



(11) **EP 3 449 643 B1**

(12) **FASCICULE DE BREVET EUROPEEN**

(45) Date de publication et mention
de la délivrance du brevet:
10.06.2020 Bulletin 2020/24

(51) Int Cl.:
H04R 3/00 (2006.01) H04S 7/00 (2006.01)
H04R 1/40 (2006.01)

(21) Numéro de dépôt: **17725294.7**

(86) Numéro de dépôt international:
PCT/FR2017/050935

(22) Date de dépôt: **20.04.2017**

(87) Numéro de publication internationale:
WO 2017/187053 (02.11.2017 Gazette 2017/44)

(54) **PROCÉDÉ ET SYSTÈME DE DIFFUSION D'UN SIGNAL AUDIO À 360°**

VERFAHREN UND SYSTEM ZUM SENDEN EINES 360°-AUDIOSIGNALS

METHOD AND SYSTEM OF BROADCASTING A 360° 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: **26.04.2016 FR 1653684**

(43) Date de publication de la demande:
06.03.2019 Bulletin 2019/10

(73) Titulaire: **Arkamys
75017 Paris (FR)**

(72) Inventeurs:
• **DEVALLEZ, Delphine
92100 BOULOGNE-BILLANCOURT (FR)**

• **AMADU, Frédéric
77500 CHELLES (FR)**

(74) Mandataire: **Ipside
7-9 Allées Haussmann
33300 Bordeaux Cedex (FR)**

(56) Documents cités:
**WO-A2-2005/015954 US-A- 6 021 206
US-A1- 2012 093 344**

• **LABORIE A ET AL: "A New Comprehensive
Approach of Surround Sound Recording", AUDIO
ENGINEERING SOCIETY CONVENTION PAPER,
NEW YORK, NY, US, 22 mars 2003 (2003-03-22),
pages 1-19, XP002280618,**

EP 3 449 643 B1

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).

Description**Domaine de l'invention**

[0001] La présente invention se rapporte au domaine du traitement du signal sonore.

Etat de la technique

[0002] On connaît dans l'état de la technique des procédés et des systèmes permettant de diffuser des signaux vidéo à 360°. Il existe un besoin dans l'état de la technique de pouvoir associer des signaux audio à ces signaux vidéo à 360°.

[0003] Jusqu'à maintenant, l'audio 3D était réservé aux professionnels du son et aux chercheurs. Cette technologie a pour objectif de capturer le maximum d'informations spatiales lors de l'enregistrement pour les restituer ensuite à l'auditeur et lui donner une sensation d'immersion dans la scène sonore. Dans le domaine de la vidéo, l'intérêt est croissant pour les vidéos filmées à 360° et reproduites via un casque de réalité virtuelle pour une immersion complète dans l'image : l'utilisateur peut tourner la tête et explorer la scène visuelle tout autour de lui. Pour obtenir la même fidélité dans le domaine du son, la solution la plus compacte est l'utilisation d'un réseau de microphones, comme par exemple l'Eigenmike de mh acoustics, le Soundfield de TSL Products, et le TetraMic de Core Sound. Le document WO 2005/015954 montre un procédé pour convertir les signaux d'un réseau de microphones placés selon une configuration sphérique en un format ambisonique. Equipés de quatre à trente-deux microphones, ces produits sont onéreux et donc réservés à des utilisations professionnelles. De récentes recherches ont permis de réduire le nombre de microphones (Palacino, J. D., & Nicol, R. (2013). "Spatial sound pick-up with a low number of microphones." ICA 2013. Montréal, Canada.), et des microphones de taille et de coût réduits peuvent être utilisés tels que ceux dont disposent les téléphones portables. La forme des réseaux de microphones, un polyèdre, reste cependant standardisée, du dodécaèdre pour l'EigenMike au tétraèdre pour le Soundfield et le TetraMic. Cette forme géométrique permet d'utiliser des formules simples pour convertir les signaux des microphones en un format ambisonique, et ont été développées par Gerzon en 1975 (Gerzon, M. (1975). « The design of precisely coincident microphone arrays for stereo and surround sound. » 50th Audio Engineering Society Conference.). Le format ambisonique est un ensemble de canaux audio qui contient toutes les informations nécessaires à la reconstruction spatiale du champ sonore. Une nouveauté apportée par le présent brevet est la possibilité d'utiliser une forme quelconque de réseau de microphones. Ainsi, il est tout à fait possible d'utiliser une forme déjà existante, telle une caméra 360° ou un téléphone portable, pour y inclure un certain nombre de microphones. On obtient alors un système complet et compact d'enregistrement du son et

de l'image à 360°.

Exposé de l'invention

[0004] La présente invention entend remédier aux inconvénients de l'art antérieur en proposant un procédé de traitement du signal sonore permettant de réaliser une captation du signal sonore dans toutes les directions, puis de restituer ledit signal sonore.

[0005] A cet effet, la présente invention concerne, dans son acception la plus générale, un procédé de traitement du signal sonore, conformément à la revendication 1.

[0006] Ainsi, grâce au procédé selon la présente invention, il est possible de réaliser une captation du signal sonore dans toutes les directions, puis de restituer ledit signal sonore.

[0007] Avantageusement, le calcul matriciel fait intervenir une matrice H calculée par la méthode des moindres carrés à partir des directivités mesurées des N microphones et des directivités idéales des composantes ambisoniques.

[0008] Selon un mode de réalisation, lesdits microphones sont disposés selon un cercle sur un plan, espacés suivant un angle égal à 360°/N ou à chaque coin d'un téléphone portable.

[0009] Selon un mode de réalisation, ledit procédé met en œuvre quatre microphones espacés suivant un angle de 90° à l'horizontal.

[0010] Selon un mode de réalisation, ledit procédé met en œuvre un filtre passe bande filtrant de 100 Hz à 6 kHz.

[0011] Selon un mode de réalisation, l'ordre R du format de type ambisonique est égal à un.

[0012] Avantageusement, au cours de ladite étape de restitution, une information relative à l'orientation de la tête d'un utilisateur écoutant le signal sonore, est exploitée.

[0013] De préférence, une captation de ladite information relative à l'orientation de la tête d'un utilisateur écoutant le signal sonore, est réalisée par un capteur au sein d'un téléphone mobile ou bien par un capteur situé dans un casque d'écoute ou un casque de réalité virtuelle.

[0014] Selon un mode de réalisation, au cours de ladite étape de restitution, les données sous format ambisonique sont transformées en données sous format binaural.

[0015] La présente invention se rapporte également à un système de traitement du signal sonore, conformément à la revendication 9.

Brève description des dessins

[0016] On comprendra mieux l'invention à l'aide de la description, faite ci-après à titre purement explicatif, d'un mode de réalisation de l'invention, en référence aux Figures dans lesquelles :

- Les Figures 1 et 3 représentent les différentes étapes du procédé selon la présente invention ;
- La Figure 2 illustre les traitements appliqués dans le

cadre de la seconde étape du procédé selon la présente invention ;

- Les Figures 4a, 4b et 4c représentent les composantes W, Y et X idéales d'un format ambisonique d'ordre 1 (sur plan horizontal) ;
- Les Figures 5a, 5b et 5c illustrent les composantes W, Y et X approximées d'un format ambisonique d'ordre 1 ; et
- La Figure 6 représente le placement de huit haut-parleurs virtuels, chacun placé à 45° autour d'un utilisateur.

Description détaillée des modes de réalisation de l'invention

[0017] La présente invention se rapporte à un procédé de traitement du signal sonore, conformément à la revendication 1.

[0018] Les Figures 1 et 3 illustrent les différentes étapes du procédé selon la présente invention.

[0019] Dans un mode de réalisation, lesdits microphones sont disposés selon un cercle sur un plan, espacés suivant un angle égal à 360°/N ou à chaque coin d'un téléphone portable.

[0020] Dans un mode de réalisation, le procédé selon la présente invention met en œuvre quatre microphones espacés suivant un angle de 90° à l'horizontal.

[0021] Dans un mode de réalisation, l'ordre R du format de type ambisonique est égal à un.

[0022] La première étape du procédé selon la présente invention consiste en l'enregistrement du signal sonore. On utilise pour cet enregistrement N microphones, N étant un entier naturel supérieur ou égal à trois, lesdits microphones étant disposés selon un cercle sur un plan, espacés suivant un angle égal à 360°/N ou à chaque coin d'un téléphone portable. Dans l'exemple de réalisation décrit ci-après, N est égal à quatre et les microphones sont espacés de 90°. Ces microphones sont disposés selon un cercle sur un plan. Dans un exemple particulier de mise en œuvre, le rayon dudit cercle est de deux centimètres, et les microphones sont omnidirectionnels.

[0023] Le signal sonore est capté par lesdits microphones, et numérisé. Il s'agit d'une captation synchrone.

[0024] On obtient à l'issue de cette première étape quatre signaux numériques échantillonnés.

[0025] La seconde étape du procédé selon la présente invention consiste en l'encodage desdits quatre signaux numériques échantillonnés, en un format de type ambisonique d'ordre R, R étant un entier naturel supérieur ou égal à un.

[0026] On rappelle ici que le format ambisonique est un format normalisé de codage audio en plusieurs dimensions.

[0027] Dans l'exemple de réalisation décrit ci-après, l'ordre R est égal à un. Cet ordre 1 permet de représenter le son avec les notions suivantes : Avant - Arrière et Gauche - Droite.

[0028] Les Figures 4a, 4b et 4c représentent les composantes W, Y et X idéales d'un format ambisonique d'ordre 1 (sur plan horizontal).

[0029] Les Figures 5a, 5b et 5c illustrent les composantes W, Y et X approximées d'un format ambisonique d'ordre 1.

[0030] La Figure 2 illustre les traitements appliqués dans le cadre de la seconde étape du procédé selon la présente invention.

[0031] On observe sur la Figure 2 que les données en entrée sont dans le domaine temporel, passent dans le domaine fréquentiel suite à une opération de transformée de Fourier rapide (FFT ou « *Fast Fourier Transform* » en terminologie anglo-saxonne), puis les données en sortie sont dans le domaine temporel suite à une opération de transformée de Fourier rapide inverse (IFFT ou « *Inverse Fast Fourier Transform* » en terminologie anglo-saxonne).

[0032] De préférence, on utilise des fenêtres de Hanning avec un recouvrement en mettant en œuvre une fonction de type « *overlap-add* ».

[0033] On observe également sur la Figure 2 que les données fréquentielles d'entrées sont modifiées à l'aide d'une multiplication matricielle. Cette matrice comporte des coefficients pondérateurs pour chaque signal de microphone et chaque fréquence.

[0034] On observe également sur la Figure 2 qu'un filtrage au moyen d'un filtre passe-bande est réalisé sur les données avant la sortie.

[0035] Dans un mode de réalisation, le procédé selon la présente invention met en œuvre un filtre passe bande filtrant de 100 Hz à 6 kHz. On élimine ainsi la partie basse et la partie aigüe.

[0036] Afin de calculer les coefficients de la matrice de pondération, on mesure des réponses impulsionnelles des N microphones, dans le cas présent des quatre microphones, avec une source positionnée tous les 5° ou tous les 10° autour du réseau de microphones.

[0037] A l'aide d'une transformée de Fourier rapide, on obtient les réponses en fréquence des N microphones en fonction des angles mesurés, ou autrement dit les directivités des N microphones en fonction de la fréquence.

[0038] Il est possible d'utiliser à ce stade les principes du procédé décrit dans la demande internationale publiée sous le numéro WO 2015/128160 « *Procédé et système d'égalisation acoustique automatisé* » pour égaliser les réponses en fréquence sur l'axe de chacun des microphones. Les mêmes filtres d'égalisation sont appliqués à tous les microphones et pour toutes les positions angulaires de source.

[0039] Les réponses des microphones sont ensuite placées dans une matrice C.

[0040] Dans le domaine fréquentiel, pour chaque index de fréquence k, on a

$$C_{D \times N} \cdot H_{N \times V} = P_{D \times V}$$

où N est le nombre de microphones (quatre dans le présent exemple de réalisation), D est le nombre de positions angulaires de source mesurées (108 dans le présent exemple de réalisation) et V est le nombre de canaux ambisoniques (trois dans le présent exemple de réalisation), $C_{D \times N}$ désigne les directivités des microphones, $H_{N \times V}$ désigne la matrice qui transforme les directivités des microphones en les directivités désirées, et $P_{D \times V}$ désigne les directivités prescrites par le format ambisonique (W, X et Y dans le présent exemple de réalisation).

[0041] On a ainsi $H_{N \times V} = P_{D \times V} / C_{D \times N}$ pour chaque index de fréquence k si $C_{D \times N}$ est inversible.

[0042] En pratique, $C_{D \times N}$ n'est pas inversible. Dans un mode de réalisation, on met en œuvre une méthode des moindres carrés pour résoudre $C_{108 \times 4} \cdot H_{4 \times 3} = P_{108 \times 3}$

[0043] La matrice H est définie une fois pour les utilisations futures du réseau de microphones considéré. Ensuite, à chaque utilisation on réalise une multiplication matricielle dans le domaine fréquentiel.

[0044] Ladite matrice H possède autant de lignes que de microphones donc quatre dans le présent exemple de réalisation, et autant de colonnes que le requiert l'ordre du format ambisonique utilisé, donc trois colonnes dans le présent exemple de réalisation dans lequel l'ordre 1 est mis en œuvre sur le plan horizontal.

[0045] On a $Out = In \times H$, où H désigne la matrice précédemment calculée, In désigne l'entrée (canaux audio provenant du réseau de microphones, passés dans le domaine fréquentiel) et Out désigne la sortie (Out étant reconverti dans le domaine temporel pour obtenir le format ambisonique).

[0046] Le procédé selon la présente invention met en œuvre, au cours de cette seconde étape, un algorithme dit algorithme des moindres carrés pour chaque fréquence, avec par exemple 512 points de fréquence.

[0047] A l'issue de cette seconde étape, on obtient des données sous le format ambisonique (dans le présent exemple de réalisation les signaux W, X et Y).

[0048] La troisième étape du procédé selon la présente invention consiste en la restitution du signal sonore, grâce à une transformation des données sous format ambisonique en deux canaux binauraux.

[0049] Au cours de cette troisième étape, l'information relative à l'orientation de la tête de l'utilisateur écoutant le signal sonore, est récupérée et exploitée. Ceci peut être réalisé par un capteur au sein d'un téléphone mobile, d'un casque d'écoute ou d'un casque de réalité virtuelle.

[0050] Cette information sur l'orientation consiste en un vecteur comportant trois valeurs d'angles, en terminologie anglo-saxonne « *pitch* », « *yaw* » et « *roll* ».

[0051] Dans le présent exemple de réalisation, sur un plan, on utilise la valeur d'angle « *yaw* ».

[0052] On transforme le format ambisonique en huit canaux audio correspondant à un placement virtuel de huit haut-parleurs, chacun placé à 45° autour de l'utilisateur.

[0053] La Figure 6 représente le placement de huit

haut-parleurs virtuels, chacun placé à 45° autour d'un utilisateur.

[0054] Chaque haut parleur virtuel restitue un signal audio issu des composantes ambisoniques selon la formule:

$$P_n = W + X \cos \theta_n + Y \sin \theta_n \quad (1)$$

où W, X et Y sont les données relatives au format ambisonique, et où θ_n représente l'angle horizontal du nième haut-parleur. Par exemple, dans le présent exemple de réalisation $\theta_0 = 0^\circ$, $\theta_1 = 45^\circ$, $\theta_2 = 90^\circ$, etc.

[0055] Ensuite, on effectue une étape de filtrage avec une paire de HRTF par haut-parleur, HRTF signifiant « *Head-related transfer fonction* » en terminologie anglo-saxonne. On associe une paire de filtres HRTF (oreille gauche et oreille droite) à chaque haut-parleur virtuel, puis on additionne (tous les canaux « oreille gauche » et tous les canaux « oreille droite » ensemble), afin de former deux canaux de sortie.

[0056] Des coefficients IIR (« *Infinite Impulse Response* ») sont mis en œuvre à ce stade, lesdits filtres HRTF étant modélisés sous formes de filtres IIR.

[0057] Lorsque l'utilisateur tourne la tête, la position des haut-parleurs virtuels est modifiée. Par exemple pour une rotation de la tête d'un angle α , l'angle des haut-parleurs virtuels devient $\beta_n = \theta_n - \alpha$. On remplace alors θ_n par $(\theta_n - \alpha)$ dans la formule (1) pour calculer le signal restitué par le nième haut-parleur virtuel.

[0058] Ainsi, grâce au procédé selon la présente invention, il est possible de réaliser une captation du signal sonore dans toutes les directions, puis de restituer ledit signal sonore.

[0059] La Figure 3 représente les différentes étapes du procédé selon la présente invention.

[0060] La présente invention se rapporte également à un système de traitement du signal sonore, conformément à la revendication 9.

[0061] Ce système de traitement du signal sonore comprend au moins une unité de calcul et une unité de mémoire.

[0062] L'invention est décrite dans ce qui précède à titre d'exemple. Il est entendu que l'homme du métier est à même de réaliser différentes variantes de l'invention sans pour autant sortir du cadre du brevet.

Revendications

1. Procédé de traitement du signal sonore, comportant les étapes suivantes :

- Captation de façon synchrone d'un signal sonore d'entrée ($S_{\text{entrée}}$) à l'aide de N microphones, N étant un entier naturel supérieur ou égal à trois;
- Encodage dudit signal sonore d'entrée

- ($S_{\text{entrée}}$) en un format de données (D) de son, ledit encodage comportant une sous-étape de transformation dudit signal d'entrée en un format de type ambisonique d'ordre R, R étant un entier naturel supérieur ou égal à un, ladite sous-étape de transformation en un format de type ambisonique étant mise en œuvre à l'aide d'une transformée de Fourier rapide, d'une multiplication matricielle, d'une transformée de Fourier rapide inverse et à l'aide d'un filtre passe-bande ; et
- Restitution d'un signal sonore de sortie (S_{sortie}) à l'aide d'un traitement numérique desdites données (D) de son ;
- et **caractérisé en ce que** la multiplication matricielle fait intervenir une matrice H calculée par la méthode des moindres carrés à partir des directivités mesurées des N microphones et des directivités idéales des composantes ambisoniques.
- Procédé de traitement du signal sonore selon la revendication 1, **caractérisé en ce que** lesdits microphones sont disposés selon un cercle sur un plan, espacés suivant un angle égal à $360^\circ/N$, ou à chaque coin d'un téléphone portable.
 - Procédé de traitement du signal sonore selon la revendication 2, **caractérisé en ce qu'il** met en œuvre quatre microphones espacés suivant un angle de 90° à l'horizontal.
 - Procédé de traitement du signal sonore selon l'une des revendications précédentes, **caractérisé en ce qu'il** met en œuvre un filtre passe bande filtrant de 100 Hz à 6 kHz.
 - Procédé de traitement du signal sonore selon l'une des revendications précédentes, **caractérisé en ce que** l'ordre R du format de type ambisonique est égal à un.
 - Procédé de traitement du signal sonore selon l'une des revendications précédentes, **caractérisé en ce qu'**au cours de ladite étape de restitution, une information relative à l'orientation de la tête d'un utilisateur écoutant le signal sonore, est exploitée.
 - Procédé de traitement du signal sonore selon la revendication 6, **caractérisé en ce qu'**une captation de ladite information relative à l'orientation de la tête d'un utilisateur écoutant le signal sonore, est réalisée par un capteur au sein d'un téléphone, d'un casque d'écoute ou d'un casque de réalité virtuelle.
 - Procédé de traitement du signal sonore selon l'une des revendications précédentes, **caractérisé en ce qu'**au cours de ladite étape de restitution, les don-

nées sous format ambisonique sont transformées en données sous format binaural.

- Système de traitement du signal sonore, comportant des moyens pour :

- Capturer de façon synchrone un signal sonore d'entrée ($S_{\text{entrée}}$) à l'aide de N microphones, N étant un entier naturel supérieur ou égal à trois ;
- Encoder ledit signal sonore d'entrée ($S_{\text{entrée}}$) en un format de données (D) de son, et des moyens pour transformer ledit signal d'entrée en un format de type ambisonique d'ordre R, R étant un entier naturel supérieur ou égal à un, lesdits moyens pour transformer en un format de type ambisonique étant mis en œuvre à l'aide d'une transformée de Fourier rapide, d'une multiplication matricielle, d'une transformée de Fourier rapide inverse et à l'aide d'un filtre passe-bande ; et
- Restituer un signal sonore de sortie (S_{sortie}) à l'aide d'un traitement numérique desdites données (D) de son ;

et **caractérisé en ce que** la multiplication matricielle fait intervenir une matrice H calculée par la méthode des moindres carrés à partir des directivités mesurées des N microphones et des directivités idéales des composantes ambisoniques.

Patentansprüche

- Verfahren zur Verarbeitung des Audiosignals, die folgenden Schritte beinhaltend:
 - Synchrones Erfassen eines Eingangsaudiosignals ($S_{\text{entrée}}$) mithilfe von N Mikrofonen, wobei N eine natürliche ganze Zahl größer oder gleich drei ist;
 - Codieren des Eingangsaudiosignals ($S_{\text{entrée}}$) in ein Tondatenformat (D), wobei das Codieren einen Unterschritt des Umwandels des Eingangssignals in ein Raumklangformat R. Ordnung ist, wobei R eine natürliche ganze Zahl größer oder gleich eins ist, wobei der Unterschritt des Umwandels in ein Raumklangformat mithilfe einer schnellen Fourier-Transformation, einer Matrixmultiplikation, einer schnellen umgekehrten Fourier-Transformation, und mithilfe eines Bandpassfilters durchgeführt wird; und
 - Wiedergeben eines Ausgangsaudiosignals (S_{sortie}) mithilfe einer digitalen Verarbeitung der Tondaten (D);

und **dadurch gekennzeichnet, dass** die Matrixmultiplikation eine Matrix H einbeziehen lässt, die durch die Methode der kleinsten Quadrate aus den gemess-

senen Richtwirkungen der N Mikrofone und den idealen Richtwirkungen der Raumklangkomponenten berechnet wird.

2. Verfahren zur Verarbeitung des Audiosignals nach Anspruch 1, **dadurch gekennzeichnet, dass** die Mikrofone in einem Kreis auf einer Ebene, entsprechend einem Winkel gleich $360^\circ/N$ beabstandet, oder an jeder Ecke eines tragbaren Telefons angeordnet sind. 5
10
3. Verfahren zur Verarbeitung des Audiosignals nach Anspruch 2, **dadurch gekennzeichnet, dass** es vier Mikrofone anwendet, die entsprechend einem Winkel von 90° zur Horizontalen beabstandet sind. 15
4. Verfahren zur Verarbeitung des Audiosignals nach einem der vorstehenden Ansprüche, **dadurch gekennzeichnet, dass** es ein Bandpassfilter anwendet, das von 100 Hz bis 6 kHz filtert. 20
5. Verfahren zur Verarbeitung des Audiosignals nach einem der vorstehenden Ansprüche, **dadurch gekennzeichnet, dass** die R. Ordnung des Raumklangsformats gleich eins ist. 25
6. Verfahren zur Verarbeitung des Audiosignals nach einem der vorstehenden Ansprüche, **dadurch gekennzeichnet, dass** im Laufe des Wiedergabeschrittes eine Information in Bezug auf die Ausrichtung des Kopfes eines Nutzers, der das Audiosignal anhört, ausgewertet wird. 30
7. Verfahren zur Verarbeitung des Audiosignals nach Anspruch 6, **dadurch gekennzeichnet, dass** das Erfassen der Information in Bezug auf die Ausrichtung des Kopfes eines Nutzers, der das Audiosignal anhört, durch einen Sensor innerhalb eines Telefons, eines Kopfhörers oder einer Brille für virtuelle Realität realisiert wird. 35
40
8. Verfahren zur Verarbeitung des Audiosignals nach einem der vorstehenden Ansprüche, **dadurch gekennzeichnet, dass** im Laufe des Wiedergabeschrittes die Daten im Raumklangsformat in Daten im Binärformat umgewandelt werden. 45
9. System zur Verarbeitung des Audiosignals, Mittel beinhaltend zum: 50
 - Synchronen Erfassen eines Eingangsaudiosignals ($S_{\text{entrée}}$) mithilfe von N Mikrofonen, wobei N eine natürliche ganze Zahl größer oder gleich drei ist;
 - Codieren des Eingangsaudiosignals ($S_{\text{entrée}}$) in ein Tondatenformat (D), wobei das Codieren einen Unterschritt des Umwandelns des Eingangssignals in ein Raumklangsformat R. Ord-

nung ist, wobei R eine natürliche ganze Zahl größer oder gleich eins ist, wobei die Mittel zum Umwandeln in ein Raumklangsformat mithilfe einer schnellen Fourier-Transformation, einer Matrixmultiplikation, einer schnellen umgekehrten Fourier-Transformation, und mithilfe eines Bandpassfilters angewendet werden; und

- Wiedergeben eines Ausgangsaudiosignals (S_{sortie}) mithilfe einer digitalen Verarbeitung der Tondaten (D);

und **dadurch gekennzeichnet, dass** die Matrixmultiplikation eine Matrix H einbeziehen lässt, die durch die Methode der kleinsten Quadrate aus den gemessenen Richtwirkungen der N Mikrofone und den idealen Richtwirkungen der Raumklangkomponenten berechnet wird.

Claims

1. Sound signal processing method, comprising the following steps:
 - synchronously acquiring an input sound signal ($S_{\text{entrée}}$) by means of N microphones, N being a natural number greater than or equal to three;
 - encoding said input sound signal ($S_{\text{entrée}}$) in a sound data format (D), said encoding comprising a sub-step of transforming said input signal into an ambisonic-type format of order R, R being a natural number greater than or equal to one, said sub-step of transforming into an ambisonic-type format being carried out by means of a Fast Fourier Transform, a matrix multiplication, an Inverse Fast Fourier Transform and by means of a band-pass filter; and
 - delivering an output sound signal (S_{sortie}) by means of digitally processing said sound data (D);

and **characterized in that** the matrix multiplication uses a matrix H calculated by the method of least squares from measured directivities of the N microphones and ideal directivities of the ambisonic components.
2. Sound signal processing method according to claim 1, **characterised in that** said microphones are positioned in a circle on a plane, spaced apart by an angle equal to $360^\circ/N$ or at each corner of a mobile phone.
3. Sound signal processing method according to claim 2, **characterised in that** it implements four microphones spaced apart by an angle of 90° to the horizontal.

4. Sound signal processing method according to any of the previous claims, **characterised in that** it implements a band-pass filter filtering frequencies from 100 Hz to 6 kHz. 5
5. Sound signal processing method according to any of the previous claims, **characterised in that** the order R of the ambisonic-type format is equal to one. 10
6. Sound signal processing method according to any of the previous claims, **characterised in that**, during said delivering step, an information item relative to the orientation of the head of a user listening to the sound signal, is exploited. 15
7. Sound signal processing method according to claim 6, **characterised in that** acquisition of said information item relative to the orientation of the head of a user listening to the sound signal, is carried out by a sensor in a telephone, an audio headset or a virtual reality headset. 20
8. Sound signal processing method according to any of the previous claims, **characterised in that**, during said delivering step, the data in ambisonic format is transformed into data in binaural format. 25
9. Sound signal processing system, comprising means for: 30
 - synchronously acquiring an input sound signal ($S_{\text{entrée}}$) by means of N microphones, N being a natural number greater than or equal to three;
 - encoding said input sound signal ($S_{\text{entrée}}$) in a sound data format (D), and means for transforming said input signal into an ambisonic-type format of order R, R being a natural number greater than or equal to one, said means for transformation into an ambisonic-type format being carried out by means of a Fast Fourier Transform, a matrix multiplication, an Inverse Fast Fourier Transform and by means of a band-pass filter; 35
 - and
 - delivering an output sound signal (S_{sortie}) by means of digitally processing the said sound data (D); 40

and **characterized in that** the matrix multiplication uses a matrix H calculated by the method of least squares from measured directivities of the N microphones and ideal directivities of the ambisonic components. 45

55

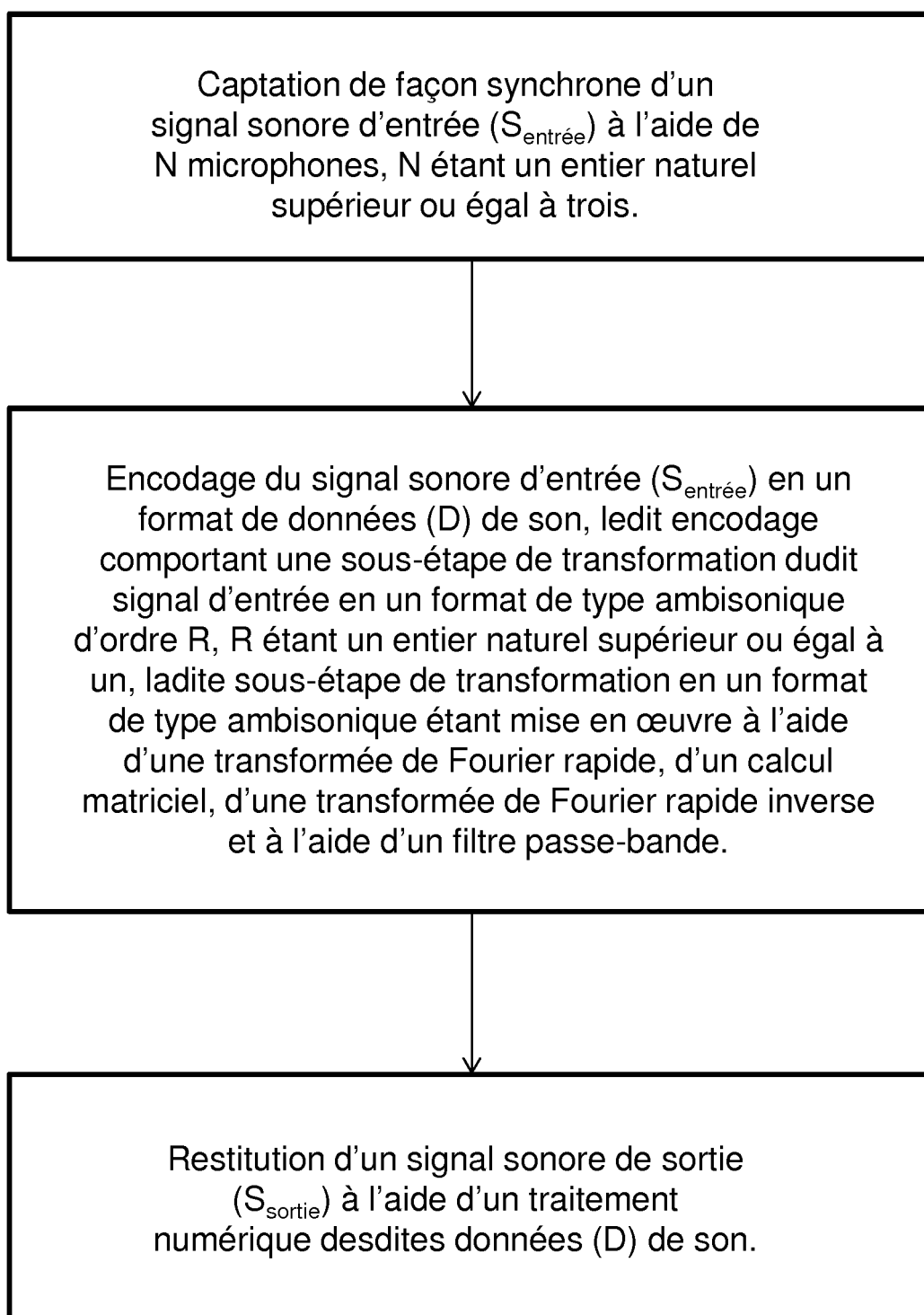


Figure 1

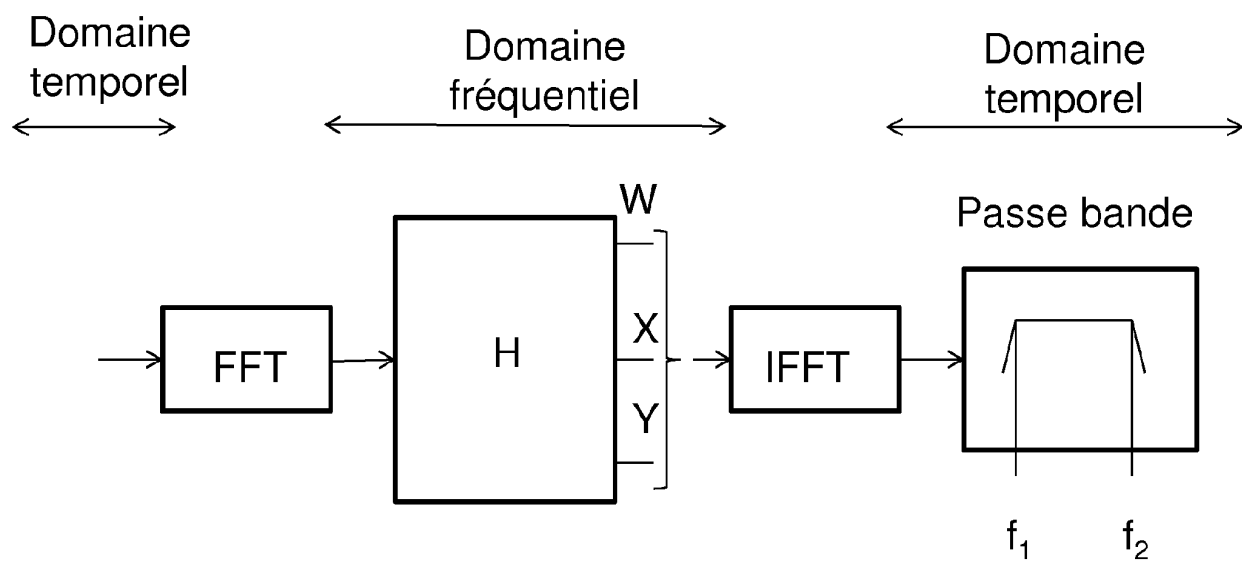


Figure 2

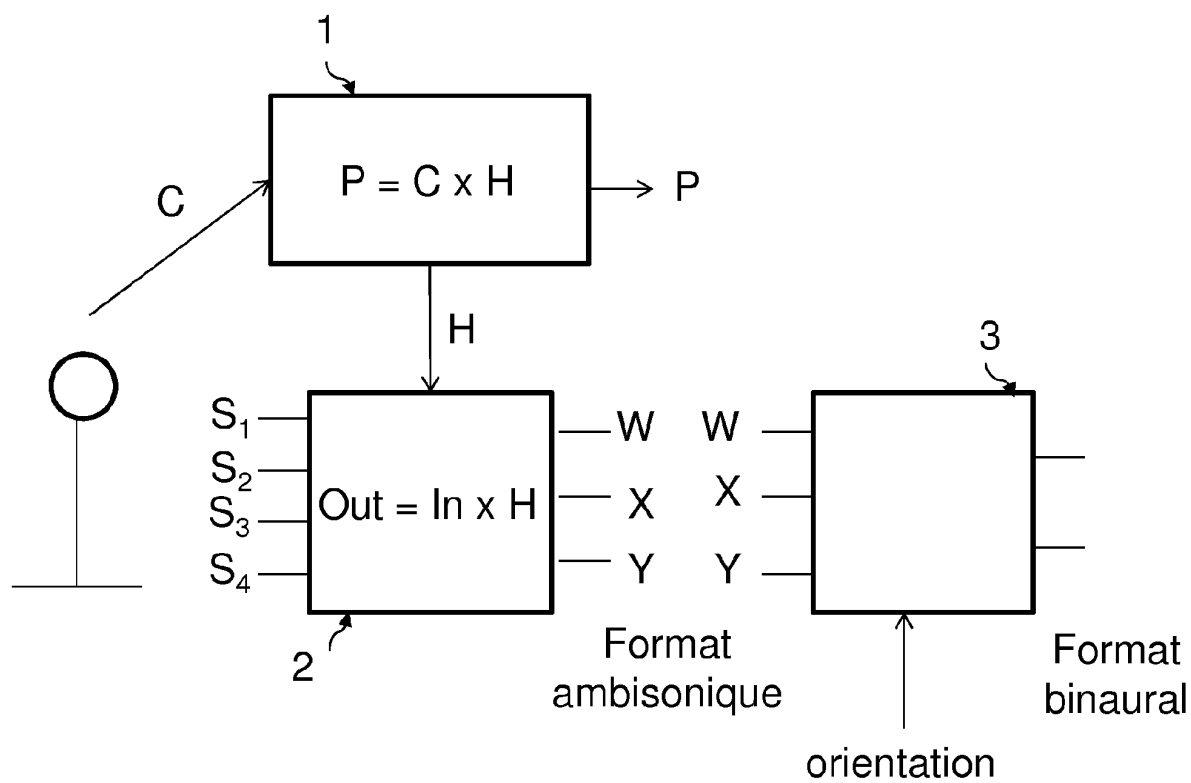
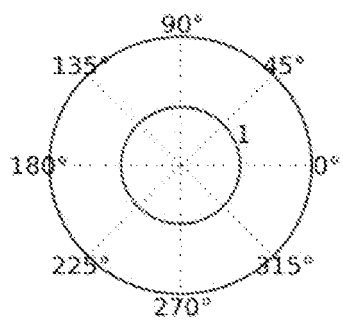
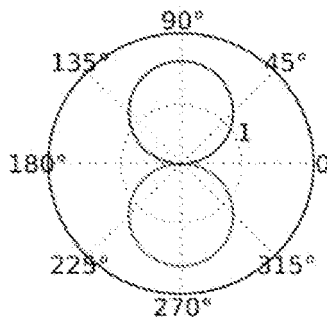


Figure 3



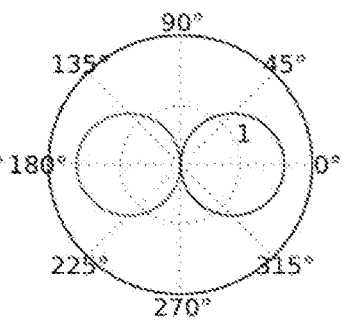
W

Figure 4a



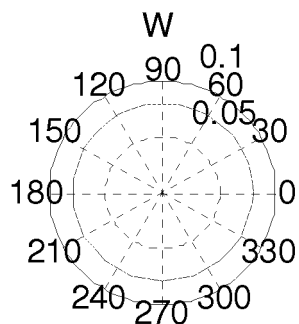
Y

Figure 4b



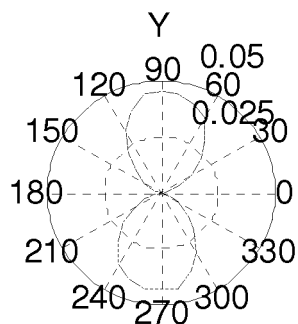
X

Figure 4c



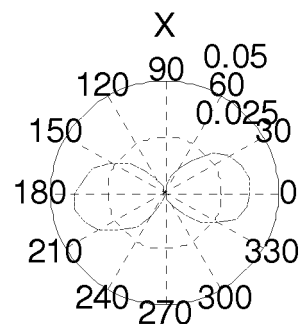
W

Figure 5a



Y

Figure 5b



X

Figure 5c

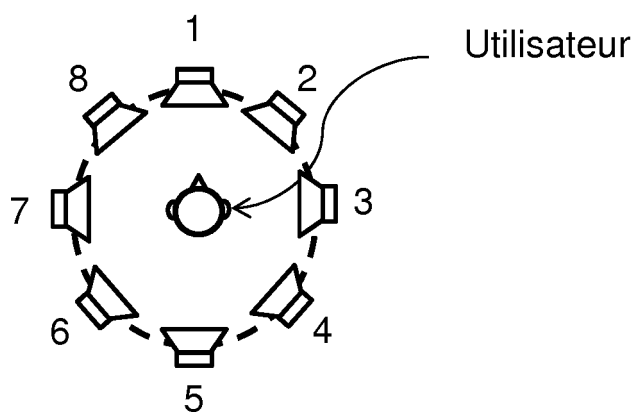


Figure 6

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 2005015954 A [0003]
- WO 2015128160 A [0038]

Littérature non-brevet citée dans la description

- **PALACINO, J. D. ; NICOL, R.** Spatial sound pick-up with a low number of microphones. *ICA 2013*, 2013 [0003]
- **GERZON, M.** The design of precisely coincident microphone arrays for stereo and surround sound. *50th Audio Engineering Society Conference*, 1975 [0003]