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

(11) **EP 2 901 718 B1**

(12) **FASCICULE DE BREVET EUROPEEN**

(45) Date de publication et mention de la délivrance du brevet:
21.12.2016 Bulletin 2016/51

(21) Numéro de dépôt: **13779299.0**

(22) Date de dépôt: **25.09.2013**

(51) Int Cl.:
H04S 3/00 *(2006.01)*

(86) Numéro de dépôt international:
PCT/FR2013/052254

(87) Numéro de publication internationale:
WO 2014/049267 (03.04.2014 Gazette 2014/14)

(54) **PROCEDE ET SYSTEME DE RESTITUTION D'UN SIGNAL AUDIO**

VERFAHREN UND VORRICHTUNG ZUR WIEDERGABE EINES AUDIOSIGNALS

METHOD AND SYSTEM FOR PLAYING BACK AN AUDIO SIGNAL

(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: **27.09.2012 FR 1259132**

(43) Date de publication de la demande:
05.08.2015 Bulletin 2015/32

(73) Titulaire: **Sonic Emotion Labs**
75015 Paris (FR)

(72) Inventeurs:
- **NGUYEN, Khoa-Van**
  33800 BORDEAUX (FR)
- **CORTEEL, Etienne**
  F-92240 Malakoff (FR)

(74) Mandataire: **Fontenelle, Sandrine et al**
Cabinet Beau de Loménie
158, rue de l'Université
75340 Paris Cedex 07 (FR)

(56) Documents cités:
EP-A1- 2 485 504 **WO-A1-2012/025580**
US-A1- 2007 269 063

- **CORTEEL ET AL: "Equalization in an Extended Area Using Multichannel Inversion and Wave Field Synthesis", JAES, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, vol. 54, no. 12, 1 décembre 2006 (2006-12-01), pages 1140-1161, XP040507980, cité dans la demande**

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).

EP 2 901 718 B1

Printed by Jouve, 75001 PARIS (FR)

Description

Arrière-plan de l'invention

[0001] L'invention se rapporte au domaine général des traitements acoustiques et de la spatialisation sonore.

[0002] Elle concerne plus particulièrement la restitution d'un signal audio multicanal sur un dispositif de restitution déterminé, équipé d'une pluralité de haut-parleurs disposés à des emplacements fixes du dispositif de restitution.

[0003] L'invention s'applique de façon privilégiée mais non limitative à un dispositif de restitution de type enceinte acoustique, aussi appelée « structure bafflante » (ou « baffling structure » en anglais). Une telle enceinte acoustique est, de façon connue, constituée d'une structure unique ou monobloc, intégrant les différents haut-parleurs utilisés pour la restitution du signal audio (les haut-parleurs ne peuvent être séparés de l'enceinte). Un exemple d'enceinte acoustique est notamment une barre de son dans laquelle sont intégrés les différents haut-parleurs.

[0004] L'invention présente par ailleurs un intérêt particulier lorsqu'elle est appliquée à une enceinte acoustique dite compacte ou plus généralement à un dispositif de restitution compact.

[0005] De façon connue en soi, un dispositif de restitution compact est un dispositif de petites dimensions (notamment par rapport aux dimensions de la pièce ou de la salle dans laquelle on envisage de placer le dispositif de restitution), et dans lequel les haut-parleurs sont montés relativement proches les uns des autres.

[0006] Il convient de noter que ce dispositif peut être monobloc (comme une enceinte acoustique) ou en variante être composé de plusieurs éléments, regroupés de sorte à former un ensemble compact, chaque élément étant équipé d'un ou de plusieurs haut-parleurs.

[0007] A titre illustratif, la plus grande dimension d'un dispositif de restitution compact n'excède pas en général 2 mètres, tandis que l'espacement entre les haut-parleurs, deux à deux, est inférieur à 50 centimètres.

[0008] Il existe dans l'état de la technique, différentes méthodes s'employant à optimiser la restitution d'un signal audio multicanal sur un dispositif de restitution, tout en prenant en compte les limites physiques du dispositif de restitution, du fait notamment de la répartition des haut-parleurs du dispositif de restitution dans l'espace.

[0009] Un exemple d'une telle méthode est décrit dans le document WO 2012/025580 en référence à plusieurs dispositifs de restitution équipés d'une pluralité de haut-parleurs répartis à divers endroits d'une pièce de sorte à couvrir une zone (spatiale) d'écoute étendue (cette zone d'écoute modélise les positions des auditeurs).

[0010] Cette méthode s'appuie sur une analyse spatiale du signal audio multicanal que l'on souhaite restituer, permettant d'extraire et de localiser les objets sonores du signal audio situés à l'intérieur d'une fenêtre de restitution sonore définie à partir de la position physique des haut-parleurs du dispositif de restitution et de la zone d'écoute étendue.

[0011] Les objets sonores extraits sont restitués à l'intérieur de la fenêtre de restitution sonore, en fonction de leur localisation dans cette fenêtre, à l'aide d'un premier traitement de restitution. Ce premier traitement de restitution est par exemple un traitement de synthèse du champ acoustique (ou traitement WFS pour « Wave Field Synthesis » en anglais), connu en soi.

[0012] Les autres composantes du signal audio multicanal sont également restituées à l'intérieur de la fenêtre de restitution sonore, selon un second traitement de restitution (tel que par exemple, un effet panoramique d'intensité). Si le document WO 2012/025580 prend en compte, lors de l'analyse spatiale et lors de la restitution du signal audio multicanal, la répartition des haut-parleurs du dispositif de restitution par l'intermédiaire notamment de la notion de fenêtre de restitution sonore, il se restreint toutefois à l'étude de dispositifs de restitution ayant des haut-parleurs éparpillés dans l'ensemble de la pièce dans laquelle on souhaite restituer le signal, et destinés à une restitution dans une zone d'écoute étendue.

[0013] Mais le document WO 2012/025580 ne s'intéresse pas à proprement parler à la restitution d'un signal audio multicanal sur un dispositif de restitution compact.

[0014] Or, l'utilisation d'un dispositif de restitution compact présente certaines contraintes, notamment en termes de dimensions de la zone d'écoute pouvant être envisagée et de la fenêtre de restitution sonore liée à la disposition physique des haut-parleurs sur le dispositif de restitution, qui sont en général plus réduites qu'avec un dispositif de restitution composé de plusieurs entités éparpillées dans l'ensemble de la pièce ou de la salle dans laquelle est placé le dispositif, et tel qu'envisagé dans le document WO 2012/025580.

[0015] Il existe donc un besoin d'une méthode de restitution d'un signal audio multicanal qui soit particulièrement bien adaptée aux dispositifs de restitution compacts et notamment aux enceintes acoustiques compactes, et qui permette d'optimiser le rendu du signal audio tout en maintenant l'intelligibilité et la clarté des composantes de ce signal.

Objet et résumé de l'invention

[0016] L'invention répond notamment à ce besoin en proposant un procédé de restitution d'un signal audio multicanal sur un dispositif de restitution équipé d'une pluralité de haut-parleurs, ces haut-parleurs étant disposés à des emplacements fixes du dispositif de restitution et définissant une fenêtre spatiale de restitution sonore par rapport à une position

spatiale dite de référence. Le procédé de restitution selon l'invention est remarquable en ce qu'il comprend :

- une étape d'analyse spatiale du signal audio multicanal comprenant :
    - ◦ l'extraction d'au moins un objet sonore du signal, et
    - ◦ pour chaque objet sonore extrait, l'estimation d'un caractère diffus ou localisé de cet objet sonore, et d'une position de cet objet sonore par rapport à la fenêtre spatiale de restitution sonore du dispositif de restitution ; et

- une étape de restitution du signal audio sur la pluralité de haut-parleurs du dispositif de restitution, au cours de laquelle on applique, à chaque objet sonore extrait du signal audio, un traitement de restitution sur au moins un haut-parleur de la pluralité de haut-parleurs du dispositif de restitution, ce traitement de restitution dépendant du caractère diffus ou localisé de l'objet sonore et de sa position par rapport à la fenêtre spatiale de restitution sonore estimés au cours de l'étape d'analyse spatiale, le traitement de restitution comprenant la création d'au moins une source virtuelle à l'extérieur de la fenêtre spatiale de restitution du dispositif de restitution, à partir des haut-parleurs du dispositif de restitution, lorsque l'objet sonore est estimé au cours de l'étape d'analyse spatiale comme étant diffus ou positionné à l'extérieur de la fenêtre spatiale de restitution du dispositif de restitution.

[0017] Corrélativement, l'invention vise également un système de restitution d'un signal audio multicanal sur un dispositif de restitution équipé d'une pluralité de haut-parleurs, ces haut-parleurs étant disposés à des emplacements fixes du dispositif de restitution et définissant une fenêtre spatiale de restitution sonore par rapport à une position de référence, ce système de restitution comprenant :

- des moyens d'analyse spatiale du signal audio multicanal comprenant :
    - ◦ des moyens d'extraction d'au moins un objet sonore du signal, et
    - ◦ des moyens d'estimation, pour chaque objet sonore extrait, d'un caractère diffus ou localisé de cet objet sonore, et d'une position de cet objet sonore par rapport à la fenêtre spatiale de restitution sonore du dispositif de restitution ; et

- des moyens de restitution du signal audio sur la pluralité de haut-parleurs du dispositif de restitution, aptes à appliquer à chaque objet sonore extrait du signal audio, un traitement de restitution sur au moins un haut-parleur de la pluralité de haut-parleurs du dispositif de restitution, ce traitement de restitution dépendant du caractère diffus ou localisé de l'objet sonore et de sa position par rapport à la fenêtre spatiale de restitution sonore estimés au cours de l'étape d'analyse spatiale,

le traitement de restitution comprenant la création d'au moins une source virtuelle à l'extérieur de la fenêtre spatiale de restitution du dispositif de restitution, à partir des haut-parleurs du dispositif de restitution, lorsque l'objet sonore est estimé par les moyens d'analyse spatiale comme étant diffus ou positionné à l'extérieur de la fenêtre spatiale de restitution du dispositif de restitution.

[0018] Par étape (respectivement moyens) de restitution sur des haut-parleurs, on entend ici l'étape (respectivement les moyens) qui consiste à générer et à fournir des signaux destinés à alimenter les haut-parleurs du dispositif de restitution. Ces signaux seront ensuite diffusés (i.e. émis) par les haut-parleurs du dispositif de restitution de sorte à restituer le signal audio multicanal.

[0019] Par ailleurs, par position spatiale de référence, on entend ici aussi bien un point de l'espace caractérisant la position d'un auditeur cible du signal audio, qu'une zone plus étendue de l'espace dans laquelle est (sont) susceptible(s) de se trouver un ou plusieurs auditeurs. Pour un dispositif de restitution compact, on s'orientera préférentiellement vers une position spatiale de référence ponctuelle même si le procédé de restitution selon l'invention permet d'atteindre une zone d'écoute particulièrement étendue.

[0020] L'invention propose donc de mettre en oeuvre une analyse spatiale du signal audio multicanal à restituer visant à séparer les objets sonores composant le signal audio en fonction d'une part, de leur caractère localisé dans l'espace (c'est-à-dire discret, généré par une source localisable) ou diffus, et d'autre part, de leur position par rapport à la fenêtre de restitution sonore définie par la position spatiale de référence et par l'emplacement physique des haut-parleurs sur (ou dans) le dispositif de restitution par rapport à cette position spatiale de référence.

[0021] Cette séparation des objets sonores est mise à profit, conformément à l'invention, en appliquant des traitements de restitution aux objets extraits qui prennent en compte leurs caractères localisés ou diffus, ainsi que les positions des sources à l'origine de ces objets à l'intérieur ou à l'extérieur de la fenêtre de restitution sonore. Autrement dit, l'invention lie les traitements de restitution appliqués aux objets sonores du signal multicanal à restituer, directement aux caractéristiques spatiales de ces objets extraites lors de l'analyse spatiale du signal multicanal.

[0022] Plus précisément, les objets sonores identifiés au cours de l'étape d'analyse spatiale comme étant diffus ou positionnés à l'extérieur de la fenêtre spatiale de restitution du dispositif de restitution, sont avantageusement restitués par l'intermédiaire des haut-parleurs du dispositif de restitution, à l'extérieur de cette fenêtre, via la mise en oeuvre d'un traitement de restitution comprenant la création de sources virtuelles à l'extérieur de cette fenêtre.

[0023] En revanche, lorsqu'un objet sonore extrait est estimé comme étant localisé et positionné à l'intérieur de la fenêtre spatiale de restitution sonore du dispositif de restitution, le traitement de restitution appliqué à cet objet sonore au cours de l'étape de restitution est préférentiellement apte à restituer cet objet sonore à l'intérieur de la fenêtre spatiale de restitution sonore du dispositif de restitution, à l'emplacement de la source à l'origine de cet objet sonore.

[0024] Cette restitution à l'intérieur de la fenêtre spatiale de restitution sonore peut se faire de façon directe, en diffusant les objets sonores sur les haut-parleurs du dispositif de restitution sans recourir à des procédés complexes de filtrage spatial. Par exemple, on diffuse l'objet tel quel sur un ou plusieurs haut-parleurs, ou en appliquant simplement un effet panoramique (ou « panning » en anglais) d'intensité. De telles techniques sont connues en soi et relativement simples à mettre en oeuvre.

[0025] En variante, le traitement de restitution à l'intérieur de la fenêtre spatiale de restitution peut comprendre la création d'une ou de plusieurs sources virtuelles à partir des haut-parleurs du dispositif de restitution, à l'intérieur de la fenêtre spatiale de restitution sonore du dispositif de restitution. Il peut s'agir notamment d'un traitement de type WFS ou dérivé.

[0026] La direction ou la position des sources virtuelles, ainsi que, le cas échéant, leur amplitude, sont alors déterminées à partir de la position estimée des sources à l'origine des objets sonores localisés extraits du signal multicanal et de leur contribution (ex. en termes de niveau sonore) dans le signal multicanal.

[0027] Un tel traitement de restitution basé sur la création de sources virtuelles permet de mieux contrôler la directivité des objets sonores ainsi restitués.

[0028] L'application, au cours de l'étape de restitution, des traitements de restitution précités choisis en fonction des caractéristiques des objets sonores déterminées au cours de l'étape d'analyse spatiale, permet d'éloigner les objets diffus ou en provenance de l'extérieur de la fenêtre de restitution, des objets localisés à l'intérieur de la fenêtre (de tels objets incluent typiquement la voix ou les dialogues).

[0029] On augmente ainsi la largeur apparente de la scène sonore vue par l'auditeur (ou les auditeurs) situé(s) au niveau de la position spatiale de référence par rapport à la fenêtre de restitution sonore nominale offerte par le dispositif de restitution, fenêtre particulièrement limitée dans le cas d'un dispositif de restitution compact. Autrement dit, en dépit de la compacité du dispositif de restitution, l'auditeur a le sentiment d'être immergé dans la scène sonore (perception d'enveloppement dans la scène sonore).

[0030] Par ailleurs, outre cet élargissement de la scène sonore perçue par l'auditeur, on établit un contraste plus marqué entre les objets sonores localisés et situés à l'intérieur de la fenêtre de restitution sonore par rapport aux objets diffus ou localisés à l'extérieur de la fenêtre. Les objets localisés et déterminés comme étant positionnés à l'intérieur de la fenêtre de restitution sont, de ce fait, restitués avec une plus grande précision et une meilleure directivité. Le contraste établi par l'invention favorise par conséquent la clarté et l'intelligibilité de ces objets sonores pour l'auditeur à la position de référence.

[0031] En d'autres mots, l'invention tire profit d'un phénomène bien connu en psycho-acoustique sous le nom d'effet « cocktail party » ou « cocktail party effect » en anglais, qui reflète la capacité du système auditif humain à sélectionner une source sonore dans un environnement bruyant et à traiter des sons même s'ils ne sont pas au coeur de l'objet de l'attention humaine.

[0032] En liant les caractéristiques des objets sonores extraits du signal audio lors de l'analyse spatiale aux traitements de restitution appliqués au cours de l'étape de restitution pour restituer ces objets sur les haut-parleurs du dispositif de restitution, l'invention permet donc une restitution du signal audio multicanal de très bonne qualité, y compris sur un dispositif de restitution compact, tout en préservant la précision et la clarté des objets sonores du signal localisés et provenant de l'intérieur de la fenêtre de restitution. Elle peut s'appliquer à n'importe quel format de signal multicanal, comme par exemple à un signal stéréo, 5.1, 7.1, 10.2, HOA (Higher Order Ambisonics), etc.

[0033] Il convient de noter que le traitement réalisé de manière générale par l'invention ne vise pas en soi à modifier les caractéristiques de la scène sonore du signal audio multicanal, mais favorise l'intelligibilité des objets sonores localisés dans la fenêtre de restitution sonore et permet d'immerger l'auditeur dans la scène sonore.

[0034] Dans une variante de réalisation, l'étape d'analyse spatiale comprend en outre l'estimation de la position de l'objet sonore par rapport au centre de la fenêtre spatiale de restitution sonore du dispositif de restitution.

[0035] De cette sorte, on peut appliquer, au cours de l'étape de restitution, un traitement de restitution distinct selon si l'objet sonore se trouve au centre de la fenêtre spatiale de restitution sonore ou à une position distincte du centre mais dans la fenêtre spatiale de restitution sonore, de sorte à mieux isoler le centre des autres objets sonores. On obtient ainsi un meilleur contraste et une meilleure intelligibilité du centre par rapport aux autres objets situés à l'intérieur de la fenêtre. On notera que le centre est souvent associé aux objets sonores tels que la voix ou les dialogues.

[0036] Comme mentionné précédemment, l'invention a une application privilégiée, mais non limitative, lorsque le

dispositif de restitution est une enceinte acoustique dans laquelle est disposée la pluralité de haut-parleurs. Une telle enceinte acoustique est par exemple une barre son équipée de plusieurs haut-parleurs.

[0037] Dans un mode particulier de réalisation de l'invention, l'étape d'analyse spatiale comprend une décomposition du signal audio reçu en une pluralité de sous-bandes fréquentielles, l'extraction dudit au moins un objet sonore étant réalisée sur au moins une sous-bande fréquentielle.

[0038] Cette décomposition en sous-bandes fréquentielles (ex. en octave, en tiers d'octave ou en bandes auditives) facilite et améliore l'extraction des objets sonores constituant le signal audio. L'analyse spatiale du signal audio est en effet réalisée par sous-bande fréquentielle : on peut ainsi mieux isoler les objets sonores composant le signal audio multicanal. On a notamment la possibilité d'isoler plusieurs objets sonores dans le signal audio multicanal, par exemple un par sous-bande fréquentielle.

[0039] Selon une variante de réalisation de l'invention, le caractère diffus ou localisé de l'objet sonore extrait est estimé à partir d'au moins une corrélation évaluée entre deux canaux distincts du signal audio multicanal.

[0040] Par ailleurs, la position de l'objet sonore extrait par rapport à la fenêtre spatiale de restitution sonore peut être estimée à partir d'au moins une différence de niveaux évaluée entre deux canaux distincts du signal audio multicanal.

[0041] La détermination des caractéristiques associées à chaque objet sonore extrait du signal audio multicanal (i.e. caractère diffus ou localisé, position par rapport à la fenêtre de restitution) peut par conséquent être réalisée de manière très simple, par le biais de calcul de corrélations et de différences de niveaux entre les signaux répartis sur les différents canaux du signal multicanal.

[0042] Selon une autre variante de réalisation, l'étape d'analyse spatiale comprend la détermination d'un vecteur de Gerzon représentatif du signal audio multicanal.

[0043] De façon connue de l'homme du métier, le vecteur de Gerzon d'un signal audio multicanal est dérivé à partir des contributions respectives (direction et intensité ou énergie) des différents canaux du signal multicanal à la scène sonore perçue par l'auditeur à la position de référence. La détermination d'un tel vecteur pour un signal audio multicanal est décrite par exemple dans le document US 2007/0269063.

[0044] Le vecteur de Gerzon d'un signal audio multicanal reflète la localisation spatiale du signal audio multicanal telle qu'elle est perçue par l'auditeur depuis la position de référence. La détermination de ce vecteur de Gerzon permet de s'affranchir du calcul de corrélations entre les différents canaux du signal multicanal pour déterminer le caractère diffus ou localisé des objets sonores extraits du signal.

[0045] Selon une autre variante de réalisation, l'étape d'analyse spatiale comprend une décomposition spatiale du signal multicanal en harmoniques sphériques.

[0046] Une telle décomposition spatiale est connue de l'homme du métier et décrite par exemple dans le document WO 2012/025580. Elle permet une analyse spatiale très précise du signal audio multicanal et des objets sonores le composant. Ainsi, notamment, plusieurs objets sonores peuvent être déterminés pour une même sous-bande fréquentielle.

[0047] Différents traitements peuvent être envisagés dans le cadre de l'invention pour la restitution des objets sonores extraits lors de l'analyse spatiale, à l'intérieur ou à l'extérieur de la fenêtre spatiale de restitution.

[0048] Ainsi, selon une première variante de réalisation de l'invention, dans laquelle la pluralité de haut-parleurs du dispositif de restitution comprend un haut-parleur central et des haut-parleurs latéraux, lorsque l'objet sonore extrait est estimé au cours de l'étape d'analyse spatiale comme étant diffus ou positionné à l'extérieur de la fenêtre spatiale de restitution du dispositif de restitution, le traitement de restitution appliqué à cet objet sonore utilise une technique transaurale de restitution de cet objet sonore sur les haut-parleurs latéraux du dispositif de restitution.

[0049] Cette première variante de réalisation a une application privilégiée dans le cas d'un dispositif de restitution équipé d'un nombre réduit de haut-parleurs, par exemple un haut-parleur central et deux haut-parleurs latéraux.

[0050] Selon une deuxième variante de réalisation de l'invention dans laquelle la pluralité de haut-parleurs du dispositif de restitution comprend un haut-parleur central et des haut-parleurs latéraux, lorsqu'un objet sonore extrait est estimé au cours de l'étape d'analyse spatiale comme étant localisé et positionné au centre de la fenêtre spatiale de restitution du dispositif de restitution, cet objet sonore est diffusé, au cours de l'étape de restitution, par le traitement de restitution, sur le haut-parleur central du dispositif de restitution.

[0051] Autrement dit, un objet sonore centré par rapport à la position spatiale de référence est rattaché au centre du dispositif de restitution de sorte à optimiser son intelligibilité. Il est préférentiellement restitué de façon directe (c'est-à-dire sans filtrage spatial) sur le haut-parleur central du dispositif de restitution, de sorte à bénéficier des propriétés de directivité naturelles du haut-parleur central.

[0052] D'autres techniques de restitution d'un objet sonore centré par rapport à la position spatiale de référence peuvent bien entendu être envisagées pour maximiser son intelligibilité. Ainsi, par exemple, on peut envisager la formation d'un faisceau (aussi connu sous le nom de « beamforming ») dirigé vers la position spatiale de référence ou une technique transaurale.

[0053] Selon une troisième variante de réalisation, lorsqu'un objet sonore extrait est estimé au cours de l'étape d'analyse spatiale comme étant localisé et positionné à l'intérieur de la fenêtre spatiale de restitution du dispositif de restitution

en une position distincte du centre de la fenêtre, le traitement de restitution appliqué au cours de l'étape de restitution diffuse cet objet sonore sur les haut-parleurs du dispositif de restitution en utilisant un effet panoramique d'intensité.

[0054] Ainsi, les objets sonores localisés et positionnés à l'intérieur de la fenêtre acoustique sont également rattachés au dispositif de restitution, et restitués de manière directe (c'est-à-dire sans filtrage spatial), à l'intérieur de la fenêtre de restitution par le biais de l'effet panoramique d'intensité appliqué sur les haut-parleurs. Cet effet panoramique d'intensité appliqué sur l'ensemble des haut-parleurs du dispositif de restitution permet de mieux distinguer les objets sonores localisés et positionnés à l'intérieur de la fenêtre acoustique des objets sonores situés au centre de la fenêtre.

[0055] L'invention ne se limite toutefois pas à l'application des traitements de restitution précités ; on peut également recourir à des traitements de restitution plus complexes, mettant en oeuvre notamment un filtrage spatial des objets sonores sur les haut-parleurs du dispositif de restitution.

[0056] Ainsi, par exemple, lorsque l'objet sonore extrait est estimé au cours de l'étape d'analyse spatiale comme étant positionné à l'extérieur de la fenêtre spatiale de restitution du dispositif de restitution, la création d'au moins une source virtuelle à l'extérieur de la fenêtre spatiale de restitution du dispositif de restitution peut comprendre la formation d'au moins un faisceau dirigé vers l'extérieur de la fenêtre spatiale de restitution (« beamforming »).

[0057] De façon similaire, lorsqu'un objet sonore extrait est estimé au cours de l'étape d'analyse spatiale comme étant localisé et positionné à l'intérieur de la fenêtre spatiale de restitution du dispositif de restitution, le traitement de restitution appliqué à cet objet sonore au cours de l'étape de restitution, peut comprendre la formation d'un faisceau dirigé vers la position spatiale de référence.

[0058] De manière générale, la création de sources virtuelles permet un meilleur contrôle et une meilleure précision de la restitution sonore d'un signal audio qu'une restitution sonore « directe » (i.e. sans filtrage spatial) sur les haut-parleurs du dispositif de restitution, limitée quant à elle par la seule capacité des haut-parleurs du dispositif de restitution. Elle offre en effet la possibilité d'avoir un meilleur contrôle de la directivité des sources sonores reconstituées.

[0059] Par ailleurs, la formation d'un faisceau (« beamforming ») pour créer une source virtuelle à l'intérieur ou à l'extérieur de la fenêtre de restitution, permet de contrôler aisément la largeur de la source virtuelle ainsi créée. Le « beamforming » est particulièrement bien adapté pour la restitution de signaux sur des réseaux de haut-parleurs denses (ex. dispositif de restitution équipé de 6 haut-parleurs ou plus), pour lesquels on dispose d'une meilleure précision pour créer les sources virtuelles du fait de l'existence d'un nombre plus important de degrés de liberté (lié à la présence d'un nombre de haut-parleurs plus important).

[0060] On peut en outre, lors de la restitution des objets sonores, interagir plus facilement via l'utilisation de techniques de « beamforming » avec les dimensions de la pièce ou de la salle dans laquelle est placé le dispositif de restitution. Ainsi par exemple, lorsque le faisceau est dirigé vers l'extérieur de la fenêtre de restitution, on peut, en jouant sur la largeur du faisceau, élargir la surface réfléchie par les murs de la pièce et créer pour l'auditeur une meilleure sensation d'enveloppement dans la scène sonore.

[0061] Dans un mode particulier de réalisation, les différentes étapes du procédé de restitution sont déterminées par des instructions de programmes d'ordinateur.

[0062] En conséquence, l'invention vise aussi un programme sur un support d'informations, ce programme étant susceptible d'être mis en oeuvre dans un système de restitution ou plus généralement dans un ordinateur, ce programme comportant des instructions adaptées à la mise en oeuvre des étapes d'un procédé de restitution tel que décrit ci-dessus.

[0063] Ce programme peut utiliser n'importe quel langage de programmation, et être sous la forme de code source, code objet, ou de code intermédiaire entre code source et code objet, tel que dans une forme partiellement compilée, ou dans n'importe quelle autre forme souhaitable.

[0064] L'invention vise aussi un support d'informations lisible par un ordinateur ou par un microprocesseur, et comportant des instructions d'un programme tel que mentionné ci-dessus.

[0065] Le support d'informations peut être n'importe quelle entité ou dispositif capable de stocker le programme. Par exemple, le support peut comporter un moyen de stockage, tel qu'une ROM, par exemple un CD ROM ou une ROM de circuit microélectronique, ou encore un moyen d'enregistrement magnétique, par exemple une disquette (floppy disc) ou un disque dur.

[0066] D'autre part, le support d'informations peut être un support transmissible tel qu'un signal électrique ou optique, qui peut être acheminé via un câble électrique ou optique, par radio ou par d'autres moyens. Le programme selon l'invention peut être en particulier téléchargé sur un réseau de type Internet.

[0067] Alternativement, le support d'informations peut être un circuit intégré dans lequel le programme est incorporé, le circuit étant adapté pour exécuter ou pour être utilisé dans l'exécution du procédé en question.

[0068] Selon un autre aspect, l'invention vise également une enceinte acoustique comprenant un système de restitution conforme à l'invention.

[0069] On peut également envisager, dans d'autres modes de réalisation, que le procédé, le système de restitution et l'enceinte acoustique selon l'invention présentent en combinaison tout ou partie des caractéristiques précitées.

Brève description des dessins

[0070] D'autres caractéristiques et avantages de la présente invention ressortiront de la description faite ci-dessous, en référence aux dessins annexés qui en illustrent des exemples de réalisation dépourvus de tout caractère limitatif.

[0071] Sur les figures :

- la figure 1 représente un système de restitution conforme à l'invention, dans un mode particulier de réalisation ;
- les figures 2, 3A et 3B illustrent des exemples de fenêtres spatiales de restitution sonore pour divers dispositifs de restitution et positions de référence ;
- la figure 4 représente schématiquement l'architecture matérielle du système de restitution de la figure 1 ; et
- la figure 5 représente les principales étapes d'un procédé de restitution selon l'invention, telles qu'elles sont mises en oeuvre, dans un mode particulier de réalisation, par le système de restitution de la figure 1.

Description détaillée de l'invention

[0072] La **figure 1** représente, dans son environnement, un système de restitution 1 d'un signal audio S multicanal sur un dispositif 2 de restitution, conforme à l'invention, dans un mode particulier de réalisation.

[0073] Le dispositif 2 de restitution est équipé d'une pluralité de haut-parleurs 2-1, 2-2, ..., 2-N (N>1). Il s'agit, dans l'exemple représenté sur la figure 1, d'un dispositif de restitution compact.

[0074] Plus précisément, le dispositif 2 de restitution est ici une enceinte acoustique compacte, autrement dit une structure monobloc ou boîte fermée unique, intégrant l'ensemble des haut-parleurs 2-1, 2-2, ..., 2-N.

[0075] Le dispositif de restitution 2 est par exemple une barre de son montée horizontale, de longueur n'excédant pas un ou deux mètres, à l'intérieur de (ou sur) laquelle sont disposés, à des emplacements fixes et rapprochés les uns des autres (à moins de 50cm les uns des autres), les haut-parleurs 2-1, 2-2,..., 2-N.

[0076] Ces hypothèses ne sont toutefois pas limitatives, et l'invention s'applique également à d'autres types de dispositifs de restitution. Ainsi, notamment l'invention s'applique aussi à un dispositif de restitution compact modulaire constitué de plusieurs éléments séparés intégrant chacun un ou plusieurs haut-parleurs.

[0077] Il convient de noter que la notion de dispositif de restitution compact est connue de l'homme du métier : un dispositif de restitution compact désigne en effet un dispositif de petites dimensions, notamment par rapport aux dimensions de la pièce ou de la salle dans laquelle on envisage la restitution du signal audio à l'aide de ce dispositif, et sur ou dans lequel les haut-parleurs sont montés relativement proches les uns des autres. A titre illustratif, la plus grande dimension d'un dispositif de restitution compact n'excède pas en général 2 mètres, tandis que les haut-parleurs sont montés sur le dispositif de restitution avec un espacement inférieur à 50 cm.

[0078] L'emplacement physique des haut-parleurs 2-1, 2-2, ..., 2-N définit, de façon connue, une fenêtre spatiale W de restitution sonore par rapport à une position dite de référence notée Pref, placée devant le dispositif de restitution 2 (par rapport notamment à l'orientation de tout ou partie des haut-parleurs et à la diffusion des sons), et modélisant la position d'un auditeur dans l'espace pris comme référence pour optimiser la restitution du signal audio S.

[0079] Le choix à proprement parler de la position de référence Pref dépend de plusieurs facteurs connus de l'homme du métier, et ne sera pas décrit ici. Pour un dispositif de restitution compact, cette position de référence Pref est choisie généralement ponctuelle.

[0080] La **figure 2** illustre la fenêtre spatiale W de restitution sonore définie par les haut-parleurs 2-1, 2-2, ..., 2-N du dispositif de restitution 2 et la position de référence Pref.

[0081] De façon connue, l'emplacement physique des haut-parleurs 2-1, 2-2, ..., 2-N sur le dispositif de restitution 2 (et plus précisément des deux haut-parleurs 2-1 et 2-N situés aux extrémités du dispositif de restitution 2), associé à la position de référence Pref, définissent une ouverture angulaire Ω de restitution sonore.

[0082] Le sous-espace délimité par cette ouverture angulaire Ω correspond à la fenêtre spatiale W de restitution sonore associée au dispositif de restitution 2 et à la position de référence Pref.

[0083] Il convient de noter que :

- la fenêtre W dépend de la position de référence Pref. Dans l'exemple de la figure 2, la position Pref est alignée par rapport au centre du dispositif de restitution 2, de sorte que la fenêtre spatiale W est définie par l'excursion angulaire $\Omega/2$ par rapport à l'axe Δ reliant le centre du dispositif 2 de restitution à la position de référence Pref ; et

- seul l'emplacement physique des haut-parleurs du dispositif de restitution 2 (et notamment des haut-parleurs situés aux extrémités du dispositif de restitution 2) par rapport à la position Pref est prise en compte dans la notion de fenêtre spatiale de restitution sonore. On ne s'attache pas à la puissance des haut-parleurs du dispositif de restitution 2 ou à d'autres caractéristiques qui pourraient influencer leur capacité à restituer un signal audio.

[0084] Les **figures 3A et 3B** illustrent respectivement, à titre d'exemples :

- la fenêtre spatiale W' de restitution sonore d'un dispositif de restitution 2' de type barre de son montée horizontale, muni de trois haut-parleurs 2-1', 2-2', 2-3' par rapport à une position spatiale Pref' de référence étendue ; et
- la fenêtre spatiale W" de restitution sonore d'un dispositif de restitution 2" muni de 8 haut-parleurs 2-1", 2-2", ..., 2-8" par rapport à une position spatiale Pref" de référence ponctuelle, les haut-parleurs 2-1" à 2-4" étant frontaux tandis que les haut-parleurs 2-5", 2-6" et 2-7", 2-8" sont disposés de chaque côté du dispositif de restitution 2".

[0085] Comme mentionné précédemment, l'invention propose un traitement d'un signal audio multicanal en deux temps : dans un premier temps, le signal audio multicanal à restituer est analysé spatialement ; puis, les caractéristiques spatiales du signal résultant de cette analyse spatiale sont utilisées pour optimiser la restitution du signal sur le dispositif de restitution 2.

[0086] Ainsi, le système 1 de restitution selon l'invention comprend :

- des moyens 3 d'analyse spatiale du signal audio S multicanal comprenant notamment des moyens d'extraction d'au moins un objet sonore du signal, et d'estimation, pour chaque objet sonore extrait, d'un caractère diffus ou localisé de cet objet sonore, et d'une position de cet objet sonore par rapport à la fenêtre spatiale W de restitution sonore du dispositif de restitution 2 (l'extraction des objets sonores et l'estimation de leurs caractéristiques sont en général réalisées conjointement) ; et
- des moyens 4 de restitution du signal S audio sur la pluralité de haut-parleurs 2-1, ..., 2-N du dispositif de restitution 2, aptes à appliquer à chaque objet sonore extrait du signal audio, un traitement de restitution sur au moins un haut-parleur de la pluralité de haut-parleurs du dispositif de restitution 2-1, ..., 2-N, ce traitement de restitution dépendant du caractère diffus ou localisé de l'objet sonore et de sa position par rapport à la fenêtre spatiale de restitution sonore estimés au cours de l'étape d'analyse spatiale.

[0087] Plus précisément, dans l'exemple envisagé ici, les moyens 4 de restitution sont aptes à appliquer les traitements de restitution T-A1, T-A2, T-B et T-C, sur les objets sonores extraits du signal S, en fonction des caractéristiques déterminées par les moyens 3 d'analyse spatiale. Aucune limitation n'est toutefois attachée au nombre de traitements différents pouvant être appliqués par le système 1 de restitution.

[0088] Il convient de noter que les traitements T-A1, T-A2, T-B et T-C, bien que dépendant des caractéristiques des objets sonores extraits, peuvent être de même nature (c'est-à-dire basés sur les mêmes techniques, comme par exemple une technique WFS ou « beamforming »). Toutefois, ils sont adaptés aux caractéristiques spatiales des objets sonores auxquels ils sont appliqués et diffèrent en ce sens les uns des autres. Ainsi par exemple, ils ne diffusent pas les signaux sur les mêmes haut-parleurs, n'envisagent pas la création de sources virtuelles dans les mêmes sous-espaces (ou ayant des caractéristiques similaires en termes de position/direction et/ou d'amplitude), les faisceaux créés peuvent être dimensionnés différemment (ex. de largeurs différentes), etc.

[0089] Ainsi, les moyens 4 de restitution comprennent ici :

- des moyens de traitement 4A aptes à appliquer un ou plusieurs traitements de restitution sur les objets sonores du signal audio S déterminés comme étant localisés et dans la fenêtre W de restitution sonore. Dans l'exemple envisagé à la figure 1, les moyens de traitement 4A sont aptes à appliquer un traitement T-A1 sur les objets sonores générés par des sources placées au centre de la fenêtre W, et un traitement T-A2 sur les objets sonores placés à l'intérieur de la fenêtre W en une position distincte du centre ;
- des moyens de traitement 4B aptes à appliquer un traitement T-B sur les objets sonores du signal audio S déterminés comme diffus ; et
- des moyens de traitement 4C aptes à appliquer un traitement T-C sur les objets sonores du signal audio S déterminés comme localisés et à l'extérieur de la fenêtre W de restitution sonore.

[0090] Les traitements de restitution T-A1, T-A2, T-B et T-C seront décrits plus en détails ultérieurement et illustrés par des exemples.

[0091] Dans le mode de réalisation décrit ici, les moyens 3 d'analyse spatiale et les moyens 4 de restitution du signal audio sont des moyens logiciels.

[0092] Plus précisément, dans le mode de réalisation décrit ici, le système de restitution 1 dispose de l'architecture matérielle d'un ordinateur, telle qu'illustrée à la **figure 4.**

[0093] Il comporte notamment un processeur (ou micro-processeur) 5, une mémoire vive 6, une mémoire morte 7, une mémoire flash non volatile 8 ainsi que des moyens de communication 9 aptes à émettre et à recevoir des signaux.

[0094] Ainsi, les moyens de communication 9 comprennent d'une part, une interface (filaire ou sans fil) avec les haut-parleurs 2-1, ..., 2-N du dispositif de restitution 2, ainsi que des moyens de réception d'un signal audio multicanal, tel

que le signal S par exemple. Ces moyens sont connus de l'homme du métier et ne seront pas décrits davantage ici.

[0095] La mémoire morte 7 du système 1 de restitution constitue un support d'enregistrement conforme à l'invention, lisible par le (micro-)processeur 5 et sur lequel est enregistré un programme d'ordinateur conforme à l'invention, comportant des instructions pour l'exécution des étapes d'un procédé de restitution décrites ultérieurement en référence à la figure 5.

[0096] Il convient de noter qu'aucune limitation n'est attachée à la nature à proprement parler du système 1 de restitution. Ainsi, notamment, le système 1 de restitution peut se présenter sous la forme d'un ordinateur ou en variante d'une puce électronique ou d'un circuit intégré, dans lequel le programme d'ordinateur comprenant les instructions pour l'exécution du procédé de restitution selon l'invention est incorporé.

[0097] Par ailleurs, le système 1 de restitution peut être une entité distincte du dispositif 2 de restitution, ou au contraire, être intégré à l'intérieur du dispositif 2 de restitution.

[0098] Nous allons maintenant décrire, en référence à la **figure 5**, les différentes étapes du procédé de restitution selon l'invention, dans un mode particulier de réalisation dans lequel il est mis en oeuvre par le système 1 de restitution pour la restitution du signal audio S multicanal sur les haut-parleurs 2-1, ..., 2-N du dispositif 2 de restitution.

[0099] On suppose que le signal audio S multicanal est fourni au système 1 de restitution par l'intermédiaire de ses moyens de communication 9. Le format et la structure d'un tel signal audio est connue de l'homme du métier et ne sera pas décrite ici.

[0100] Sur réception du signal S (étape E10), le système 1 de restitution initie une première phase $\sum$I d'analyse spatiale du signal S réalisée à l'aide de ses moyens 3 d'analyse spatiale.

[0101] Plus précisément, dans le mode de réalisation décrit ici, au cours de cette première phase $\sum$I, le système 1 de restitution décompose le signal S multicanal en une pluralité K de sous-bandes fréquentielles désignées par BW1, ..., BWK (étape E20), chaque sous-bande fréquentielle BWi, i=1, ..., K intégrant les différents canaux composant le signal S. Autrement dit, le signal noté Si résultant de la décomposition du signal S et associé à la sous-bande fréquentielle BWi est lui-même un signal multicanal.

[0102] Aucune limitation n'est attachée à la largeur de chaque sous-bande : on peut par exemple envisager une décomposition en octave, en tiers d'octave, ou encore en bandes auditives (i.e. adaptées à l'audition), en fonction d'un compromis complexité/précision notamment.

[0103] La décomposition en sous-bandes fréquentielles du signal S est effectuée par l'intermédiaire d'une transformation de Fourier appliquée au signal S, et ne présente pas de difficulté en soi pour l'homme du métier.

[0104] Suite à cette décomposition, le système 1 de restitution analyse les signaux Si, i=1, ..., K associés à chaque sous-bande fréquentielle BWi, i=1, ..., K.

[0105] Durant cette analyse, pour chaque sous-bande fréquentielle BWi, il extrait les objets sonores compris dans le signal Si (c'est-à-dire de manière équivalente les sons ou les éléments sonores présents dans le signal Si), et estime, pour chaque objet sonore extrait (étape E30) :

- s'il s'agit d'un objet localisé (l'objet est créé par une source localisée et identifiable dans l'espace) ou diffus (i.e. l'objet ne vient pas d'une source localisable, mais semble provenir de tout autour l'auditeur) ; et
- lorsqu'il s'agit d'un objet localisé, sa position (i.e. la position de la source à l'origine de cet objet) par rapport à la fenêtre spatiale W de restitution sonore.

[0106] Dans le mode de réalisation décrit ici, les amplitudes des objets sonores extraits sont contenues directement dans les signaux Si, et correspondent respectivement aux niveaux des sous-bandes fréquentielles.

[0107] L'extraction des objets sonores et l'estimation des caractéristiques précitées de chaque objet (localisé/diffus, position par rapport à la fenêtre spatiale W) sont réalisées conjointement par les moyens 3 d'analyse spatiale.

[0108] Différentes techniques peuvent être utilisées à cette fin par les moyens 3 du système 1 de restitution.

[0109] Ainsi, selon une première variante de réalisation de l'invention, les moyens 3 d'analyse spatiale du système 1 de restitution mettent en oeuvre une analyse temporelle du signal multicanal Si.

[0110] Lors de cette analyse temporelle, le système 1 de restitution évalue, pour chaque paire de canaux distincts du signal multicanal Si, la corrélation normalisée entre ces canaux (i.e. entre les signaux représentatifs des canaux), définie par l'équation suivante :

$$R_{x,y}(p) = \begin{cases} \frac{1}{M} \sum_{m=0}^{M-p-1} x(m+p)y^*(m) \quad pour\ p \geq 0 \\ R^*_{x,y}(-p) pour\ p < 0 \end{cases}$$

où x et y désignent respectivement deux canaux distincts du signal multicanal Si, [.]* désigne l'opérateur de conjugaison

complexe, et M est une constante définissant le nombre d'échantillons de signal sur lequel la corrélation est évaluée.

[0111] De façon alternative, lors de l'analyse temporelle, le système 1 de restitution peut se contenter d'évaluer une corrélation normalisée entre deux canaux distincts du signal multicanal Si pour une sélection seulement de paires de canaux prédéterminés du signal Si.

[0112] Par exemple, pour un signal multicanal de format 5.1, composé d'un centre à 0°, de canaux gauche L et droite R situés à +/- 30° par rapport au centre, et de canaux arrière gauche Ls et arrière droit Rs situés à +/-110° par rapport au centre, cette sélection peut comprendre uniquement quatre paires de canaux, à savoir, la paire constituée des canaux L et R, la paire constituée des canaux Ls et Rs, la paire constituée des canaux L et Ls et la paire constituée des canaux R et Rs.

[0113] Chaque corrélation $R_{x,y}$ ainsi évaluée est ensuite comparée à un seuil prédéfini noté THR.

[0114] Si la corrélation est supérieure au seuil THR, le système 1 de restitution estime que le signal Si (et donc a fortiori le signal S) contient un objet sonore localisé.

[0115] Au contraire, si la corrélation est inférieure au seuil THR, le système 1 de restitution estime que le signal Si contient un objet sonore diffus.

[0116] La valeur du seuil THR est déterminée de façon empirique : elle est choisie préférentiellement entre 0.5 et 0.8.

[0117] On peut donc ainsi extraire autant d'objets sonores du signal Si que de paires de canaux examinées ou de façon équivalente, que de corrélations évaluées entre les canaux du signal Si.

[0118] Lorsqu'un objet sonore est estimé comme localisé par le système 1 de restitution, celui-ci estime la position de cet objet sonore par rapport à la fenêtre spatiale W de restitution sonore (par définition, un objet diffus n'a pas de position précise ou identifiable dans l'espace. Il n'est donc pas nécessaire d'estimer sa position par rapport à la fenêtre W spatiale de restitution).

[0119] A cet effet, le système 1 de restitution estime ici la fenêtre spatiale W de restitution à partir de la position de référence Pref et des emplacements physiques des haut-parleurs du dispositif 2 de restitution.

[0120] La fenêtre spatiale W peut être déterminée géométriquement par le système 1 de restitution, en termes d'excursion angulaire par rapport à l'axe Δ passant par le centre du dispositif 2 de restitution et la position de référence Pref, à partir de la connaissance de la position Pref et des emplacements physiques des haut-parleurs du dispositif 2 placés aux extrémités (i.e. 2-1 et 2-N). Dans l'exemple représenté à la figure 2, la fenêtre spatiale W est associée par le système 2 de restitution à une excursion angulaire de $\Omega/2$ par rapport à l'axe Δ.

[0121] La position Pref et les emplacements physiques des haut-parleurs du dispositif peuvent être préalablement configurés dans la mémoire flash 7 non volatile du système 1 de restitution, par exemple lors de la construction du système 1 de restitution si celui-ci est intégré dans le dispositif 2 ou lors d'une étape préalable de configuration du système 1 de restitution.

[0122] En variante, la fenêtre W peut être estimée par le système 1 de restitution à l'aide d'une technique similaire ou identique à celle décrite dans le document d'E. Corteel intitulé « Equalization in extended area using multichannel inversion and wave field synthesis », Journal of the Audio Engineering Society n°54(12), décembre 2006, lorsque la position Pref est une zone étendue.

[0123] D'autres techniques connues de l'homme du métier peuvent bien entendu être utilisées en variante des deux techniques précitées. Par ailleurs, dans une autre variante encore, la fenêtre spatiale W peut être prédéterminée, et stockée par exemple dans la mémoire flash 7 non volatile du système 1 de restitution.

[0124] Le système 1 de restitution évalue également, pour chaque paire de canaux distincts du signal Si, la différence de niveaux (ou d'énergie), entre ces canaux, par exemple en décibels, selon l'équation suivante :

$$10 \log_{10} \frac{\sum_{p=p0}^{P} \|x\|^2(p)}{\sum_{p=p0}^{P} \|y\|^2(p)}$$

où x et y désignent respectivement deux canaux distincts du signal multicanal Si, $||x||$ désigne la norme du signal x, P et p0 désignant des constantes encadrant le nombre d'échantillons de signal sur lequel l'énergie est évaluée.

[0125] Les différences de niveaux ainsi obtenues lui permettent de déterminer la direction de l'objet localisé par rapport à la position de référence.

[0126] Cette direction est évaluée ici en termes d'excursion angulaire par rapport à l'axe Δ.

[0127] A cette fin, le système 1 de restitution associe à une différence de niveaux prédéfinie entre deux canaux, par exemple -30dB (respectivement de 30dB), une direction de l'objet sonore de 90° (respectivement de -90°) par rapport à l'axe Δ. Les directions comprises entre -90° et 90° sont ensuite estimées à partir d'une fonction d'interpolation croissante (ex. une fonction linéaire croissante) définie entre les deux valeurs -90° et 90°.

[0128] Le système 1 de restitution compare ensuite la direction de l'objet sonore ainsi évaluée par rapport à l'excursion angulaire $\Omega/2$ définissant la fenêtre spatiale W, afin de déterminer si l'objet se trouve à l'intérieur ou à l'extérieur de la fenêtre spatiale W : ainsi, un objet sonore pour lequel on a estimé une direction en valeur absolue supérieure à $\Omega/2$ par

rapport à l'axe Δ, est considéré par le système 1 comme à l'extérieur de la fenêtre spatiale W, tandis qu'un objet sonore pour lequel on a estimé une direction en valeur absolue inférieure ou égale à $\Omega/2$ par rapport à l'axe Δ, est considéré par le système 1 comme étant positionné à l'intérieur de la fenêtre spatiale W.

[0129] Dans le mode de réalisation décrit ici, le système 1 de restitution utilise également la direction estimée de l'objet sonore pour déterminer si cet objet se trouve au centre de la fenêtre spatiale W (à un delta de précision près), afin de mieux distinguer lors de la restitution, les objets situés au centre de la fenêtre W des autres objets situés dans la fenêtre W (étape E40).

[0130] Ainsi, un objet est considéré par le système 1 de restitution comme étant positionné au centre de la fenêtre spatiale W si sa direction est comprise dans un intervalle [0 ; δ] autour de l'axe Δ, où δ désigne un angle prédéfini, par exemple 2.5°.

[0131] Cette étape est toutefois optionnelle.

[0132] Des techniques alternatives peuvent être utilisées en variante pour extraire les objets sonores des signaux Si et estimer leurs caractéristiques (caractère diffus ou localisé, direction et position par rapport à la fenêtre W, et le cas échéant, amplitude) au cours des étapes E30 et E40.

[0133] Ainsi selon une deuxième variante de réalisation, la phase d'analyse spatiale $\sum$I comprend la détermination d'un vecteur de Gerzon représentatif de chaque signal Si audio multicanal (un vecteur est estimé pour chaque sous-bande fréquentielle BWi).

[0134] De façon connue de l'homme du métier, le vecteur de Gerzon d'un signal audio multicanal est dérivé à partir des contributions respectives (direction et intensité ou énergie) des différents canaux du signal multicanal à la scène sonore perçue par l'auditeur situé à la position de référence Pref. La détermination d'un tel vecteur pour un signal audio multicanal (ou de façon équivalente d'un vecteur de Gerzon normalisé) est décrite dans le document US 2007/0269063 et ne sera pas décrite plus en détail ici. On suppose ici que dans la deuxième variante de réalisation, le système 1 de restitution procède de manière identique à celle décrite dans ce document.

[0135] Le vecteur de Gerzon d'un signal audio multicanal reflète la localisation spatiale du signal audio multicanal telle qu'elle est perçue par l'auditeur depuis la position de référence. La détermination de ce vecteur de Gerzon permet de s'affranchir du calcul de corrélations entre les différents canaux du signal multicanal pour déterminer le caractère diffus ou localisé des objets sonores extraits du signal, et la position de ces objets par rapport à le fenêtre spatiale W.

[0136] Comme décrit dans le document US 2007/0269063, le vecteur de Gerzon associé à un signal multicanal Si s'écrit sous la forme d'un vecteur directif, donnant la direction de l'objet sonore associée à la sous-bande fréquentielle BWi, et d'un vecteur non directif (i.e. diffus).

[0137] Autrement dit, à partir des vecteurs de Gerzon associés aux signaux Si, le système 1 de restitution sonore est capable d'extraire les objets sonores localisés et diffus composant le signal S, et de déterminer la position des objets localisés par rapport à la fenêtre spatiale W (à partir de la direction des vecteurs de Gerzon, et notamment des vecteurs « directifs »), ainsi que leur amplitude (déterminée à partir de la norme des vecteurs de Gerzon et de la contribution des vecteurs directif/non directif).

[0138] Il procède à cet effet de manière similaire à celle décrite pour l'analyse temporelle des signaux Si, par comparaison de la norme des vecteurs par rapport à un ou plusieurs seuils prédéfinis, et de leur direction par rapport à l'excursion angulaire $\Omega/2$.

[0139] Plus précisément, pour chaque vecteur de Gerzon normalisé, la norme du vecteur directif et la norme du vecteur non directif sont comparées à un seuil dit inférieur, noté THR_inf, et à un seuil dit supérieur, noté THR_sup :

- si les normes des vecteurs directif et non directif du vecteur de Gerzon normalisé sont toutes les deux comprises entre THR_inf et THR_sup, les deux objets sonores (i.e. l'objet localisé correspondant au vecteur directif et l'objet diffus correspondant au vecteur non directif) sont extraits et restitués ;
- si, en revanche, l'un des vecteurs a une norme supérieure à THR_sup, seul l'objet correspondant à ce vecteur est extrait et restitué (i.e. on restitue uniquement un objet localisé ou un objet totalement diffus).

[0140] Les seuils THR_inf et THR_sup sont choisis de manière empirique, en fonction du compromis complexité versus perception de l'auditeur souhaité. Par exemple, THR_inf = 0.3 et THR_sup = 0.7 pour des amplitudes normalisées.

[0141] L'amplitude associée à chaque objet sonore ainsi extrait est alors dérivée de l'amplitude du vecteur directif ou non directif correspondant.

[0142] De façon alternative, les objets diffus et localisé donnés par le vecteur non directif et le vecteur directif dérivés du vecteur de Gerzon sont extraits tous les deux (pas de comparaison préalable par rapport à un seuil pour estimer si la contribution de l'un et/ou l'autre est suffisamment significative pour être restituée) afin d'être restitués sur les haut-parleurs du dispositif 2 de restitution.

[0143] La direction des vecteurs (i.e. directifs) correspondant aux objets sonores extraits est ensuite comparée par rapport à l'excursion angulaire $\Omega/2$, afin de déterminer leur position par rapport à la fenêtre W.

[0144] Par ailleurs, de façon similaire à l'analyse temporelle, le système 1 de restitution peut identifier les objets situés

au centre de la fenêtre spatiale W, de sorte à mieux les distinguer lors de la restitution par rapport aux autres objets localisés à l'intérieur de la fenêtre spatiale W.

[0145] Il convient de noter que les techniques d'analyse s'appuyant sur la détermination des vecteurs de Gerzon n'offrent pas la possibilité d'extraire plus d'un objet sonore localisé par sous-bande fréquentielle.

[0146] Pour remédier à cette limitation, dans une troisième variante de réalisation de l'invention, les moyens 3 d'analyse spatiale du système 1 de restitution mettent en oeuvre, pour extraire les objets sonores des signaux Si et estimer leurs caractéristiques au cours des étapes E30 et E40, une technique s'appuyant sur une décomposition spatiale de chaque signal multicanal Si en harmoniques sphériques.

[0147] De façon connue, pour chaque bande de fréquence, le champ sonore $p(r,\omega)$ dérivé de chaque signal multicanal Si peut être décomposé selon le formalisme des harmoniques sphériques, de la façon suivante :

$$p(r,\omega) = \sum_{n=0}^{+\infty} i^n j_n(kr) \sum_{m=-n}^{n} B_{mn}(\omega) Y_{mn}(\varphi,\theta),$$

où $Y_{mn}(\varphi,\theta)$ désigne l'harmonique sphérique de degré m et d'ordre n définie par :

$$Y_{mn}(\varphi,\theta) = \sqrt{(2n+1)\varepsilon_n \frac{(n-m)!}{(n+m)!}} P_{mn}(\sin\theta) \times \begin{cases} \cos(m\varphi) & si\ m \geq 0 \\ \sin(-m\varphi) & si\ m < 0 \end{cases},$$

$B_{mn}(\omega)$ désigne le coefficient (à la fréquence ω) associé à l'harmonique sphérique $Y_{mn}(\varphi,\theta)$ dans la décomposition, et :

$$i^2=-1,$$

k est une constante,

$$\varepsilon_n = \begin{cases} 1 & si\ n = 0 \\ 2 & sinon \end{cases}$$

$j_n(kr)$ est une fonction sphérique de Bessel de première espèce d'ordre n, $P_{mn}(\sin\theta)$ est la fonction de Legendre associée définie par :

$$P_{mn}(\sin\theta) = \frac{dP_n(\sin\theta)}{d(\sin\theta)^m}$$

où $P_n(sin\theta)$ désigne le polynôme de Legendre de première espèce d'ordre n.

[0148] Dans le cas particulier d'une onde plane de magnitude O_{pw} en provenance d'une direction $(\varphi_{pw}, \theta_{pw})$, les coefficients $B_{mn}(\omega)$ de la décomposition en harmoniques sphériques sont donnés par :

$$B_{mn}(\omega) = \frac{O_{pw}}{4\pi}$$

et sont indépendants de la fréquence.

[0149] Ainsi, dans cette troisième variante de réalisation, les moyens 3 d'analyse spatiale appliquent par exemple la technique d'extraction des objets sonores d'un signal multicanal à partir de sa décomposition spatiale en harmoniques sphériques décrite dans le document WO 2012/025580.

[0150] Cette technique s'appuie sur une représentation de la matrice $B(\omega,t)$, construite à partir des coefficients $B_{mn}(\omega)$ de la décomposition en harmoniques sphériques auxquels on a appliqué une transformée de Fourier STFT (pour « Short Time Fourier Transform ») à l'instant t, sous la forme d'une somme de deux termes, i.e., un premier terme modélisant les objets sonores localisés compris dans le signal Si, et un second terme modélisant les objets sonores diffus.

[0151] Les directions des objets sonores localisés sont obtenues à partir de matrice de corrélation $S_{BB}(\omega,t) = E\{B(\omega,t)B^H(\omega,t)\}$.

[0152] Une fois les objets sonores localisés extraits, leur contribution est retranchée du signal Si de sorte à obtenir, le cas échéant, les objets sonores diffus contenus dans le signal. Comme dans la deuxième variante basée sur la représentation du signal selon un vecteur de Gerzon, des seuils inférieur et supérieur peuvent être introduits pour se limiter à l'extraction d'objets sonores d'amplitude suffisante.

[0153] L'amplitude associée aux objets sonores localisés est déterminée à partir de la somme des coefficients d'harmoniques sphériques associés à ces objets en fonction de la direction estimée. L'amplitude des objets diffus est estimée à partir des coefficients des harmoniques sphériques résiduels obtenus après avoir retranché la contribution des objets sonores localisés.

[0154] Cette technique étant décrite dans le détail dans le document WO 2012/025580, elle ne sera pas détaillée davantage ici.

[0155] Pour déterminer la position des objets sonores localisés par rapport à la fenêtre spatiale W, le système 1 de restitution procède de manière similaire à celle décrite dans la première variante pour l'analyse temporelle des signaux Si, par comparaison de leur direction par rapport à l'excursion angulaire $\Omega/2$.

[0156] Par ailleurs, de façon similaire à l'analyse temporelle, le système 1 de restitution peut identifier les objets situés au centre de la fenêtre spatiale W, de sorte à mieux les distinguer lors de la restitution par rapport aux autres objets localisés à l'intérieur de la fenêtre spatiale W.

[0157] Il convient de noter que dans le mode de réalisation décrit ici (quelle que soit la technique retenue pour l'analyse spatiale), le système 1 de restitution ne s'attache pas à proprement parler à la position des objets sonores extraits des signaux Si par rapport au dispositif 2 de restitution, i.e., il n'établit pas de distinction entre les objets sonores selon si ceux-ci se situent derrière ou devant le dispositif 2 de restitution par rapport à la position de référence Pref. De façon alternative, l'analyse spatiale effectuée par le système 1 de restitution peut se limiter aux objets sonores situés derrière le dispositif 2 de restitution, quelle que soit la technique d'analyse spatiale retenue parmi les techniques précitées notamment.

[0158] Par ailleurs, dans le mode de réalisation décrit ici, une décomposition en sous-bandes fréquentielles du signal multicanal S est réalisée, puis le système 1 de restitution examine chaque sous-bande fréquentielle pour extraire les objets sonores du signal multicanal S. Ceci permet d'extraire plus précisément les objets sonores constituant le signal S (on peut identifier plus d'objets sonores notamment). Toutefois, cette hypothèse n'est pas limitative et on pourrait envisager dans le cadre de l'invention de travailler directement sur le signal multicanal S sans réaliser de décomposition en sous-bandes fréquentielles.

[0159] A l'issue de l'analyse spatiale $\sum$I, le système 1 de restitution a extrait et identifié plusieurs catégories d'objets sonores dans le signal multicanal S, à savoir :

- une première catégorie d'objets sonores, notée OBJLocIntW, regroupant les objets sonores localisés et situés à l'intérieur de la fenêtre spatiale W ;
- une deuxième catégorie d'objets sonores, notée OBJLocExtW, regroupant les objets sonores localisés et situés à l'extérieur de la fenêtre spatiale W ; et
- une troisième catégorie d'objets sonores, notée OBJDiff, regroupant les objets sonores diffus.

[0160] Le système 1 de restitution dispose par ailleurs, pour la première et la deuxième catégorie d'objets sonores, de la position de ces objets dans la fenêtre spatiale W.

[0161] Dans le mode de réalisation décrit ici, le système 1 de restitution a également identifié, au sein de la catégorie d'objets sonores OBJLocIntW, les objets sonores en provenance de sources positionnées au centre de la fenêtre spatiale W.

[0162] L'ensemble de ces informations est par exemple stocké dans la mémoire vive 6 ou dans la mémoire flash 7 non volatile du système 1 de restitution afin de pouvoir être utilisé en temps réel.

[0163] Comme mentionné précédemment, conformément à l'invention, selon une seconde phase $\sum$II dite de restitution du signal audio multicanal S, le système 1 va restituer les objets sonores extraits du signal S en fonction de leur catégorie, et des caractéristiques de ces objets déterminées au cours des étapes E30 et E40.

[0164] Plus précisément, dans le mode de réalisation décrit ici, les moyens 4 de restitution du système 1 de restitution appliquent quatre traitements distincts T-A1, TA-2, T-B et T-C sélectionnés en fonction des caractéristiques des objets sonores extraits par les moyens 3 d'analyse spatiale du système 1 de restitution au cours de la phase $\sum$I (étape E50).

[0165] Ainsi, dans le mode de réalisation décrit ici, les objets sonores identifiés comme appartenant à la première catégorie OBJLocIntW, sont restitués par les moyens 4 de restitution (et plus précisément par les moyens 4A), en appliquant les traitements TA1 ou T-A2 selon s'ils sont situés respectivement au centre ou non de la fenêtre spatiale W (étape E51).

[0166] Conformément à l'invention, les traitements T-A1 et T-A2 restituent les objets sonores de la catégorie OBJLo-

cIntW à l'intérieur de la fenêtre spatiale W.

[0167] Différents types de traitements T-A1 et T-A2 peuvent être envisagés pour une telle restitution. Ces traitements peuvent mettre en oeuvre ou non un filtrage des objets sonores avant leur diffusion sur tout ou partie des haut-parleurs du dispositif 2 de restitution.

[0168] Ainsi, par exemple, lorsque le dispositif 2 de restitution comprend un haut-parleur central et des haut-parleurs latéraux :

- le traitement T-A1 peut être apte à diffuser les objets sonores extraits du signal S identifiés au centre de la fenêtre spatiale W, directement sur le haut-parleur central du dispositif 2 ; et
- le traitement de restitution T-A2 peut être apte à diffuser les objets sonores extraits du signal S et positionnés à une position distincte du centre de la fenêtre spatiale W sur l'ensemble des haut-parleurs du dispositif 2 de restitution en utilisant un effet panoramique d'intensité, choisi de sorte à préserver la position des objets sonores perçue par l'auditeur à la position de référence.

[0169] En variante, les traitements de restitution T-A1 et/ou T-A2 appliqué sur les objets sonores localisés à l'intérieur de la fenêtre spatiale W peuvent être des traitements plus complexes de filtrage spatial comprenant par exemple la création de sources virtuelles 10 à partir des haut-parleurs du dispositif 2 de restitution à l'intérieur de la fenêtre spatiale W, les sources virtuelles étant positionnées en accord avec les caractéristiques des objets sonores estimées aux étapes E30 et/ou E40 (c'est-à-dire dans les directions et le cas échéant, selon les amplitudes estimées aux étapes E30 et E40).

[0170] La création de sources virtuelles à partir de haut-parleurs d'un dispositif de restitution est connue de l'homme du métier et ne sera pas décrite ici. Un traitement de restitution comprenant la création de sources virtuelles aux positions identifiées lors des étapes E30 et/ou E40 est par exemple un traitement de synthèse de champ acoustique aussi appelé traitement WFS connu de l'homme du métier ou une technique de formation de faisceau (ou « beamforming » en anglais), le faisceau étant dirigé par exemple vers la position de référence.

[0171] Les objets sonores appartenant respectivement aux catégories OBJLocExtW et OBJDiff sont restitués à l'extérieur de la fenêtre spatiale W par les moyens 4 de restitution (respectivement par les moyens 4-B et 4-C), en appliquant les traitements T-B et T-C (étapes E52 et E53).

[0172] Plus précisément, conformément à l'invention, les traitements de restitution T-B et T-C comprennent la création d'au moins une source virtuelle 11, 12 à l'extérieur de la fenêtre spatiale W de restitution du dispositif 2 de restitution.

[0173] Pour les objets sonores de la catégorie OBJLocExtW (étape E52), ces sources virtuelles 11 sont reconstituées à partir des positions des objets sonores identifiées à l'étape E30, via par exemple une technique transaurale (particulièrement bien adaptée pour une configuration du dispositif 2 de restitution avec un haut-parleur central et deux haut-parleurs latéraux), une technique WFS ou dérivée, telle que décrite par exemple dans la demande de brevet européen EP 1 116 572.0 non publiée, ou encore la formation d'un faisceau dirigé vers l'extérieur de la fenêtre spatiale de restitution, et dont la largeur peut être configurée de sorte à optimiser le rendu sonore.

[0174] Pour les objets sonores de la catégorie OBJDiff (étape E53), le traitement T-C permet la création de sources virtuelles 12 diffuses. On utilisera préférentiellement à cette fin des techniques T-C de « beamforming » pour créer ces sources virtuelles, pour lesquelles on contrôle aisément l'orientation et la largeur des faisceaux de sorte à créer des réflexions sur les murs de la pièce dans laquelle est positionnée le dispositif 2 de restitution et ainsi créer davantage de sensation d'enveloppement pour l'auditeur placé à la position de référence.

[0175] Afin de mieux comprendre l'invention, nous allons maintenant décrire trois exemples de mise en oeuvre, illustrant notamment différentes techniques d'analyse spatiale et différents traitements de restitution pouvant être envisagées au cours des différentes étapes de la figure 5.

Exemple 1 :

[0176] Dans ce premier exemple, on suppose que le dispositif 2 de restitution est une enceinte acoustique de type barre de son horizontale équipée de trois haut-parleurs 2-1, 2-2 et 2-3 (un haut-parleur central et deux haut-parleurs latéraux).

[0177] La position Pref est choisie ponctuelle, centrée par rapport au dispositif 2 de restitution.

[0178] On suppose par ailleurs que le signal multicanal S fourni au système 1 de restitution au cours de l'étape E10 est un signal audio stéréo, autrement dit, composé de deux canaux distincts.

[0179] Dans ce premier exemple, les étapes suivantes sont mises en oeuvre par le système 1 de restitution à partir du signal S :

(1) Décomposition du signal S en sous-bandes fréquentielles à l'étape E20 à l'aide d'une transformée de Fourier appliquée au signal S, chaque sous-bande fréquentielle comprenant un signal Si composé de deux canaux.

(2) Analyse spatiale $\sum$I du signal S, ou de façon équivalente de chaque signal Si sur chaque sous-bande fréquentielle

comprenant une analyse temporelle du signal Si au cours de l'étape E30 permettant d'extraire un objet sonore du signal Si, cette analyse temporelle incluant notamment :

- l'évaluation de la corrélation normalisée entre les deux canaux du signal Si et la comparaison de cette corrélation par rapport au seuil prédéfini THR afin d'estimer le caractère local ou diffus de l'objet sonore inclus dans le signal Si ;
- l'évaluation de la différence de niveaux entre les deux canaux du signal Si, et la transformation de cette différence de niveaux en excursion angulaire par rapport à l'axe Δ reliant la position Pref au centre du dispositif 2 de restitution. On considère dans ce premier exemple qu'une différence de niveaux de -30dB (respectivement 30dB) correspond à une excursion angulaire de 90°, (respectivement -90°C), les valeurs intermédiaires étant estimées à l'aide d'une fonction linéaire entre ces deux bornes ;
- l'estimation de la fenêtre spatiale W de restitution sonore (et de l'excursion angulaire associée à cette fenêtre), définie par la position de référence Pref et les haut-parleurs latéraux du dispositif 2 de restitution. A titre illustratif, si l'on considère une position de référence Pref placée à une distance de 2 à 4m du dispositif 2 de restitution et un dispositif de restitution de largeur 1m, les haut-parleurs latéraux de ce dispositif étant placés aux extrémités du dispositif, l'excursion angulaire $\Omega/2$ correspondant à la fenêtre spatiale W est comprise entre 7 et 15° ; et
- à partir de l'excursion angulaire obtenue pour l'objet sonore extrait du signal Si et l'excursion angulaire $\Omega/2$ correspondant à la fenêtre spatiale W, la détermination de la direction de l'objet sonore et de sa position par rapport à la fenêtre W. Ainsi, si l'objet sonore extrait de Si présente une excursion angulaire inférieure ou égale à $\Omega/2$, il est estimé comme étant positionné dans la fenêtre spatiale W. Inversement, si l'objet sonore extrait de Si a présente une excursion angulaire supérieure à $\Omega/2$, il est estimé comme étant positionné à l'extérieur de la fenêtre spatiale W.

L'amplitude de chaque objet sonore extrait sur chaque sous-bande fréquentielle est donnée par le niveau du signal Si sur cette sous-bande.
L'analyse spatiale du signal S comprend également, dans le premier exemple envisagé ici, l'identification E40 des objets sonores localisés au centre de la fenêtre spatiale W en comparant l'excursion angulaire associée à chaque objet sonore extrait des signaux Si à l'intervalle [0 ; 2.5°], un objet sonore étant considéré comme étant au centre de la fenêtre si son excursion angulaire est comprise entre 0 et 2.5° (en valeur absolue).
(3) Restitution $\sum$II/E50 du signal S, et plus précisément des objets sonores extraits lors de l'analyse spatiale $\sum$I :

- au cours de l'étape E51, restitution à l'intérieur de la fenêtre spatiale W, des objets sonores localisés estimés comme étant positionnés à l'intérieur de la fenêtre spatiale W (catégorie OBJLocIntW), à l'aide des traitements de restitution T-A1 et T-A2 suivants :
  - traitement T-A1 appliqué aux objets sonores estimés au centre de la fenêtre spatiale W : diffusion des objets sonores directement (i.e. sans filtrage spatial) sur le haut-parleur central du dispositif 2 de restitution, autrement dit, les objets sonores ainsi restitués sont rattachés au centre du dispositif 2 de restitution ;
  - traitement T-A2 appliqué aux objets sonores non centrés localisés dans la fenêtre spatiale W : diffusion des objets sonores sur les trois haut-parleurs du dispositif 2 de restitution en utilisant un effet panoramique d'intensité ;
- au cours de l'étape E52, restitution à l'extérieur de la fenêtre spatiale W, des objets sonores localisés estimés comme étant positionnés à l'extérieur de la fenêtre spatiale W (catégorie OBJLocExtW), à l'aide d'une technique de restitution T-B transaurale. Plus précisément, on crée, à l'aide des deux haut-parleurs latéraux du dispositif 2 de restitution, des sources virtuelles transaurales placées en dehors de la fenêtre W, par exemple à 30° et 60° (respectivement à -30° et -60°) par rapport à l'axe Δ. Les objets sonores de la catégorie OBJLocExtW sont alors diffusés à travers ces sources virtuelles, dans les directions déterminées à l'étape E30 ;
- au cours de l'étape E53, restitution à l'extérieur de la fenêtre spatiale W, des objets sonores diffus (catégorie OBJDiff), à l'aide d'une technique de restitution T-C transaurale. Plus précisément, on crée à l'aide des deux haut-parleurs latéraux du dispositif 2 de restitution, des sources virtuelles transaurales placées en dehors de la fenêtre W, à un angle supérieur à 60° (respectivement inférieur à -60°) par rapport à l'axe Δ. Les objets sonores de la catégorie OBJDiff sont alors diffusés à travers ces sources virtuelles.

Les techniques de restitution transaurales sont connues de l'homme du métier, et décrites par exemple dans le document de J. Bauck et D.H. Cooper, intitulé « Generalized Transaural Stereo and Applications », Journal Audio Engineering Society, vol. 44 n° 9, 1996. De telles techniques consistent à appliquer un filtre sur chacun des haut-

parleurs latéraux du dispositif 2 de restitution, chaque filtre comprenant un filtre de spatialisation et un filtre d'annulation de la propagation croisée entre les deux haut-parleurs.

Exemple 2 :

[0180] Dans ce deuxième exemple, on suppose que le dispositif 2 de restitution est une enceinte acoustique compacte de type barre de son horizontale équipée de 15 haut-parleurs 2-1, 2-2, ..., 2-15 d'une longueur d'environ 1m.

[0181] La position Pref est choisie ponctuelle, centrée par rapport au dispositif 2 de restitution.

[0182] On suppose par ailleurs que le signal multicanal S fourni au système 1 de restitution au cours de l'étape E10 est un signal audio 5.1. Un tel signal contient déjà intrinsèquement une information de spatialisation. Plus spécifiquement, la norme ITU-R BS.775-1 définissant le format des signaux 5.1 sous-entend un centre situé à 0°, des canaux gauche L et droite R situés à +/- 30° par rapport au centre, et des canaux arrière gauche Ls et arrière droit Rs situés à +/-110° par rapport au centre.

[0183] Dans ce deuxième exemple, les étapes suivantes sont mises en oeuvre par le système 1 de restitution à partir du signal S :

(1) Décomposition du signal S en sous-bandes fréquentielles à l'étape E20 à l'aide d'une transformée de Fourier appliquée au signal S, chaque sous-bande fréquentielle comprenant un signal Si composé de cinq canaux.

(2) Analyse spatiale $\sum$I du signal S, ou de façon équivalente de chaque signal Si sur chaque sous-bande fréquentielle comprenant, au cours de l'étape E30, la détermination d'un vecteur de Gerzon associé à chaque signal Si, de façon similaire à celle décrite dans le document US2007269063.

Les objets sonores situés au centre de la fenêtre spatiale W sont présents dans le canal central par définition du format 5.1. Ils sont donc « extraits » aisément à partir de ce canal central déjà isolé.

Le système 1 de restitution considère ensuite le signal Si' composé des quatre canaux L, R, Ls et Rs du signal Si, et les quatre vecteurs « canaux » reliant la position de référence Pref aux quatre canaux L, R, Ls et Rs. Il affecte à chaque vecteur canal un poids correspondant à l'énergie du canal associé. Le vecteur de Gerzon associé au signal Si' (ou de façon équivalente au signal Si) est défini comme le barycentre des points L, R, Ls et Rs ainsi pondérés.

Le vecteur de Gerzon ainsi défini s'écrit sous la forme d'un vecteur directif (égal à la somme des deux vecteurs canaux adjacents au vecteur de Gerzon : ainsi par exemple, si la direction du vecteur de Gerzon est de 15° par rapport à l'axe Δ, le vecteur directif est la somme des vecteurs canaux associés respectivement aux canaux L et R), et d'un vecteur non directif.

Le vecteur directif caractérise un objet sonore localisé du signal Si et sa position (donnée par la direction du vecteur) par rapport à la fenêtre W. Le système 1 de restitution compare cette position par rapport à l'excursion angulaire $\Omega/2$ de façon similaire à l'exemple 1, pour estimer si l'objet sonore ainsi identifié appartient à la catégorie OBJLocIntW ou à la catégorie OBJLocExtW.

Le vecteur non directif caractérise un objet sonore diffus du signal Si, classé par le système 1 de restitution dans la catégorie OBJDiff.

Le système 1 de restitution associe à chaque objet sonore extrait une amplitude évaluée à partir de l'amplitude du vecteur correspondant (directif ou non directif et composant le vecteur de Gerzon).

(3) Restitution II/E50 du signal S, et plus précisément des objets sonores extraits lors de l'analyse spatiale $\sum$I, selon les directions et amplitudes estimées à l'étape E30 :

- au cours de l'étape E51, restitution à l'intérieur de la fenêtre spatiale W, des objets sonores localisés estimés comme étant positionnés à l'intérieur de la fenêtre spatiale W (catégorie OBJLocIntW), à l'aide des traitements de restitution T-A1 et T-A2 suivants :
  - traitement T-A1 appliqué aux objets sonores estimés au centre de la fenêtre spatiale W (i.e. objets contenus dans le canal central du signal S) : diffusion des objets sonores directement (i.e. sans filtrage spatial) sur le haut-parleur central du dispositif 2 de restitution, autrement dit, les objets sonores ainsi restitués sont rattachés au centre du dispositif 2 de restitution ;
  - traitement T-A2 appliqué aux objets sonores non centrés localisés dans la fenêtre spatiale W : diffusion des objets sonores à l'aide d'une technique WFS de synthèse de champ acoustique comprenant la création de sources virtuelles via les haut-parleurs du dispositif 2 de restitution, ces sources virtuelles étant positionnées (en agissant sur les retards et les gains appliqués à chaque haut-parleur) dans les directions estimées par les vecteurs directifs extraits des vecteurs de Gerzon dérivés lors de l'analyse spatiale de sorte à respecter la même organisation spatiale que lors du mixage du signal multicanal. Les amplitudes des objets sonores restitués sont conformes aux amplitudes évaluées à l'étape E30 ;

◦ au cours de l'étape E52, restitution à l'extérieur de la fenêtre spatiale W, des objets sonores localisés estimés comme étant positionnés à l'extérieur de la fenêtre spatiale W (catégorie OBJLocExtW), à l'aide d'une technique WFS comprenant la création de six sources virtuelles entourant la position de référence Pref :

- deux sources virtuelles sont positionnées aux extrémités du dispositif 2 de restitution,
- quatre sources virtuelles sont positionnées à l'extérieur de la fenêtre spatiale W, parmi lesquelles : deux sources virtuelles sont positionnées entre 30°C et 60°C par rapport à l'axe Δ, et entre -30° et -60°, par exemple à l'aide de deux ondes places dirigées vers les murs latéraux de la pièce dans laquelle est placé le dispositif 2 de restitution ; et deux sources virtuelles sont positionnées entre entre 135° et 150° et entre -135° et -150°, par exemple à l'aide de deux ondes places dirigées vers les murs arrière de la pièce dans laquelle est placé le dispositif 2 de restitution.

Les sources virtuelles ainsi positionnées sont utilisées pour restituer les objets sonores de la catégorie OBJLocExtW selon les directions et les amplitudes estimées à l'étape E30 ;

◦ au cours de l'étape E53, restitution à l'extérieur de la fenêtre spatiale W, des objets sonores diffus (catégorie OBJDiff), à l'aide d'une technique WFS de restitution T-C, comprenant la création de quatre sources virtuelles à l'extérieur de la fenêtre W à l'aide par exemple de quatre ondes planes dirigées vers les murs de la pièce dans laquelle est placé le dispositif 2 de restitution de manière à créer deux réflexions sur les murs latéraux situés entre 60° et 80° (respectivement -60° et -80°) par rapport à l'axe Δ.

Les techniques de synthèse de champ acoustique ou WFS sont connues de l'homme du métier, et décrites par exemple dans le document de A.J. Berkhout et al. intitulé « A holographic approach to acoustic control », J. Audio. Eng. Soc. Vol. 36, 1988. De telles techniques consistent à appliquer gain et un retard à chaque haut-parleur du dispositif 2 de restitution. Elles reposent uniquement sur la position relative des sources virtuelles que l'on souhaite créer (i.e. sources ponctuelles ou ondes planes) par rapport à la position physique des différents haut-parleurs du dispositif 2 de restitution.

Exemple 3 :

[0184] Dans ce troisième exemple, on suppose que le dispositif 2 de restitution est une enceinte acoustique compacte équipée de 8 haut-parleurs 2-1, 2-2, ..., 2-8 de largeur environ 80cm, avec quatre haut-parleurs frontaux 2-1,..., 2-4, et deux haut-parleurs 2-5 et 2-6, respectivement 2-7 et 2-8, situés de chaque côté du dispositif 2 (dispositif similaire au dispositif 2" illustré à la figure 3B).

[0185] La position Pref est choisie ponctuelle, centrée par rapport au dispositif 2 de restitution.

[0186] On suppose par ailleurs que le signal multicanal S fourni au système 1 de restitution au cours de l'étape E10 est un signal audio composé de quatre canaux distincts.

[0187] Dans ce troisième exemple, les étapes suivantes sont mises en oeuvre par le système 1 de restitution à partir du signal S :

(1) Décomposition du signal S en sous-bandes fréquentielles à l'étape E20 à l'aide d'une transformée de Fourier appliquée au signal S, chaque sous-bande fréquentielle comprenant un signal Si composé de quatre canaux.
(2) Analyse spatiale $\sum$I du signal S, ou de façon équivalente de chaque signal Si sur chaque sous-bande fréquentielle comprenant, au cours de l'étape E30 :

◦ la décomposition spatiale en harmoniques sphériques
◦ de chaque signal Si, l'extraction des objets sonores diffus et localisés de chaque signal ainsi que la détermination de leurs caractéristiques (directions et amplitudes) selon la technique décrite dans le document WO 2012/025580 (cette étape peut éventuellement inclure le codage du signal Si dans un format audio de type HOA, connu en soi) ;
◦ la séparation des objets sonores localisés détectés lors du balayage selon les catégories OBJLocIntW et OBJLocExtW en comparant la direction examinée dans laquelle ces objets ont été détectés par rapport à l'excursion angulaire $\Omega/2$ associés à la fenêtre spatiale W, comme décrit précédemment pour les exemples 1 et 2 ;

(3) Restitution $\sum$II/E50 du signal S, et plus précisément des objets sonores extraits lors de l'analyse spatiale $\sum$I :

◦ au cours de l'étape E51, restitution à l'intérieur de la fenêtre spatiale W, des objets sonores localisés estimés comme étant positionnés à l'intérieur de la fenêtre spatiale W (catégorie OBJLocIntW), à l'aide d'un traitement T-A de restitution combinant une technique WFS et un contrôle de rayonnement prenant en compte le rayon-

nement de chaque haut-parleur et l'influence de l'enceinte acoustique à proprement parler contenant les différents haut-parleurs. Le champ de restitution sonore de chaque objet est contrôlé par l'intermédiaire de filtrages. Un tel traitement est décrit notamment dans la demande de brevet européen non encore publiée EP 1116572.0. Ainsi, plus précisément, dans ce troisième exemple, le traitement T-A comprend la création de sources virtuelles derrière le dispositif 2 de restitution via la technique WFS, et l'application d'un filtrage aux haut-parleurs 2-1, ..., 2-8 du dispositif 2 déterminé de sorte que l'énergie des objets sonores restitués par ces sources virtuelles est dirigée vers la position de référence et est en accord avec les amplitudes déterminées à l'étape E30 ;

◦ au cours de l'étape E52, restitution à l'extérieur de la fenêtre spatiale W, des objets sonores localisés estimés comme étant positionnés à l'extérieur de la fenêtre spatiale W (catégorie OBJLocExtW), à l'aide d'un traitement de restitution T-B telle que décrit dans la demande de brevet européen non encore publiée EP 1116572.0 et combinant :

- une technique WFS comprenant la création de sources virtuelles à l'extérieur de la fenêtre spatiale W via la formation de deux faisceaux fins se réfléchissant sur les murs latéraux de la pièce dans laquelle est installée le dispositif 2 de restitution à une position ponctuelle prédéterminée ; et
- un filtrage appliqué aux haut-parleurs 2-1, ..., 2-8 du dispositif 2 déterminé de sorte que l'énergie des objets sonores restitués par ces sources virtuelles est dirigée concentrée vers les murs latéraux de la pièce.

Les sources virtuelles ainsi positionnées sont utilisées pour restituer les objets sonores de la catégorie OBJLocExtW selon les directions et les amplitudes estimées à l'étape E30 ;

◦ au cours de l'étape E53, restitution à l'extérieur de la fenêtre spatiale W, des objets sonores diffus (catégorie OBJDiff), à l'aide d'un traitement de restitution T-C telle que décrit dans la demande de brevet européen non encore publiée EP 1116572.0 et combinant :

- une technique WFS comprenant la création de sources virtuelles à l'extérieur de la fenêtre spatiale W via la formation de deux faisceaux larges se réfléchissant sur une zone étendue prédéterminée des murs latéraux de la pièce dans laquelle est installée le dispositif 2 de restitution ; et
- un filtrage appliqué aux haut-parleurs 2-1, ..., 2-8 du dispositif 2 déterminé de sorte que l'énergie des objets sonores restitués par ces sources virtuelles est dirigée concentrée vers les murs latéraux de la pièce.

[0188] Bien entendu, ces trois exemples ne sont donnés qu'à titre illustratif et d'autres configurations de dispositif de restitution, ainsi que d'autres techniques d'analyse spatiale et d'autres traitements de restitution peuvent être utilisés dans le cadre de l'invention.

Revendications

1. Procédé de restitution d'un signal audio (S) multicanal sur un dispositif (2) de restitution équipé d'une pluralité de haut-parleurs (2-1,...,2-N), lesdits haut-parleurs étant disposés à des emplacements fixes du dispositif de restitution et définissant une fenêtre spatiale (W) de restitution sonore par rapport à une position spatiale (Pref) dite de référence, ledit procédé de restitution comprenant :

- une étape ($\sum$I) d'analyse spatiale du signal audio (S) multicanal comprenant :

◦ l'extraction (E30) d'au moins un objet sonore du signal, et
◦ pour chaque objet sonore extrait, l'estimation (E30) d'un caractère diffus ou localisé de cet objet sonore, et d'une position de cet objet sonore par rapport à la fenêtre spatiale de restitution sonore du dispositif de restitution ; et ledit procédé étant **caractérisé par**

- une étape ($\sum$II,E50,E51,E52,E53) de restitution du signal audio sur la pluralité (2-1,...,2-N) de haut-parleurs du dispositif de restitution (2), au cours de laquelle on applique, à chaque objet sonore extrait du signal audio, un traitement (T-A1, T-A2,T-B,T-C) de restitution sur au moins un haut-parleur de la pluralité de haut-parleurs du dispositif de restitution (2), ce traitement de restitution dépendant du caractère diffus ou localisé de l'objet sonore et de sa position par rapport à la fenêtre spatiale de restitution sonore estimés au cours de l'étape d'analyse spatiale,
le traitement de restitution (T-B,T-C) comprenant la création d'au moins une source virtuelle (11,12) à l'extérieur de la fenêtre spatiale (W) de restitution du dispositif de restitution, à partir des haut-parleurs du dispositif de restitution, lorsque l'objet sonore est estimé au cours de l'étape d'analyse spatiale (E50) comme étant diffus

ou positionné à l'extérieur de la fenêtre spatiale de restitution du dispositif de restitution.

2. Procédé selon la revendication 1 dans lequel ledit dispositif de restitution (2) est une enceinte acoustique (2) dans laquelle est disposée ladite pluralité de haut-parleurs (2-1,...,2-N).

3. Procédé selon la revendication 1 ou 2 dans lequel l'étape d'analyse spatiale (∑I) comprend en outre l'estimation (E40) de la position de l'objet sonore par rapport au centre de la fenêtre spatiale (W) de restitution sonore du dispositif de restitution (2).

4. Procédé selon l'une quelconque des revendications 1 à 3 dans lequel l'étape d'analyse spatiale (∑I) comprend une décomposition (E20) du signal audio reçu en une pluralité de sous-bandes fréquentielles, l'extraction (E30) dudit au moins un objet sonore étant réalisée sur au moins une sous-bande fréquentielle.

5. Procédé selon l'une quelconque des revendications 1 à 4 dans lequel le caractère diffus ou localisé de l'objet sonore extrait est estimé à partir d'au moins une corrélation évaluée entre deux canaux distincts du signal (S) audio multicanal.

6. Procédé selon l'une quelconque des revendications 1 à 5 dans lequel la position de l'objet sonore extrait par rapport à la fenêtre spatiale (W) de restitution sonore est estimée à partir d'au moins une différence de niveaux évaluée entre deux canaux distincts du signal (S) audio multicanal.

7. Procédé selon l'une quelconque des revendications 1 à 5 dans lequel l'étape d'analyse spatiale (∑I) comprend la détermination d'un vecteur de Gerzon représentatif du signal audio multicanal.

8. Procédé selon l'une quelconque des revendications 1 à 4 dans lequel l'étape d'analyse spatiale (∑I) comprend une décomposition spatiale du signal multicanal (S) en harmoniques sphériques.

9. Procédé selon l'une quelconque des revendications 1 à 8 dans lequel lorsqu'un objet sonore extrait est estimé comme étant localisé et positionné à l'intérieur de la fenêtre spatiale (W) de restitution sonore du dispositif de restitution, le traitement de restitution (T-A1, T-A2) appliqué à cet objet sonore au cours de l'étape de restitution est apte à restituer cet objet sonore à l'intérieur de la fenêtre spatiale (W) de restitution sonore du dispositif de restitution.

10. Procédé selon la revendication 9 dans lequel ledit traitement de restitution (T-A1, T-A2) comprend la création d'au moins une source virtuelle (10) à partir des haut-parleurs (2-1,...,2-N) du dispositif de restitution (2) à l'intérieur de la fenêtre spatiale (W) de restitution sonore du dispositif de restitution.

11. Procédé selon l'une quelconque des revendications 1 à 10 dans lequel lorsque l'objet sonore extrait est estimé au cours de l'étape d'analyse spatiale (∑I) comme étant positionné à l'extérieur de la fenêtre spatiale (W) de restitution du dispositif de restitution (2), la création (T-B) d'au moins une source virtuelle (11) à l'extérieur de la fenêtre spatiale (W) de restitution du dispositif de restitution comprend la formation d'au moins un faisceau dirigé vers l'extérieur de la fenêtre spatiale de restitution.

12. Procédé selon l'une quelconque des revendications 1 à 11 dans lequel :

- la pluralité de haut-parleurs (2-1,...,2-N) du dispositif de restitution (2) comprend un haut-parleur central et des haut-parleurs latéraux ; et
- lorsque l'objet sonore extrait est estimé au cours de l'étape d'analyse spatiale (∑I) comme étant diffus ou positionné à l'extérieur de la fenêtre spatiale (W) de restitution du dispositif de restitution, le traitement de restitution (T-B, T-C) appliqué à cet objet sonore utilise une technique transaurale de restitution de cet objet sonore sur les haut-parleurs latéraux du dispositif de restitution.

13. Procédé selon l'une quelconque des revendications 1 à 12 dans lequel lorsqu'un objet sonore extrait est estimé au cours de l'étape d'analyse spatiale (∑I) comme étant localisé et positionné à l'intérieur de la fenêtre spatiale (W) de restitution du dispositif de restitution (2), le traitement (T-A1, T-A2) de restitution appliqué à cet objet sonore au cours de l'étape de restitution (E50,E51), comprend la formation d'un faisceau dirigé vers ladite position spatiale (Pref) de référence.

14. Procédé selon l'une quelconque des revendications 1 à 12 dans lequel :

- la pluralité (2-1,...,2-N) de haut-parleurs du dispositif de restitution comprend un haut-parleur central et des haut-parleurs latéraux ; et
- lorsqu'un objet sonore extrait est estimé au cours de l'étape d'analyse spatiale (∑I) comme étant localisé et positionné au centre de la fenêtre spatiale (W) de restitution du dispositif de restitution, cet objet sonore est diffusé, au cours de l'étape de restitution (E50,E51), par le traitement de restitution (T-A1), sur le haut-parleur central du dispositif de restitution.

15. Procédé selon l'une quelconque des revendications 1 à 12 et 14 dans lequel lorsqu'un objet sonore extrait est estimé au cours de l'étape d'analyse spatiale (∑I) comme étant localisé et positionné à l'intérieur de la fenêtre spatiale (W) de restitution du dispositif de restitution (2) en une position distincte du centre de la fenêtre (W), le traitement de restitution (T-A2) appliqué au cours de l'étape de restitution (E50,E51) diffuse cet objet sonore sur les haut-parleurs (2-1,...,2-N) du dispositif de restitution en utilisant un effet panoramique d'intensité.

16. Programme comportant des instructions pour l'exécution des étapes du procédé de restitution selon l'une quelconque des revendications 1 à 15 lorsque ledit programme est exécuté par un ordinateur ou par un microprocesseur.

17. Système (1) de restitution d'un signal (S) audio multicanal sur un dispositif (2) de restitution équipé d'une pluralité de haut-parleurs (2-1,...,2-N), lesdits haut-parleurs étant disposés à des emplacements fixes du dispositif de restitution (2) et définissant une fenêtre spatiale (W) de restitution sonore par rapport à une position de référence (Pref), ledit système de restitution comprenant :

- des moyens (3) d'analyse spatiale du signal (S) audio multicanal comprenant :

  ◦ des moyens d'extraction d'au moins un objet sonore du signal, et
  ◦ des moyens d'estimation, pour chaque objet sonore extrait, d'un caractère diffus ou localisé de cet objet sonore, et d'une position de cet objet sonore par rapport à la fenêtre spatiale (W) de restitution sonore du dispositif de restitution ; et ledit système étant **caractérisé par**

- des moyens (4,4A,4B,4C) de restitution du signal (S) audio sur la pluralité de haut-parleurs (2-1,...,2-N) du dispositif de restitution (2), aptes à appliquer à chaque objet sonore extrait du signal audio, un traitement (T-A1,T-A2,T-B,T-C) de restitution sur au moins un haut-parleur de la pluralité de haut-parleurs du dispositif de restitution, ce traitement de restitution dépendant du caractère diffus ou localisé de l'objet sonore et de sa position par rapport à la fenêtre spatiale (W) de restitution sonore estimés par les moyens d'analyse spatiale, le traitement de restitution (T-B,T-C) comprenant la création d'au moins une source virtuelle (11,12) à l'extérieur de la fenêtre spatiale (W) de restitution du dispositif de restitution, à partir des haut-parleurs du dispositif de restitution, lorsque l'objet sonore est estimé par les moyens (3) d'analyse spatiale comme étant diffus ou positionné à l'extérieur de la fenêtre spatiale de restitution du dispositif de restitution.

Patentansprüche

1. Verfahren zur Wiedergabe eines Mehrkanalaudiosignals (S) an einer mit einer Vielzahl von Lautsprechern (2-1, ..., 2-N) ausgestatteten Wiedergabevorrichtung (2), wobei die Lautsprecher an festen Orten der Wiedergabevorrichtung angeordnet sind und ein Raumfenster (W) zur Tonwiedergabe in Bezug auf eine sogenannte räumliche Referenzposition (Pref) definieren, wobei das Wiedergabeverfahren umfasst:

- einen Schritt (ΣI) zur Raumanalyse des Mehrkanalaudiosignals (S), umfassend:

  • die Extraktion (E30) wenigstens eines Klangobjektes des Signals und
  • für jedes extrahierte Klangobjekt die Schätzung (E30) eines diffusen oder lokalisierten Merkmals dieses Klangobjektes und einer Position dieses Klangobjektes in Bezug auf das Tonwiedergabe-Raumfenster der Wiedergabevorrichtung, und

wobei das Verfahren **gekennzeichnet ist durch**
- einen Schritt (ΣII, E50, E51, E52, E53) zur Wiedergabe des Audiosignals an der Vielzahl (2-1, ..., 2-N) von Lautsprechern der Wiedergabevorrichtung (2), im Laufe dessen auf jedes extrahierte Klangobjekt des Audiosignals eine Wiedergabebehandlung (T-A1, T-A2, T-B, T-C) an wenigstens einem Lautsprecher der Vielzahl von Lautsprechern der Wiedergabevorrichtung (2) angewandt wird, wobei diese Wiedergabebehandlung von

dem diffusen oder lokalisierten Merkmal des Klangobjektes und von seiner Position in Bezug auf das Tonwiedergabe-Raumfenster, welche im Laufe des Schrittes der Raumanalyse geschätzt werden, abhängt,
wobei die Wiedergabebehandlung (T-B, T-C) das Einrichten wenigstens einer virtuellen Quelle (11, 12) außerhalb des Wiedergabe-Raumfensters (W) der Wiedergabevorrichtung anhand der Lautsprecher der Wiedergabevorrichtung umfasst, wenn das Klangobjekt im Laufe des Schrittes der Raumanalyse (E50) als diffus oder als außerhalb des Wiedergabe-Raumfensters der Wiedergabevorrichtung positioniert geschätzt wird.

2. Verfahren nach Anspruch 1, bei dem die Wiedergabevorrichtung (2) eine Lautsprecherbox (2) ist, in der die Vielzahl von Lautsprechern (2-1, ..., 2-N) angeordnet ist.

3. Verfahren nach Anspruch 1 oder 2, bei dem der Schritt der Raumanalyse (ΣI) ferner die Schätzung (E40) der Position des Klangobjektes in Bezug auf das Zentrum des Tonwiedergabe-Raumfensters (W) der Wiedergabevorrichtung (2) umfasst.

4. Verfahren nach einem der Ansprüche 1 bis 3, bei dem der Schritt der Raumanalyse (ΣI) eine Zerlegung (E20) des empfangenen Audiosignals in eine Vielzahl von Frequenzunterbändern umfasst, wobei die Extraktion (E30) des wenigstens einen Klangobjektes auf wenigstens einem Frequenzunterband durchgeführt wird.

5. Verfahren nach einem der Ansprüche 1 bis 4, bei dem das diffuse oder lokalisierte Merkmal des extrahierten Klangobjektes anhand wenigstens einer bewerteten Korrelation zwischen zwei getrennten Kanälen des Mehrkanalaudiosignals (S) geschätzt wird.

6. Verfahren nach einem der Ansprüche 1 bis 5, bei dem die Position des extrahierten Klangobjektes in Bezug auf das Tonwiedergabe-Raumfenster (W) anhand wenigstens eines bewerteten Pegelunterschiedes zwischen zwei getrennten Kanälen des Mehrkanalaudiosignals (S) geschätzt wird.

7. Verfahren nach einem der Ansprüche 1 bis 5, bei dem der Schritt der Raumanalayse (Y-1) die Bestimmung eines für das Mehrkanalaudiosignal repräsentativen Gerzon-Vektors umfasst.

8. Verfahren nach einem der Ansprüche 1 bis 4, bei dem der Schritt der Raumanalayse (ΣI) eine räumliche Zerlegung des Mehrkanalsignals (S) in sphärische Harmonische umfasst.

9. Verfahren nach einem der Ansprüche 1 bis 8, bei dem, wenn ein extrahiertes Klangobjekt als lokalisiert und als innerhalb des Tonwiedergabe-Raumfensters (W) der Wiedergabevorrichtung positioniert geschätzt wird, die Wiedergabebehandlung (T-A1, T-A2), die im Laufe des Wiedergabeschrittes auf dieses Klangobjekt angewandt wird, geeignet ist, dieses Klangobjekt innerhalb des Tonwiedergabe-Raumfensters (W) der Wiedergabevorrichtung wiederzugeben.

10. Verfahren nach Anspruch 9, bei dem die Wiedergabebehandlung (T-A1, T-A2) das Einrichten wenigstens einer virtuellen Quelle (10) anhand der Lautsprecher (2-1, ..., 2-N) der Wiedergabevorrichtung (2) innerhalb des Tonwiedergabe-Raumfensters (W) der Wiedergabevorrichtung umfasst.

11. Verfahren nach einem der Ansprüche 1 bis 10, bei dem, wenn das extrahierte Klangobjekt im Laufe des Schrittes der Raumanalyse (ΣI) als außerhalb des Wiedergabe-Raumfensters (W) der Wiedergabevorrichtung (2) positioniert geschätzt wird, das Einrichten (T-B) wenigstens einer virtuellen Quelle (11) außerhalb des Wiedergabe-Raumfensters (W) der Wiedergabevorrichtung die Bildung wenigstens eines zum Äußeren des Wiedergabe-Raumfensters gerichteten Strahls umfasst.

12. Verfahren nach einem der Ansprüche 1 bis 11, bei dem:

- die Vielzahl von Lautsprechern (2-1, ..., 2-N) der Wiedergabevorrichtung (2) einen zentralen Lautsprecher und seitliche Lautsprecher umfasst und,
- wenn das extrahierte Klangobjekt im Laufe des Schrittes der Raumanalyse (ΣI) als diffus oder als außerhalb des Wiedergabe-Raumfensters (W) der Wiedergabevorrichtung positioniert geschätzt wird, die Wiedergabebehandlung (T-B, T-C), welche auf dieses Klangobjekt angewandt wird, eine transaurale Wiedergabetechnik dieses Klangobjektes an den seitlichen Lautsprechern der Wiedergabevorrichtung verwendet.

13. Verfahren nach einem der Ansprüche 1 bis 12, bei dem, wenn ein extrahiertes Klangobjekt im Laufe des Schrittes

der Raumanalyse (ΣI) als lokalisiert und als innerhalb des Wiedergabe-Raumfensters (W) der Wiedergabevorrichtung (2) positioniert geschätzt wird, die Wiedergabebehandlung (T-A1, T-A2), welche im Laufe des Wiedergabeschrittes (E50, E51) auf dieses Klangobjekt angewandt wird, die Bildung eines in Richtung der räumlichen Referenzposition (Pref) gerichteten Strahls umfasst.

14. Verfahren nach einem der Ansprüche 1 bis 12, bei dem:

- die Vielzahl (2-1, ..., 2-N) von Lautsprechern der Wiedergabevorrichtung einen zentralen Lautsprecher und seitliche Lautsprecher umfasst und,
- wenn ein extrahiertes Klangobjekt im Laufe des Schrittes der Raumanalyse (ΣI) als lokalisiert und als im Zentrum des Wiedergabe-Raumfensters (W) der Wiedergabevorrichtung positioniert geschätzt wird, dieses Klangobjekt im Laufe des Wiedergabeschrittes (E50, E51) durch die Wiedergabebehandlung (T-A1) an dem zentralen Lautsprecher der Wiedergabevorrichtung verteilt wird.

15. Verfahren nach einem der Ansprüche 1 bis 12 und 14, bei dem, wenn ein extrahiertes Klangobjekt im Laufe des Schrittes der Raumanalyse (ΣI) als lokalisiert und als innerhalb des Wiedergabe-Raumfensters (W) der Wiedergabevorrichtung (2) an einer von dem Zentrum des Fensters (W) verschiedenen Position positioniert geschätzt wird, die Wiedergabebehandlung (T-A2), welche im Laufe des Wiedergabeschrittes (E50, E51) angewandt wird, dieses Klangobjekt an den Lautsprechern (2-1, ..., 2-N) der Wiedergabevorrichtung unter Verwendung eines Panorama-Intensitätseffektes verteilt.

16. Programm, umfassend Befehle für die Durchführung der Schritte des Wiedergabeverfahrens nach einem der Ansprüche 1 bis 15, wenn das Programm durch einen Computer oder durch einen Mikroprozessor ausgeführt wird.

17. System (1) zur Wiedergabe eines Mehrkanalaudiosignals (S) an einer mit einer Vielzahl von Lautsprechern (2-1, ..., 2-N) ausgestatteten Wiedergabevorrichtung (2), wobei die Lautsprecher an festen Orten der Wiedergabevorrichtung (2) angeordnet sind und ein Raumfenster (W) zur Tonwiedergabe in Bezug auf eine Referenzposition (Pref) definieren, wobei das Wiedergabesystem umfasst:

- Mittel (3) zur Raumanalyse des Mehrkanalaudiosignals (S), umfassend:

  • Mittel zur Extraktion wenigstens eines Klangobjektes des Signals und
  • Mittel zum Schätzen, für jedes extrahierte Klangobjekt, eines diffusen oder lokalisierten Merkmals dieses Klangobjektes und einer Position dieses Klangobjektes in Bezug auf das Tonwiedergabe-Raumfenster (W) der Wiedergabevorrichtung, und

wobei das System **gekennzeichnet ist durch**
- Mittel (4, 4A, 4B, 4C) zur Wiedergabe des Audiosignals (S) an der Vielzahl von Lautsprechern (2-1, ..., 2-N) der Wiedergabevorrichtung (2), die geeignet sind, auf jedes extrahierte Klangobjekt des Audiosignals eine Wiedergabebehandlung (T-A1, T-A2, T-B, T-C) an wenigstens einem Lautsprecher der Vielzahl von Lautsprechern der Wiedergabevorrichtung anzuwenden, wobei diese Wiedergabebehandlung von dem diffusen oder lokalisierten Merkmal des Klangobjektes und von seiner Position in Bezug auf das Tonwiedergabe-Raumfenster (W), welche **durch** die Raumanalysemittel geschätzt werden, abhängt,
wobei die Wiedergabebehandlung (T-B, T-C) das Einrichten wenigstens einer virtuellen Quelle (11, 12) außerhalb des Wiedergabe-Raumfensters (W) der Wiedergabevorrichtung anhand der Lautsprecher der Wiedergabevorrichtung umfasst, wenn das Klangobjekt **durch** die Raumanalysemittel (3) als diffus oder als außerhalb des Wiedergabe-Raumfensters der Wiedergabevorrichtung positioniert geschätzt wird.

Claims

1. A method of playing back a multichannel audio signal (S) via a playback device (2) having a plurality of loudspeakers (2-1, ..., 2-N), said loudspeakers being arranged at fixed locations of the playback device and defining a spatial window (W) for playing back sound relative to a "reference" spatial position (Pref), said playback method comprising:

• a spatial analysis step (▪ I) of spatially analyzing the multichannel audio signal (S), this step comprising:

• extracting (E30) at least one sound object from the signal; and

• for each extracted sound object, estimating (E30) a diffuse or localized nature of this sound object, and estimating a position of this sound object relative to the sound playback spatial window of the playback device;

said method being **characterized by** :

• a playback step (▪ II, E50, E51, E52, E53) of playing back the audio signal via the plurality (2-1, ..., 2-N) of loudspeakers of the playback device (2), during which step, playback treatment (T-A1, T-A2, T-B, T-C) is applied to each sound object extracted from the audio signal for playing that object back via at least one loudspeaker of the plurality of loudspeakers of the playback device (2), this playback treatment depending on the diffuse or localized nature of the sound object and on its position relative to the sound playback spatial window as estimated during the spatial analysis step;

the playback treatment (T-B, T-C) including creating at least one virtual source (11, 12) outside the playback spatial window (W) of the playback device, from the loudspeakers of the playback device, whenever the sound object is estimated during the spatial analysis step (E50) as being diffuse or as being positioned outside the playback spatial window of the playback device.

2. A method according to claim 1, wherein the playback device (2) is an acoustic enclosure (2) having said plurality of loudspeakers (2-1, ..., 2-N) arranged therein.

3. A method according to claim 1 or claim 2, wherein the spatial analysis step (·I) further comprises estimating (E40) the position of the sound object relative to the center of the sound playback spatial window (W) of the playback device (2).

4. A method according to any one of claims 1 to 3, wherein the spatial analysis step (▪ I) comprises decomposing (E20) the received audio signal into a plurality of frequency sub-bands, with said at least one sound object being extracted (E30) on at least one frequency sub-band.

5. A method according to any one of claims 1 to 4, wherein the diffuse or localized nature of the extracted sound object is estimated from at least one correlation evaluated between two distinct channels of the multichannel audio signal (S).

6. A method according to any one of claims 1 to 5, wherein the position of the extracted sound object relative to the sound playback spatial window (W) is estimated from at least one difference of level as evaluated between two distinct channels of the multichannel audio channel (S).

7. A method according to any one of claims 1 to 5, wherein the spatial analysis step (▪ I) comprises determining a Gerzon vector representative of the multichannel audio signal.

8. A method according to any one of claims 1 to 4, wherein the spatial analysis step (▪ I) comprises spatially decomposing the multichannel signal (S) into spherical harmonics.

9. A method according to any one of claims 1 to 8, wherein when an extracted sound object is estimated as being localized and as being positioned inside the sound playback spatial window (W) of the playback device, the playback treatment (T-A1, T-A2) applied to the sound object during the playback step is suitable for playing back the sound object inside the sound playback spatial window (W) of the playback device.

10. A method according to claim 9, wherein said playback treatment (T-A1, T-A2) comprises creating at least one virtual source (10) from the loudspeakers (2-1, ..., 2-N) of the playback device (2) inside the sound playback spatial window (W) of the playback device.

11. A method according to any one of claims 1 to 10, wherein when the extracted sound object is estimated during the spatial analysis step (▪ I) as being positioned outside the playback spatial window (W) of the playback device (2), creating (T-B) at least one virtual source (11) outside the playback spatial window (W) of the playback device comprises forming at least one beam that is directed to the outside of the playback spatial window.

12. A method according to any one of claims 1 to 11, wherein:

▪ the plurality of loudspeakers (2-1, ..., 2-N) of the playback device (2) comprises a central loudspeaker and lateral loudspeakers; and

- when the extracted sound object is estimated during the spatial analysis step (▪ I) as being diffuse or as being positioned outside the playback spatial window (W) of the playback device, the playback treatment (T-B, T-C) applied to the sound object uses a transaural technique for playing back the sound object via the lateral loudspeakers of the playback device.

13. A method according to any one of claims 1 to 12, wherein when an extracted sound object is estimated during the spatial analysis step (▪ I) as being localized and positioned inside the playback spatial window (W) of the playback device (2), the playback treatment (T-A1, T-A2) applied to the sound object during the playback step (E50, E51) comprises forming a beam that is directed towards said reference spatial position (Pref).

14. A method according to any one of claims 1 to 12, wherein:

- the plurality (2-1, ..., 2-N) of loudspeakers of the playback device comprises a central loudspeaker and lateral loudspeakers; and
- when an extracted sound object is estimated during the spatial analysis step (▪ I) as being localized and as being positioned at the center of the playback spatial window (W) of the playback device, the sound object is diffused during the playback step (E50, E51) by playback treatment (T-A1) via the central loudspeaker of the playback device.

15. A method according to any one of claims 1 to 12 and 14, wherein, when an extracted sound object is estimated during the spatial analysis step (▪ I) as being localized and positioned inside the playback spatial window (W) of the playback device (2) at a position that is distinct from the center of the window (W), the playback treatment (T-A2) applied during the playback step (E50, E51) diffuses this sound object via the loudspeakers (2-1, ..., 2-N) of the playback device while using an intensity panning effect.

16. A program including instructions for executing steps of the playback method according to any one of claims 1 to 15 when said program is executed by a computer or by a microprocessor.

17. A system (1) for playing back a multichannel audio signal (S) via a playback device (2) having a plurality of loudspeakers (2-1, ..., 2-N), said loudspeakers being arranged at fixed locations of the playback device (2) and defining a spatial window (W) for playing back sound relative to a reference position (Pref), said playback system comprising:

- spatial analysis means (3) for spatially analyzing the multichannel audio signal (S), these means comprising:
- extraction means for extracting at least one sound object from the signal; and
- estimation means for estimating for each extracted sound object a diffuse or localized nature of this sound object, and estimating a position of this sound object relative to the sound playback spatial window of the playback device;

said system being **characterized by** :

- playback means (4, 4A, 4B, 4C) for playing back the audio signal (S) via the plurality of loudspeakers (2-1, ..., 2-N) of the playback device (2), which means are suitable for applying playback treatment (T-A1, T-A2, T-B, T-C) to each sound object extracted from the audio signal for playing that object back via at least one loudspeaker of the plurality of loudspeakers of the playback device, this playback treatment depending on the diffuse or localized nature of the sound object and on its position relative to the sound playback spatial window (W) as estimated by the spatial analysis means;
the playback treatment (T-B, T-C) including creating at least one virtual source (11, 12) outside the playback spatial window (W) of the playback device from the loudspeakers of the playback device whenever the sound object is estimated by the spatial analysis means (3) as being diffuse or as being positioned outside the playback spatial window of the playback device.

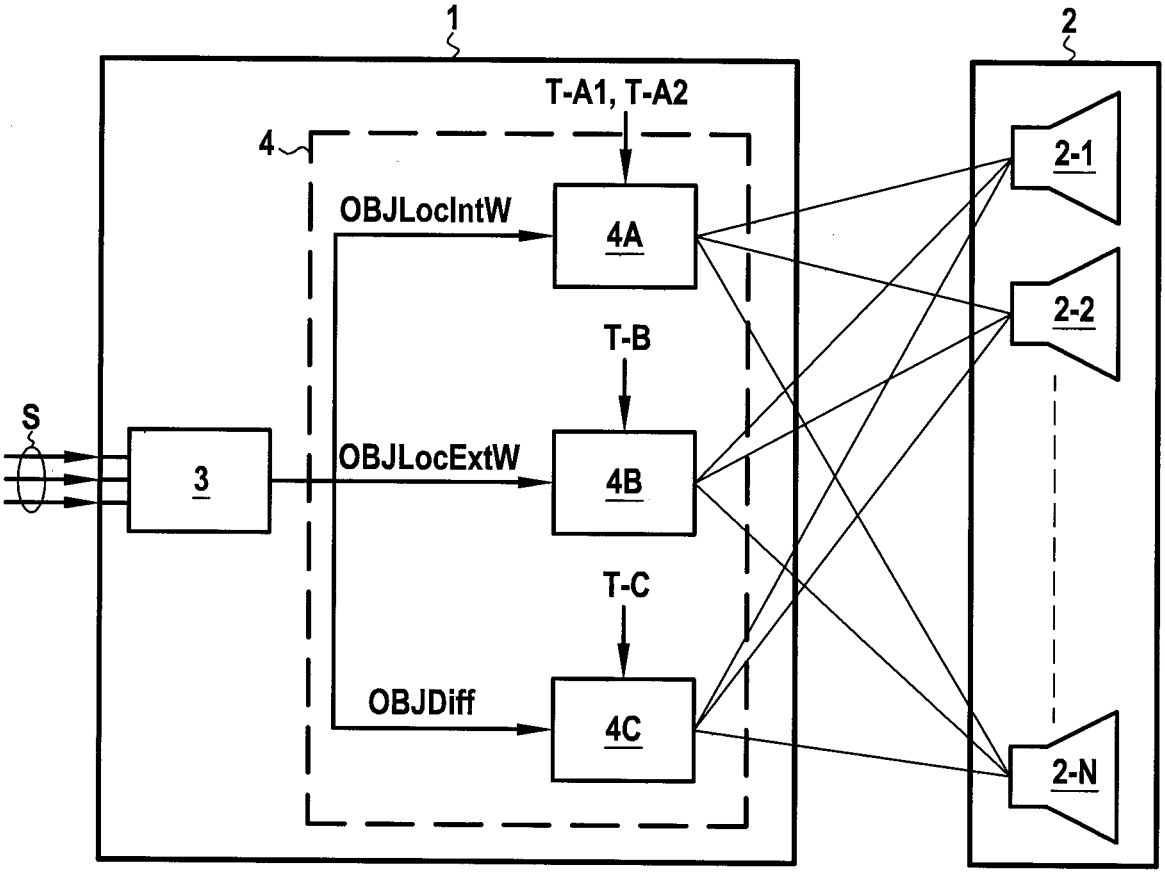
1
2
T-A1, T-A2
4
OBJLocIntW
4A
2-1
2-2
T-B
S
3
OBJLocExtW
4B
T-C
OBJDiff
4C
2-N

FIG.1

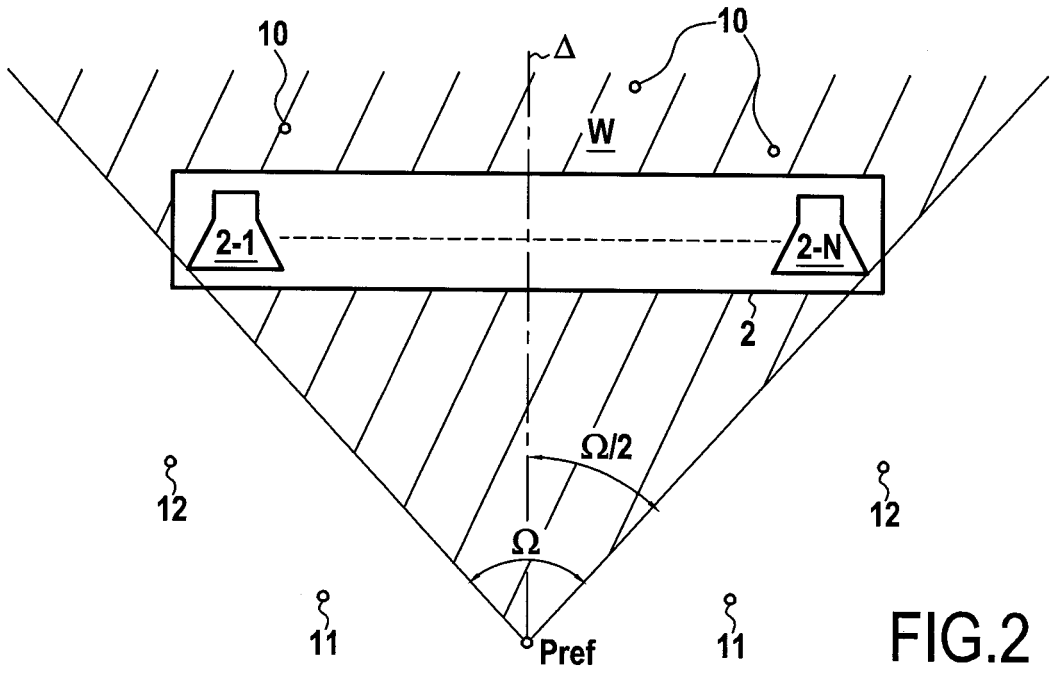
10
Δ
10
W
2-1
2-N
2
Ω/2
12
12
Ω
11
Pref
11

FIG.2

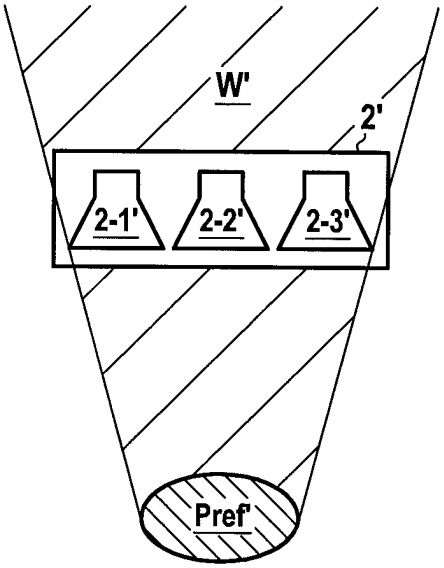

FIG.3A

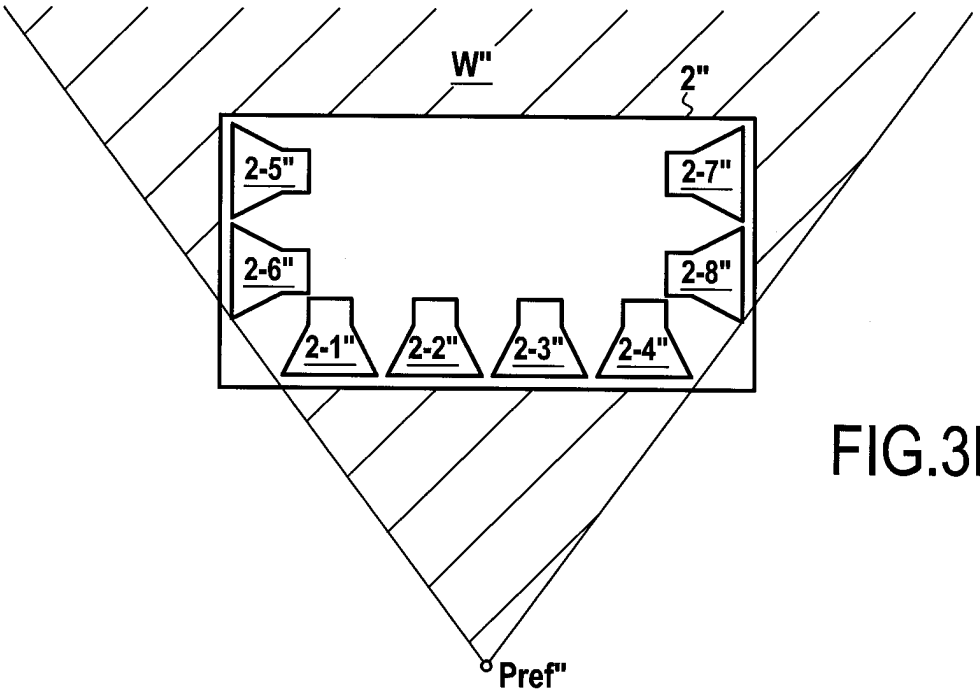

FIG.3B

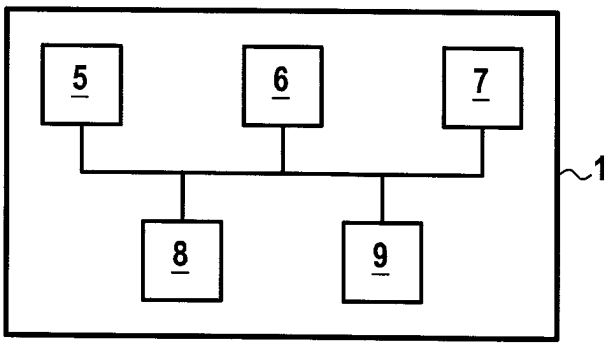

FIG.4

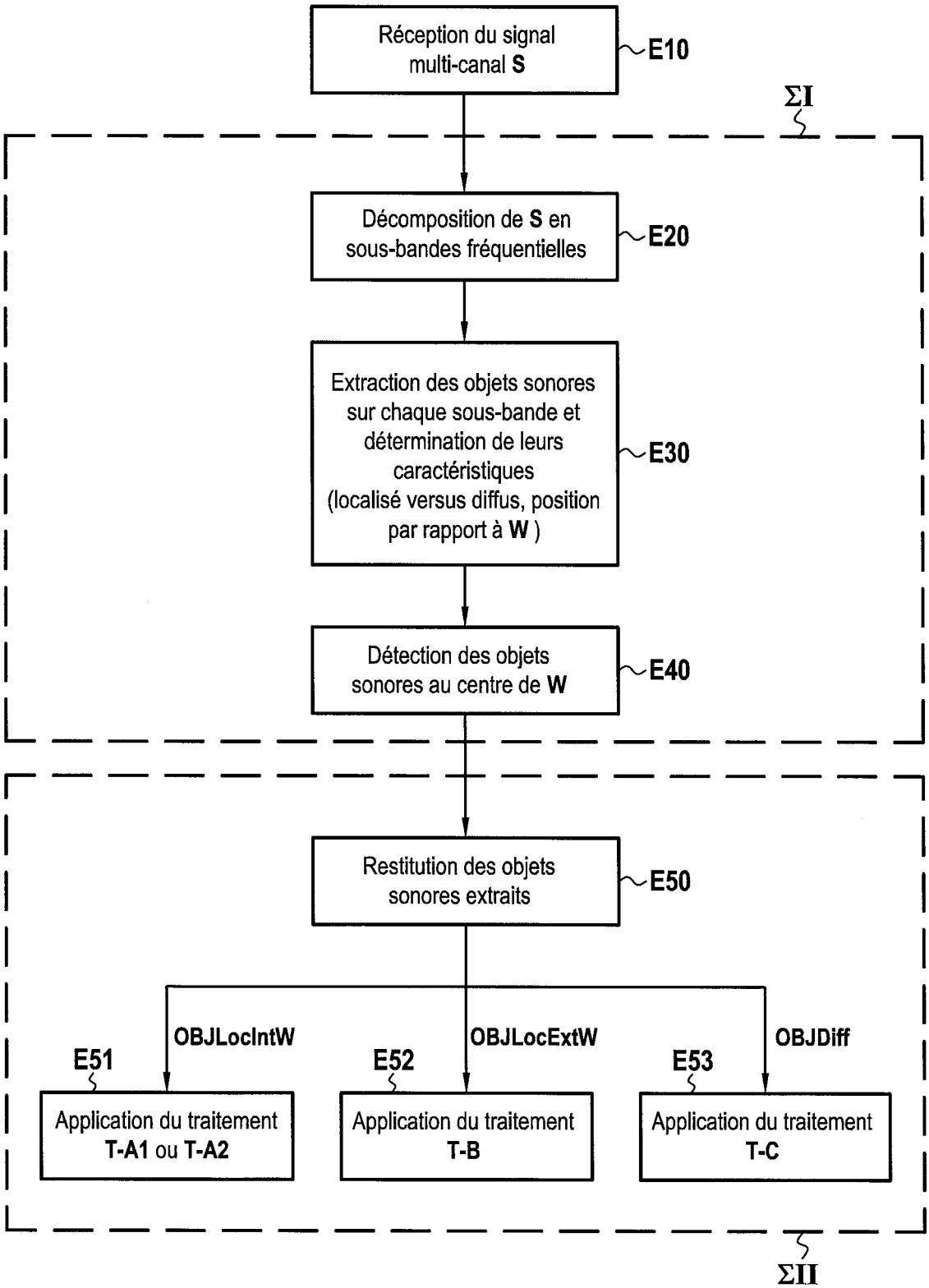
Réception du signal multi-canal S
E10
ΣI
Décomposition de S en sous-bandes fréquentielles
E20
Extraction des objets sonores sur chaque sous-bande et détermination de leurs caractéristiques (localisé versus diffus, position par rapport à W)
E30
Détection des objets sonores au centre de W
E40
Restitution des objets sonores extraits
E50
OBJLocIntW
OBJLocExtW
OBJDiff
E51
E52
E53
Application du traitement T-A1 ou T-A2
Application du traitement T-B
Application du traitement T-C
ΣII

FIG.5

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.

Documents brevets cités dans la description

- WO 2012025580 A **[0009] [0012] [0013] [0014] [0046] [0149] [0154] [0187]**
- US 20070269063 A **[0043] [0134] [0136]**
- EP 1116572 A **[0173] [0187]**
- US 2007269063 A **[0183]**

Littérature non-brevet citée dans la description

- **E. CORTEEL.** Equalization in extended area using multichannel inversion and wave field synthesis. *Journal of the Audio Engineering Society,* Décembre 2006 **[0122]**
- **J. BAUCK ; D.H. COOPER.** Generalized Transaural Stereo and Applications. *Journal Audio Engineering Society,* 1996, vol. 44 (9 **[0179]**
- **A.J. BERKHOUT et al.** A holographic approach to acoustic control. *J. Audio. Eng. Soc.,* 1988, vol. 36 **[0183]**