RMT pour « Reflection Masked Threshold ») obtenu par la table décrite ci-dessus en référence à la figure 4. Les réflexions dont le niveau est inférieur à la courbe de seuil sont ainsi identifiées. On note que dans le cas illustré, les premières réflexions survenant dans les 15 premières ms ne sont pas perceptibles.

[0088] A partir de cette identification des réflexions non perceptibles, l'étape E205 effectue une modification des réponses impulsionnelles $h_j(t)$ obtenues à l'étape E201 pour les j=1 à N haut-parleurs, pour obtenir des réponses impulsionnelles perceptives $hp_j(t)$. Pour cela, la modification consiste à éliminer les réflexions non perceptibles identifiées à l'étape E203 dans les réponses impulsionnelles.

[0089] De façon plus détaillée, cette opération s'effectue par exemple par une opération de seuillage. A chaque instant t, la valeur du seuil de perceptibilité Se est retranchée au signal de réponse impulsionnelle qui a été obtenue à l'étape E201.

[0090] Préférentiellement ce traitement est appliqué sur le spectre spatial défini par les K composantes $h_j(t) = [H_{j1}(t) \ldots H_{jl}(t) \ldots H_{jK}(t)]$ dans le domaine de représentation spatiale choisi, correspondant par exemple à la représentation sur la base des harmoniques sphériques. Cependant le traitement peut aussi s'appliquer dans le domaine dual des coordonnées d'espace. Dans la suite, nous allons décrire l'opération réalisée dans le cas du spectre spatial.

[0091] L'opération de seuillage consiste à comparer pour chaque réflexion identifiée son amplitude au seuil de perceptibilité Se associé à ses caractéristiques. Ainsi, pour la ième réflexion identifiée pour le jième haut-parleur, le seuil Se(i) est déterminé en fonction de ses caractéristiques $[\tau_{Ri}(j), C_{Ri}(j)]$. Cette réflexion est localisée à l'instant t_i donné par:

$$t_i = T_D(j) + \tau_{Ri}(j) \ .$$

[0092] Pour réaliser le seuillage, on considère donc la réponse impulsionnelle à cet instant, soit $h_j(t_i)$, ou plus exactement sur le spectre spatial associé et constitué des K composantes $[H_{j1}(t_i) \ldots H_{jl}(t_i) \ldots H_{jK}(t_i)]$. Plusieurs stratégies sont alors possibles. La plus simple consiste à préserver l'amplitude relative des composantes du spectre spatial, c'est-à-dire qu'on applique un traitement identique à toutes les composantes. Dans ce cas, pour chaque composante $H_{jl}(t_i)$, l'opération de seuillage peut se traduire par les équations suivantes:

$$HP_{jl}(t_i) = 0 \qquad \text{si } AN_{Ri}(j) \leq 10^{0.05\,Se}$$

$$HP_{jl}(t_i) = \left(\left|H_{jl}(t_i) - 10^{0.05Se}\right|\right) \frac{H_{jl}(t_i)}{|H_{jl}(t_i)|} \qquad \text{si } AN_{Ri}(j) > 10^{0.05\,Se}$$

où $HP_{jl}(t)$ désigne la réponse impulsionnelle perceptive associée à $H_{jl}(t)$.

[0093] Ainsi, les réponses impulsionnelles perceptives ne conservent que les réflexions ayant un impact significatif sur la perception de l'onde directe.

[0094] Ces réponses impulsionnelles perceptives sont alors utilisées pour déterminer la matrice de filtrage, à l'étape E206. Cette matrice de filtrage est ensuite utilisée pour traiter le signal audio multi canal avant sa restitution sonore par l'ensemble de restitution du système.

[0095] Pour obtenir l'ensemble de filtres constituant la matrice de filtrage Filt du dispositif de traitement, un mode de réalisation possible comporte une étape de détermination d'un signal d'erreur défini par la différence entre un signal de réponse cible prédéterminé de l'ensemble de restitution et un signal de réponse reconstruit à partir des réponses impulsionnelles perceptives et une étape d'inversion multicanale par minimisation du signal d'erreur ainsi déterminé.

[0096] Le signal d'erreur ainsi obtenu ne prend donc en compte que les réflexions perceptibles puisque qu'il est calculé à partir d'un signal reconstruit basé sur les réponses impulsionnelles perceptives.

[0097] L'inversion peut être réalisée par un algorithme de descente de gradient ou ses variantes. Un exemple d'algorithme d'inversion possible est celui de type ISTA (pour "Iterative Shrinkage-Thresholding algorithm) tel que décrit dans le document intitulé " A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems" des auteurs Amir Beck & Marc Teboulle, publié dans SIAM J. IMAGING SCIENCES, Vol. 2, No. 1, pp. 183-202 en 2009.

[0098] D'une façon générale, le problème qui se pose pour calculer les filtres de la matrice de traitement, est le suivant. Il y a N haut-parleurs qui constituent le système réel de reproduction. Dans le contexte de spatialisation ambisonique d'ordre supérieur (HOA), l'espace de représentation spatiale est de dimension K. L'information spatiale est donc décrite

par K coefficients. L'objectif est de reproduire avec le système de N haut-parleurs, un ensemble de V signaux définissant le signal audio multicanal d'entrée. Ces V signaux sont dédiés à un système idéal de reproduction constitués de V haut-parleurs. Ce système idéal définit les V signaux cibles qu'on souhaite reproduire et qui correspondent donc aux réponses d'un système fictif de V haut-parleurs virtuels. Dans le cas le plus simple, le système réel de reproduction comporte aussi N=V haut-parleurs. Mais dans le cas général, on est capable d'émuler un système de V haut-parleurs virtuels à partir d'un dispositif de N haut-parleurs réels.

[0099] L'équation à résoudre est la suivante:

$$T(t)=H*W(t)$$

avec H, la matrice de dimension KxN comportant les réponses impulsionnelles des N éléments du système de restitution dans le domaine d'analyse spatiale,
W, la matrice comportant les filtres de correction à calculer, de dimension NxV,
T, la matrice contenant les V réponses cibles définies dans le domaine d'analyse spatiale, de dimension KxV,
et l'opération dénotée par « * » est un produit matriciel convolutif où un élément T_{ij} de la matrice T est obtenu de la façon suivante :

$$T_{ij}(t) = \sum_{k=1}^{N} H_{ik} * W_{kj}(t)$$

Chaque matrice est une matrice de vecteurs, au sens où la troisième dimension correspond à l'échelle des temps. L'objectif de l'opération d'inversion est de trouver les éléments de la matrice W.

[0100] La résolution de cette opération peut s'effectuer en deux temps. Tout d'abord, on calcule les filtres de correction en ne corrigeant que l'effet de salle du lieu de restitution, c'est-à-dire qu'on prend en compte le dispositif réel de haut-parleurs, soit N haut-parleurs. Dans une seconde étape, on compense la disposition des haut-parleurs pour adapter les V signaux à une restitution selon une configuration non idéale de N haut-parleurs. Dans ce but, les V signaux sont répartis par matriçage sur les N canaux associés au système réel de reproduction afin d'émuler un système de V haut-parleurs virtuels.

[0101] Dans le cas présent, pour mettre en œuvre l'invention, les éléments de la matrice H comportent les réponses impulsionnelles perceptives telles qu'obtenues à l'étape E205.

[0102] Les réponses cibles peuvent varier selon le résultat de restitution sonore attendue.

[0103] Dans un mode de réalisation, cette réponse cible correspond à la réponse impulsionnelle donnée par l'onde directe seule sans aucune réflexion. Cela revient à supprimer tout l'effet de salle dans le signal attendu.

[0104] Dans une première variante de réalisation, le signal de réponse cible correspond à la réponse d'une onde directe associée à des réflexions représentatives d'un lieu d'écoute prédéterminé.

[0105] Un lieu d'écoute caractéristique qui présente une bonne qualité d'écoute peut être souhaitée (par exemple le lieu d'écoute de la salle Pleyel™). Dans ce cas, les filtres de traitement seront calculés pour obtenir une restitution sonore proche de cette qualité d'écoute.

[0106] Dans une deuxième variante de réalisation, le signal de réponse cible correspond à la réponse d'une onde directe associée à des réflexions représentatives d'un ensemble de restitution différent de celui utilisé pour restituer le signal résultant.

[0107] Ainsi, un système de restitution souhaité, par exemple comportant plus de haut-parleurs, est pris comme référence pour obtenir une restitution proche de celle qui aurait été obtenu avec un tel système.

[0108] D'autres signaux de réponse cibles peuvent bien évidemment être choisi selon l'effet de la restitution souhaitée.

[0109] Ainsi, la mise en œuvre du procédé décrit permet d'obtenir une meilleure qualité d'écoute lors de la restitution d'un signal audio multi canal grâce à la prise en compte seule des réflexions perceptibles des signaux par l'ensemble de restitution dans le lieu d'écoute.

[0110] La **figure 7** représente un exemple de réalisation matérielle d'un dispositif de calibration selon l'invention. Celui-ci peut faire partie intégrante d'un décodeur audio/vidéo, d'un serveur de traitement, d'un pont de conférence ou de tout autre équipement de lecture ou de diffusion audio ou vidéo.

[0111] Ce type de dispositif comporte un processeur μP coopérant avec un bloc mémoire MEM comportant une mémoire de stockage et/ou de travail.

[0112] Le bloc mémoire peut avantageusement comporter un programme informatique comportant des instructions de code pour la mise en œuvre des étapes du procédé de calibration au sens de l'invention, lorsque ces instructions

sont exécutées par le processeur, et notamment les étapes d'obtention de réponses impulsionnelles multidirectionnelles des haut-parleurs de l'ensemble de restitution à la reproduction d'un signal audio prédéterminé, d'analyse des réponses impulsionnelles multidirectionnelles obtenues, dans un domaine de représentation spatio-temporelle, sur au moins une fenêtre temporelle englobant les instants d'arrivée des premières réflexions du signal audio prédéterminé reproduit pour déterminer un ensemble de caractéristiques des premières réflexions, de comparaison de l'amplitude de chacune des réflexions à un seuil de perceptibilité prédéterminé et d'identification des réflexions non perceptibles pour lesquelles l'amplitude est inférieure au seuil prédéterminé, de modification des réponses impulsionnelles obtenues pour obtenir des réponses impulsionnelles perceptives, par suppression des réflexions identifiées comme non perceptibles et de détermination d'une matrice de filtrage à partir des réponses impulsionnelles perceptives pour une application de cette matrice de filtrage au signal audio multi canal avant restitution sonore.

[0113] Typiquement, la description de la figure 2 reprend les étapes d'un algorithme d'un tel programme informatique. Le programme informatique peut également être stocké sur un support mémoire lisible par un lecteur du dispositif ou téléchargeable dans l'espace mémoire de celui-ci.

[0114] La mémoire MEM enregistre une table de valeurs de seuil de perceptibilité en fonction de caractéristiques des composantes sonores constituées de l'onde directe et des réflexions utilisée dans le procédé selon un mode de réalisation de l'invention et de manière générale, toutes les données nécessaires à la mise en œuvre du procédé.

[0115] Un tel dispositif comporte un module d'entrée I apte à recevoir des réponses impulsionnelles d'un ensemble de restitution et un module de sortie S apte à transmettre à un module de traitement, les filtres calculés d'une matrice de filtrage.

[0116] Dans un mode possible de réalisation, le dispositif ainsi décrit peut également comporter les fonctions de traitement par la mise en œuvre de la matrice de traitement à la réception en I d'un signal multi canal Si pour transmettre en sortie des signaux traités SCi aptes à être restitués par l'ensemble de restitution.

Revendications

1. Procédé de calibration d'un ensemble de restitution sonore d'un signal sonore multi canal comportant une pluralité de haut-parleurs, le procédé comportant les étapes suivantes:

    - obtention (E201) de réponses impulsionnelles multidirectionnelles des haut-parleurs de l'ensemble de restitution à la reproduction d'un signal audio de référence;
    - analyse (E202) des réponses impulsionnelles multidirectionnelles obtenues, dans un domaine de représentation spatio-temporelle, sur au moins une fenêtre temporelle englobant les instants d'arrivée des premières réflexions du signal audio de référence reproduit pour déterminer un ensemble de caractéristiques des ondes directes et des premières réflexions associées comprenant au moins l'amplitude;

    le procédé étant **caractérisé en ce qu'**il comporte les étapes suivantes :

    - comparaison (E203) de l'amplitude de chacune des réflexions à un seuil de perceptibilité déterminé (E204) en fonction de caractéristiques de l'onde directe et des premières réflexions du signal audio de référence et identification (E203) des réflexions non perceptibles pour lesquelles l'amplitude est inférieure au seuil déterminé;
    - modification (E205) des réponses impulsionnelles obtenues pour obtenir des réponses impulsionnelles perceptives, par suppression des réflexions identifiées comme non perceptibles;
    - détermination (E206) d'une matrice de filtrage par les étapes de :

        - détermination d'un signal d'erreur défini par la différence entre un signal de réponse cible d'un ensemble de restitution et un signal de réponse reconstruit à partir des réponses impulsionnelles perceptives;
        - inversion multicanale par minimisation du signal d'erreur ainsi déterminé pour obtenir les filtres de la matrice de filtrage,

    pour une application de cette matrice de filtrage au signal audio multi canal avant restitution sonore.

2. Procédé selon la revendication 1, **caractérisé en ce que** le seuil de perceptibilité est déterminé en fonction de la direction d'incidence de l'onde directe (C_D) et/ou de son amplitude (A_D), et des directions d'incidences des premières réflexions (C_{Ri}) et/ou de leurs délais d'arrivée (τ_{Ri}) par rapport à l'onde directe.

3. Procédé selon la revendication 1, **caractérisé en ce que** le signal de réponse cible correspond à la réponse de l'onde directe seule sans aucune réflexion.

4. Procédé selon la revendication 1, **caractérisé en ce que** le signal de réponse cible correspond à la réponse d'une onde directe associée à des réflexions représentatives d'un lieu d'écoute prédéterminé.

5. Procédé selon la revendication 1, **caractérisé en ce que** le signal de réponse cible correspond à la réponse d'une onde directe associée à des réflexions représentatives d'un ensemble de restitution différent.

6. Dispositif de calibration d'un ensemble de restitution sonore d'un signal sonore multi canal comportant une pluralité de haut-parleurs, le dispositif comportant:

- un module d'obtention (110) de réponses impulsionnelles multidirectionnelles des haut-parleurs de l'ensemble de restitution à la reproduction d'un signal audio de référence;
- un module d'analyse (120) des réponses impulsionnelles multidirectionnelles obtenues, dans un domaine de représentation spatio-temporelle, sur au moins une fenêtre temporelle englobant les instants d'arrivée des premières réflexions du signal audio de référence reproduit pour déterminer un ensemble de caractéristiques des ondes directes et des premières réflexions associées comprenant au moins l'amplitude;

le dispositif étant **caractérisé en ce qu'**il comporte :

- un module de comparaison (120) de l'amplitude de chacune des réflexions à un seuil de perceptibilité déterminé (140) en fonction de caractéristiques de l'onde directe et des premières réflexions du signal audio de référence et d'identification (120) des réflexions non perceptibles pour lesquelles l'amplitude est inférieure au seuil déterminé;
- un module de modification (150) des réponses impulsionnelles obtenues pour obtenir des réponses impulsionnelles perceptives, par suppression des réflexions identifiées comme non perceptibles par le module d'identification;
- un module de calcul (130) d'une matrice de filtrage apte à mettre en œuvre les étapes de:

  - détermination d'un signal d'erreur défini par la différence entre un signal de réponse cible d'un ensemble de restitution et un signal de réponse reconstruit à partir des réponses impulsionnelles perceptives;
  - inversion multicanale par minimisation du signal d'erreur ainsi déterminé pour obtenir les filtres de la matrice de filtrage ;

pour une application de cette matrice de filtrage au signal audio multi canal avant restitution sonore.

7. Décodeur audio comportant un dispositif de calibration selon la revendication 6.

8. Programme informatique comportant des instructions de code pour la mise en œuvre des étapes du procédé de calibration selon l'une des revendications 1 à 5, lorsque ces instructions sont exécutées par un processeur.

9. Support de stockage, lisible par un processeur, sur lequel est stocké un programme informatique comprenant des instructions de code pour l'exécution des étapes du procédé de calibration selon l'une des revendications 1 à 5.

Patentansprüche

1. Verfahren zur Kalibrierung einer akustischen Wiedergabeeinheit eines akustischen Mehrkanalsignals, die eine Vielzahl von Lautsprechern aufweist, wobei das Verfahren die folgenden Schritte aufweist:

- Erhalt (E201) von multidirektionalen Impulsantworten der Lautsprecher der Wiedergabeeinheit bei der Reproduktion eines Bezugs-Audiosignals;
- Analyse (E202) der erhaltenen multidirektionalen Impulsantworten in einem Bereich raum-zeitlicher Darstellung, in mindestens einem Zeitfenster, das die Ankunftsmomente der ersten Reflexionen des reproduzierten Bezugs-Audiosignals umfasst, um eine Einheit von Merkmalen der direkten Wellen und der zugeordneten ersten Reflexionen zu bestimmen, die mindestens die Amplitude enthält;

wobei das Verfahren **dadurch gekennzeichnet ist, dass** es die folgenden Schritte aufweist:

- Vergleich (E203) der Amplitude jeder der Reflexionen mit einer bestimmten Wahrnehmbarkeitsschwelle (E204)

abhängig von Merkmalen der direkten Welle und der ersten Reflexionen des Bezugs-Audiosignals und Erkennung (E203) der nicht wahrnehmbaren Reflexionen, für die die Amplitude niedriger als die bestimmte Schwelle ist;
- Änderung (E205) der erhaltenen Impulsantworten, um wahrnehmungsfähige Impulsantworten zu erhalten, durch Unterdrückung der als nicht wahrnehmbar erkannten Reflexionen;
- Bestimmung (E206) einer Filtermatrix durch die folgenden Schritte:

- Bestimmung eines Fehlersignals definiert durch die Differenz zwischen einem Ziel-Antwortsignal einer Wiedergabeeinheit und einem ausgehend von den wahrnehmungsfähigen Impulsantworten rekonstruierten Antwortsignal;
- Mehrkanalumschaltung durch Minimierung des so bestimmten Fehlersignals, um die Filter der Filtermatrix zu erhalten,

für eine Anwendung dieser Filtermatrix an das Mehrkanal-Audiosignal vor der akustischen Wiedergabe.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** die Wahrnehmbarkeitsschwelle abhängig von der Einfallrichtung der direkten Welle (C_D) und/oder ihrer Amplitude (A_D) und den Einfallrichtungen der ersten Reflexionen (C_{Ri}) und/oder ihren Ankunftsverzögerungen (τ_{Ri}) bezüglich der direkten Welle bestimmt wird.

3. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** das Ziel-Antwortsignal der Antwort der direkten Welle alleine ohne jede Reflexion entspricht.

4. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** das Ziel-Antwortsignal der Antwort einer direkten Welle zugeordnet zu Reflexionen entspricht, die für einen vorbestimmten Hörort repräsentativ sind.

5. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** das Ziel-Antwortsignal der Antwort einer direkten Welle zugeordnet zu Reflexionen entspricht, die für eine andere Wiedergabeeinheit repräsentativ sind.

6. Vorrichtung zur Kalibrierung einer akustischen Wiedergabeeinheit eines akustischen Mehrkanalsignals, die eine Vielzahl von Lautsprechern aufweist, wobei die Vorrichtung aufweist:

- ein Modul (110) zum Erhalt multidirektionaler Impulsantworten der Lautsprecher der Wiedergabeeinheit bei der Reproduktion eines Bezugs-Audiosignals;
- ein Modul (120) zur Analyse der erhaltenen multidirektionalen Impulsantworten in einem Bereich raum-zeitlicher Darstellung, in mindestens einem Zeitfenster, das die Ankunftsmomente der ersten Reflexionen des reproduzierten Bezugs-Audiosignals umfasst, um eine Einheit von Merkmalen der direkten Wellen und der zugeordneten ersten Reflexionen zu bestimmen, die mindestens die Amplitude enthält;

wobei die Vorrichtung **dadurch gekennzeichnet ist, dass** sie aufweist:

- ein Modul (120) zum Vergleich der Amplitude jeder der Reflexionen mit einer bestimmten Wahrnehmbarkeitsschwelle (140) abhängig von Merkmalen der direkten Welle und der ersten Reflexionen des Bezugs-Audiosignals und zur Erkennung (120) der nicht wahrnehmbaren Reflexionen, für die die Amplitude niedriger als die bestimmte Schwelle ist;
- ein Modul (150) zur Änderung der erhaltenen Impulsantworten, um wahrnehmungsfähige Impulsantworten zu erhalten, durch Unterdrückung der vom Erkennungsmodul als nicht wahrnehmbar erkannten Reflexionen;
- ein Modul (130) zur Berechnung einer Filtermatrix, das die folgenden Schritte durchführen kann:

- Bestimmung eines Fehlersignals definiert durch die Differenz zwischen einem Ziel-Antwortsignal einer Wiedergabeeinheit und einem ausgehend von den wahrnehmungsfähigen Impulsantworten rekonstruierten Antwortsignal;
- Mehrkanalumschaltung durch Minimierung des so bestimmten Fehlersignals, um die Filter der Filtermatrix zu erhalten,

für eine Anwendung dieser Filtermatrix an das Mehrkanal-Audiosignal vor der akustischen Wiedergabe.

7. Audio-Decoder, der eine Kalibriervorrichtung nach Anspruch 6 aufweist.

8. EDV-Programm, das Codeanweisungen zur Durchführung der Schritte des Kalibrierverfahrens nach einem der Ansprüche 1 bis 5 aufweist, wenn diese Anweisungen von einem Prozessor ausgeführt werden.

9. Prozessorlesbarer Speicherträger, auf dem ein EDV-Programm gespeichert ist, das Codeanweisungen zur Ausführung der Schritte des Kalibrierverfahrens nach einem der Ansprüche 1 bis 5 enthält.

Claims

1. Method for calibrating an assembly for sound playback of a multi-channel sound signal having a plurality of loudspeakers, the method having the following steps:

    - obtaining (E201) multi-directional impulse responses from the loudspeakers of the playback assembly upon reproduction of a reference audio signal;
    - analyzing (E202) the multi-directional impulse responses obtained, in a domain of spatio-temporal representation, over at least one time window encompassing the instants of arrival of the early reflections of the reproduced reference audio signal in order to determine a set of characteristics of the direct waves and of the associated early reflections comprising at least the amplitude;

    the method being **characterized in that** is has the following steps:

    - comparing (E203) the amplitude of each of the reflections with a determined perceptibility threshold (E204) as a function of characteristics of the direct wave and of the early reflections of the reference audio signal and identifying (E203) the non-perceptible reflections for which the amplitude is below the determined threshold;
    - modifying (E205) the impulse responses obtained in order to obtain perceptual impulse responses, by suppression of the reflections identified as non-perceptible;
    - determining (E206) a filtering matrix with the steps of:

        - determination of an error signal defined by the difference between a target response signal for a playback assembly and a response signal reconstructed from the perceptual impulse responses;
        - multi-channel inversion by minimization of the error signal thus determined in order to obtain the filters of the filtering matrix,

    for an application of this filtering matrix to the multi-channel audio signal before sound playback.

2. Method according to Claim 1, **characterized in that** the perceptibility threshold is determined as a function of the direction of incidence of the direct wave (C_D) and/or its amplitude (A_D), and the directions of incidence of the early reflections (C_{Ri}) and/or their arrival times (τ_{Ri}) with respect to the direct wave.

3. Method according to Claim 1, **characterized in that** the target response signal corresponds to the response of the direct wave alone without any reflection.

4. Method according to Claim 1, **characterized in that** the target response signal corresponds to the response of a direct wave associated with reflections representing a predetermined listening site.

5. Method according to Claim 1, **characterized in that** the target response signal corresponds to the response of a direct wave associated with reflections representing a different playback assembly.

6. Device for calibrating an assembly for sound playback of a multi-channel sound signal having a plurality of loudspeakers, the device having:

    - a module (110) for obtaining multi-directional impulse responses from the loudspeakers of the playback assembly upon reproduction of a reference audio signal;
    - a module (120) for analyzing the multi-directional impulse responses obtained, in a domain of spatio-temporal representation, over at least one time window encompassing the instants of arrival of the early reflections of the reproduced reference audio signal in order to determine a set of characteristics of the direct waves and of the associated early reflections comprising at least the amplitude;

the device being **characterized in that** it comprises:

- a module (120) for comparing the amplitude of each of the reflections with a determined perceptibility threshold (140) as a function of characteristics of the direct wave and of the early reflections of the reference audio signal and for identifying (120) the non-perceptible reflections for which the amplitude is below the determined threshold;
- a module (150) for modifying the impulse responses obtained in order to obtain perceptual impulse responses, by suppression of the reflections identified as non-perceptible by the identification module;
- a module (130) for computing a filtering matrix able to implement the steps of:

  - determination of an error signal defined by the difference between a target response signal for a playback assembly and a response signal reconstructed from the perceptual impulse responses;
  - multi-channel inversion by minimization of the error signal thus determined in order to obtain the filters of the filtering matrix;

for an application of this filtering matrix to the multi-channel audio signal before sound playback.

7. Audio decoder having a calibration device according to Claim 6.

8. Computer program having code instructions for the implementation of the steps of the calibration method according to one of Claims 1 to 5 when these instructions are executed by a processor.

9. Storage medium, readable by a processor, on which a computer program is stored comprising code instructions for the execution of the steps of the calibration method according to one of Claims 1 to 5.

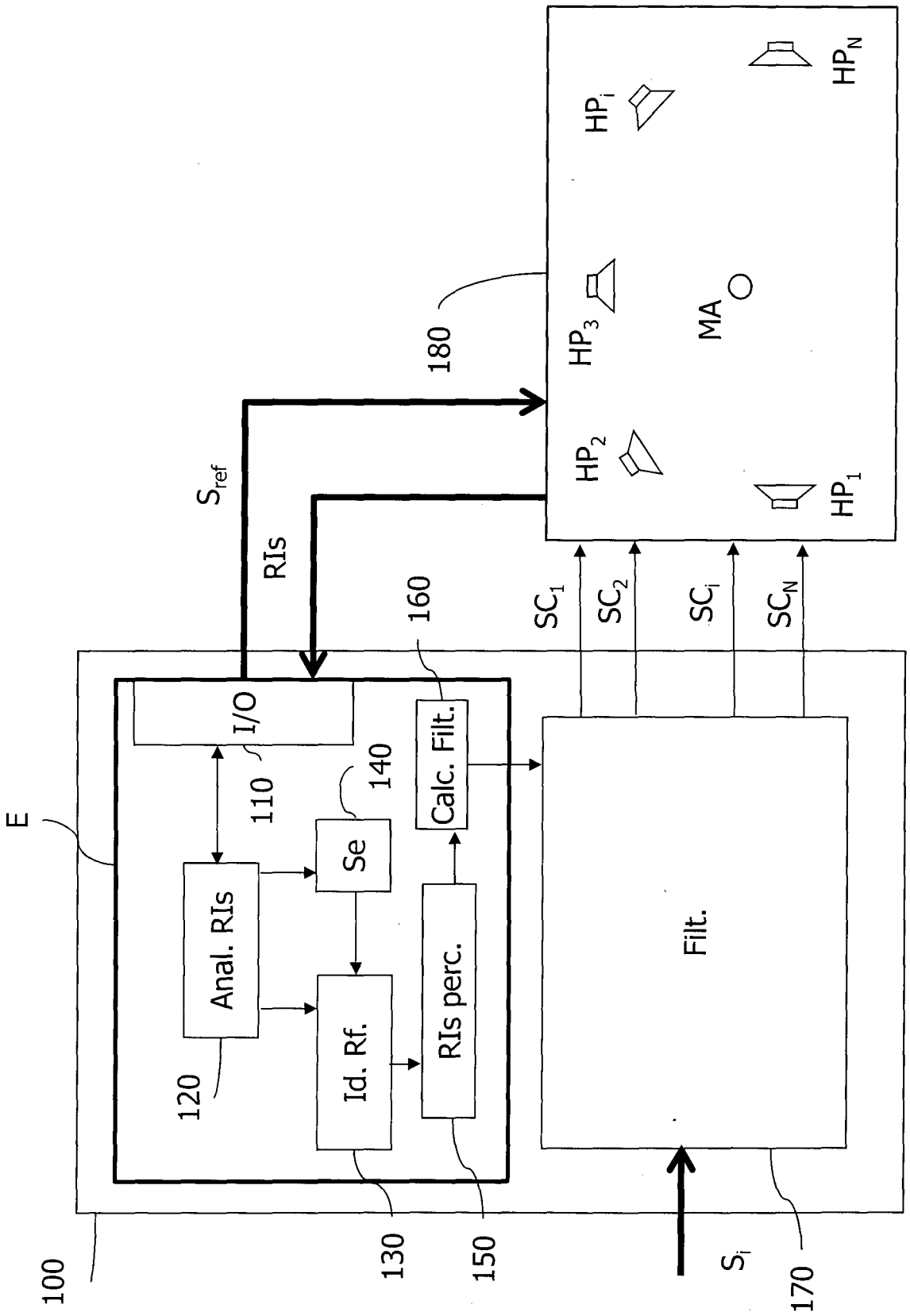

E
100
I/O
110
Anal. RIs
120
Se
140
Id. Rf.
130
Calc. Filt.
160
RIs perc.
150
Filt.
S_i
170
S_{ref}
RIs
180
HP_1
HP_2
HP_3
HP_i
HP_N
MA
SC_1
SC_2
SC_i
SC_N

Fig.1

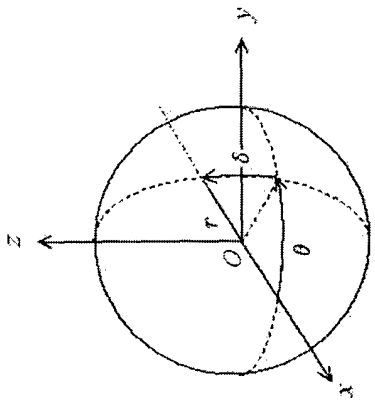


Fig.3a

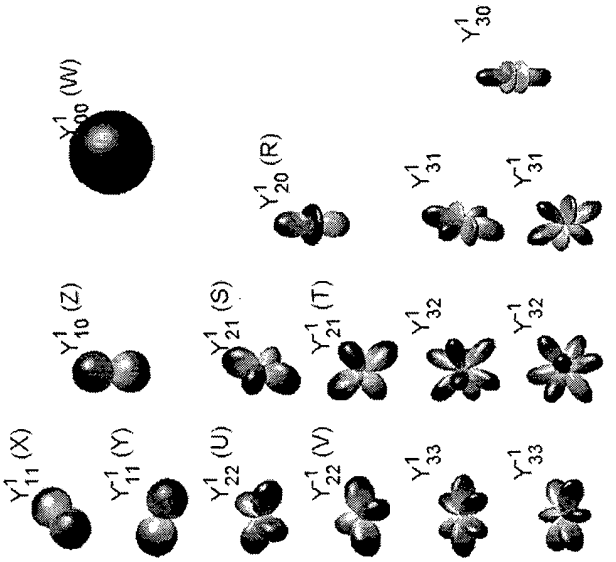


Fig.3b

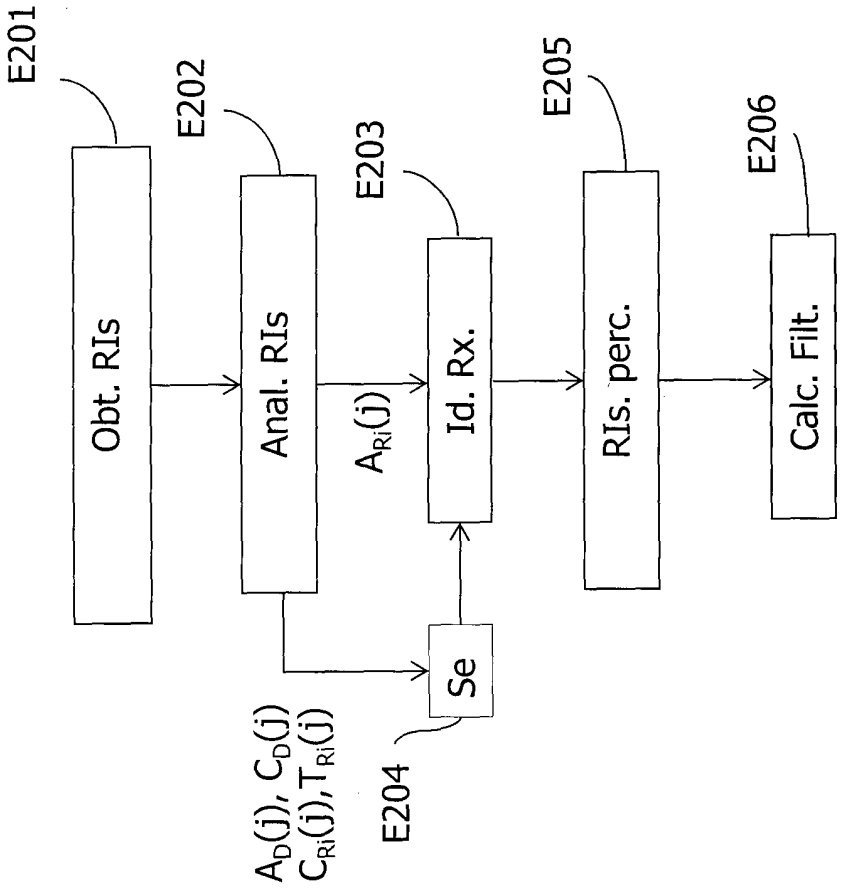


Fig.2

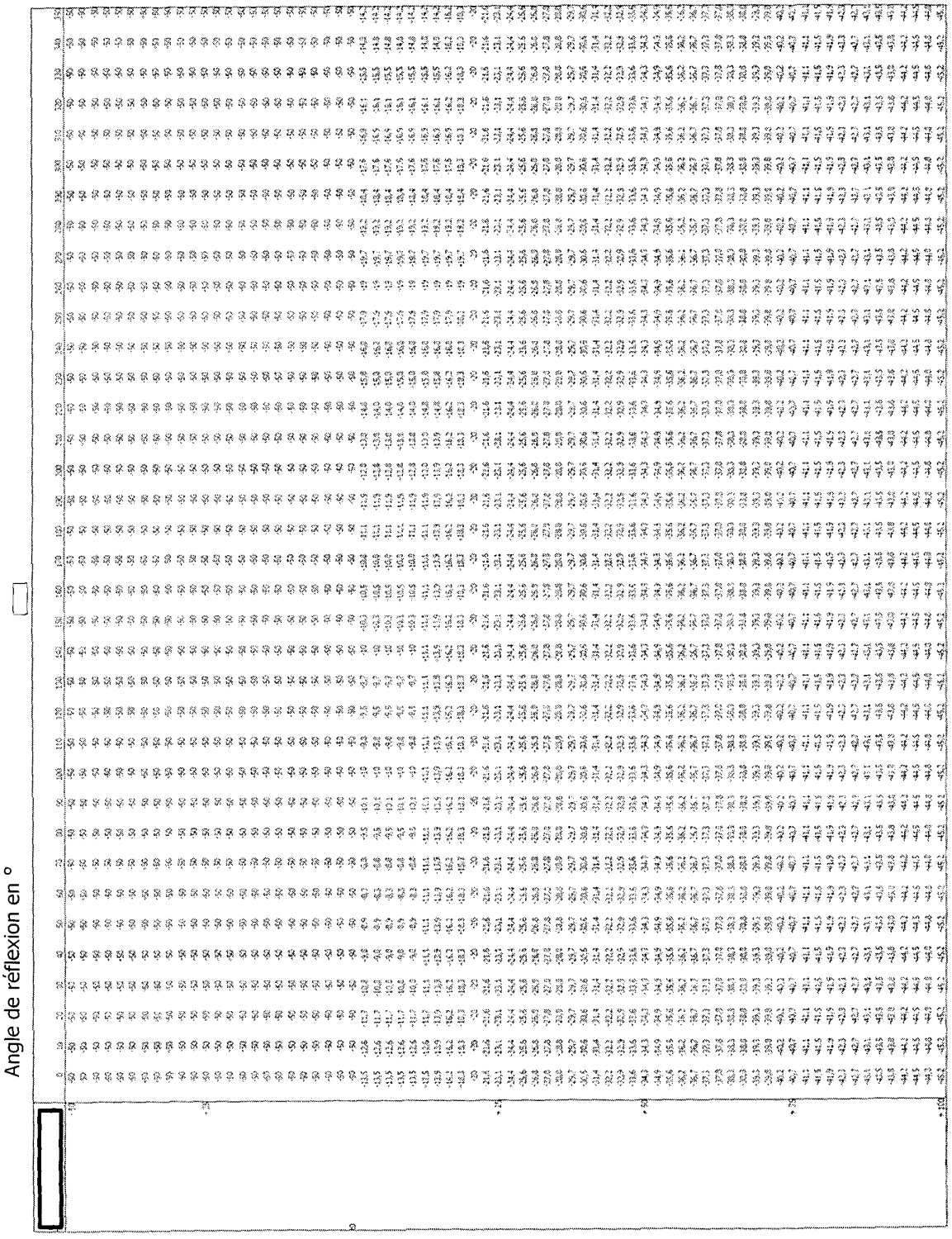


Fig.4

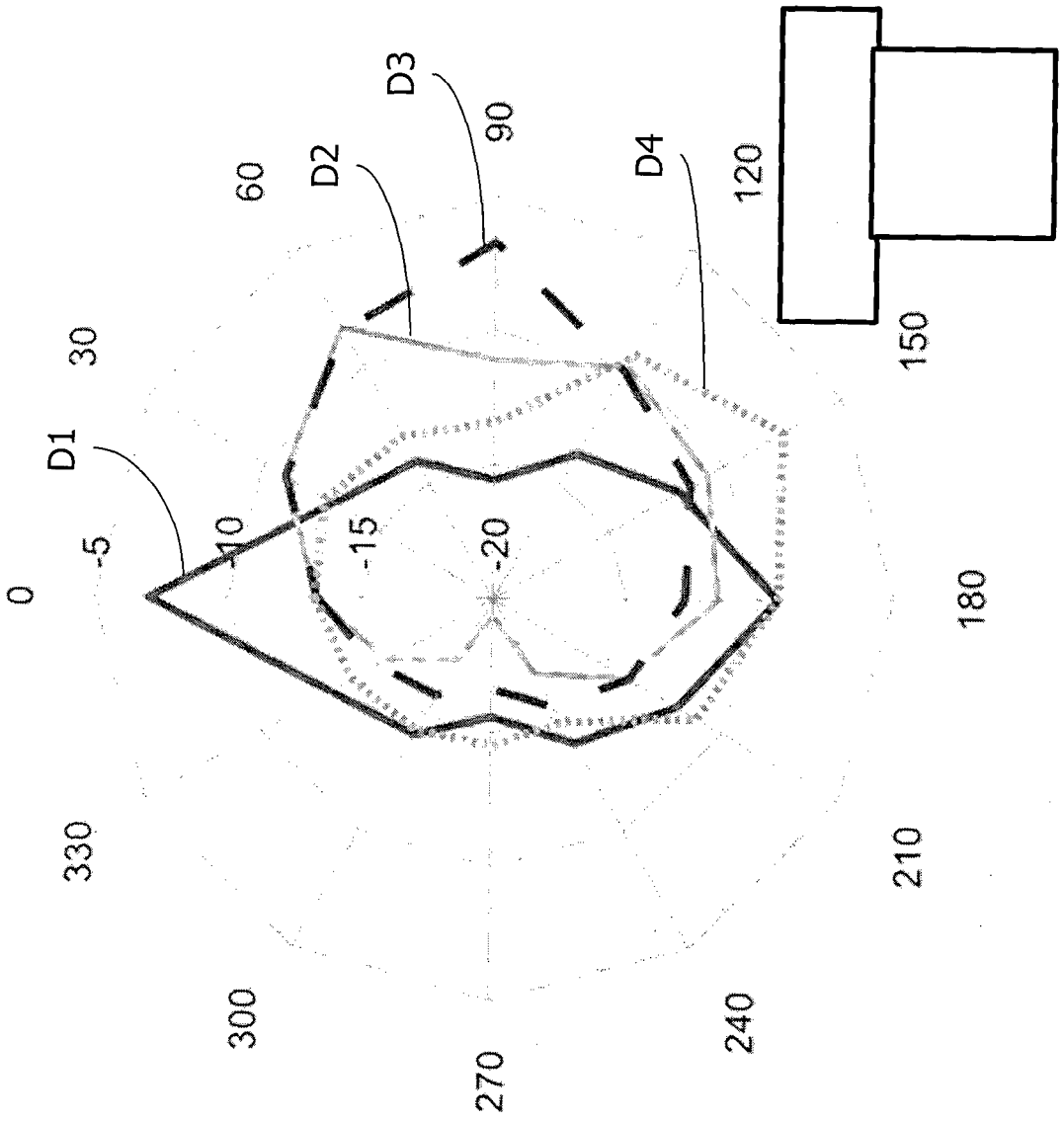


Fig.5

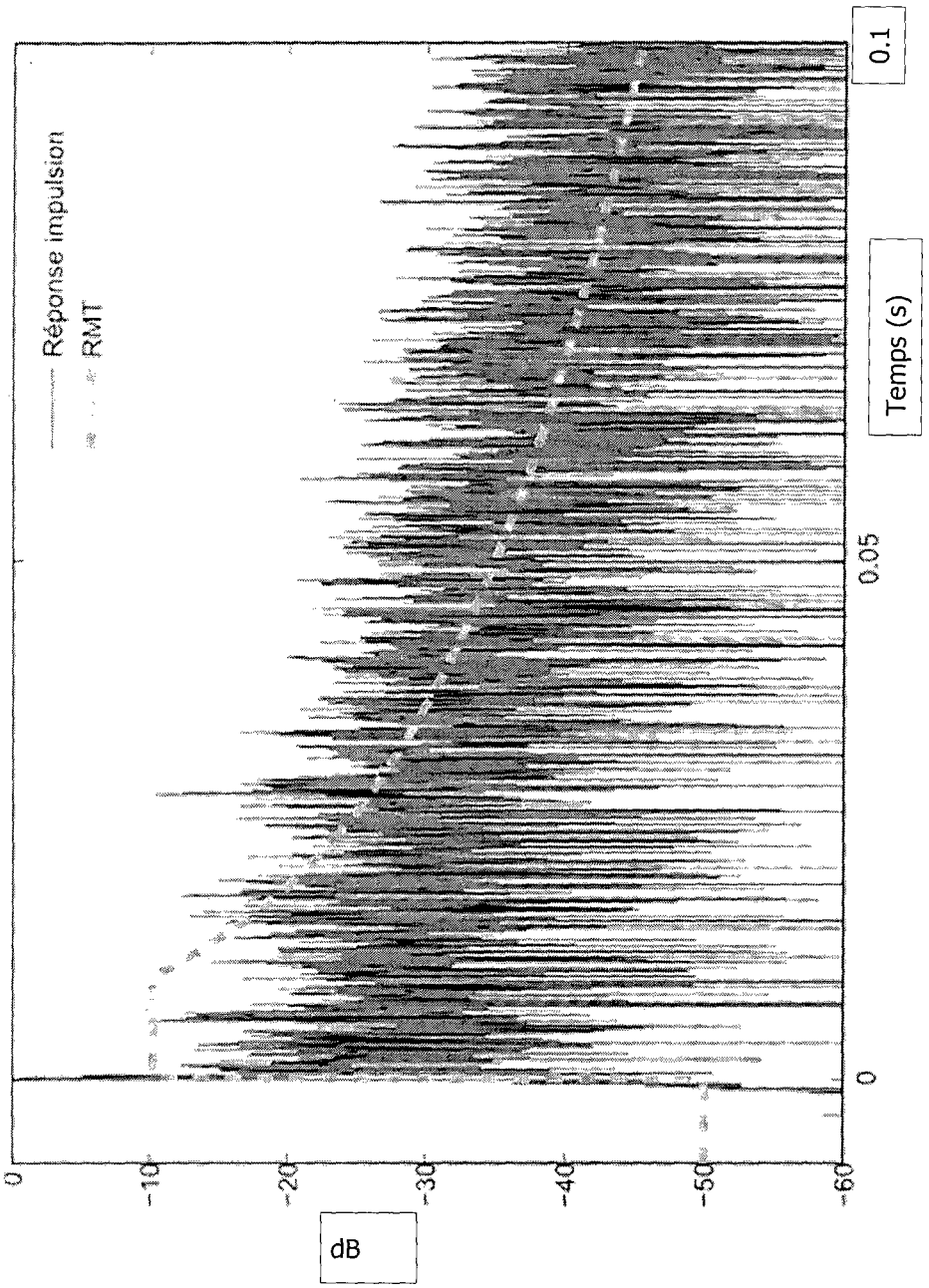


Fig.6

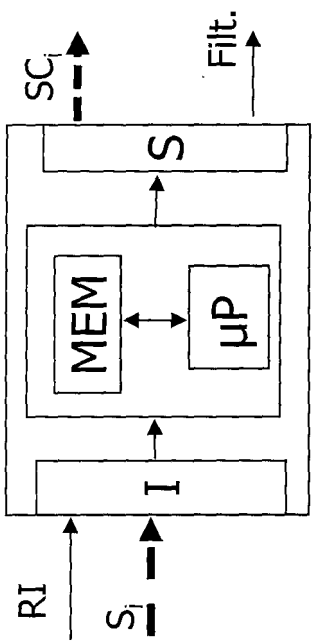


Fig.7

RÉFÉRENCES CITÉES DANS LA DESCRIPTION

Cette liste de références citées par le demandeur vise uniquement à aider le lecteur et ne fait pas partie du document de brevet européen. Même si le plus grand soin a été accordé à sa conception, des erreurs ou des omissions ne peuvent être exclues et l'OEB décline toute responsabilité à cet égard.

Littérature non-brevet citée dans la description

- **KIRKEBY ; NELSON.** Digital Filter Design for Inversion Problems in Sound Reproduction. *JAES,* 1999, vol. 7/8, 583-595 **[0005]**
- **ROMAIN DEPREZ et al.** *10ième congrès Français d'Acoustique,* 12 Avril 2010, 1-6 **[0011]**
- **S. MOREAU.** Etude et réalisation d'outils avancés d'encodage spatial pour la technique de spatialisation sonore Higher Order Ambisonics : microphone 3D et contrôle de la distance. *Univ. du Maine, PhD thesis,* 2006 **[0047]**
- **JÉRÔME DANIEL ; JEAN-BERNARD RAULT ; JEAN-DOMINIQUE POLACK.** Ambisonics encoding of other audio formats for multiple listening conditions. *AES 105th Convention,* Septembre 1998 **[0072]**
- **AMIR BECK ; MARC TEBOULLE.** A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. *SIAM J. IMAGING SCIENCES,* 2009, vol. 2 (1), 183-202 **[0097]**