

12

DEMANDE DE BREVET D'INVENTION

A1

22 Date de dépôt : 24.07.13.

30 Priorité :

43 Date de mise à la disposition du public de la demande : 30.01.15 Bulletin 15/05.

56 Liste des documents cités dans le rapport de recherche préliminaire : *Se reporter à la fin du présent fascicule*

60 Références à d'autres documents nationaux apparentés :

○ Demande(s) d'extension :

71 Demandeur(s) : **ORANGE Société anonyme** — FR.

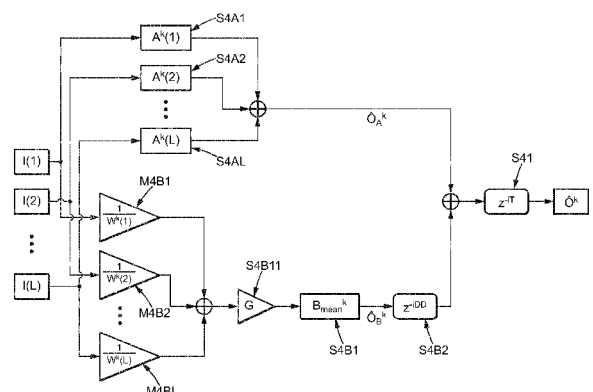
72 Inventeur(s) : **PALLONE GREGORY et EMERIT MARC.**

73 Titulaire(s) : **ORANGE Société anonyme.**

74 Mandataire(s) : **CABINET PLASSERAUD.**

54 **SPATIALISATION SONORE AVEC EFFET DE SALLE.**

57 L'invention concerne un procédé de spatialisation sonore, dans lequel on applique au moins un filtrage, avec sommation, à au moins deux signaux d'entrée $I(1), I(2), \dots, I(L)$, caractérisé en ce que le filtrage comporte:
 - l'application d'au moins une première fonction de transfert à effet de salle ($A^k(1), A^k(2), \dots, A^k(L)$), la première fonction de transfert étant spécifique à chaque signal d'entrée,
 - et l'application d'au moins une deuxième fonction de transfert à effet de salle (B_{mean}^k), la deuxième fonction de transfert étant commune à tous les signaux d'entrée.



Spatialisation sonore avec effet de salle

L'invention est relative au traitement de données sonores, et plus particulièrement à la spatialisation (dite « rendu 3D ») de signaux audio.

5 Une telle opération est par exemple exécutée lors du décodage d'un signal audio 3D codé, représenté sur un certain nombre de canaux, vers un nombre de canaux différents, deux par exemple, pour permettre la restitution des effets 3D audio sur un casque d'écoute.

10 L'invention est également relative à la transmission et à la restitution de signaux audio multicanaux et à leur conversion vers un dispositif de restitution, transducteur, imposé par l'équipement d'un utilisateur. C'est par exemple le cas pour la restitution d'une scène sonore 5.1 par un casque d'écoute audio, ou par une paire de hauts parleurs.

L'invention est également relative à la restitution, dans le cadre d'un jeu ou enregistrement vidéo par exemple, d'un ou plusieurs échantillons sonores stockés dans des fichiers, en vue de leur spatialisation.

15 Dans le cas d'une source monophonique statique, la binauralisation est basée sur le filtrage du signal monophonique par la fonction de transfert entre la position désirée de la source et chacune des deux oreilles. Le signal binaural (deux canaux) obtenu peut alors alimenter un casque audio et fournir la sensation à l'auditeur d'une source à la position simulée. Ainsi, le terme « binaural » vise la restitution d'un signal sonore avec des effets de spatialisation.

20 Chacune des fonctions de transfert simulant différentes positions peuvent être mesurées en chambre sourde, aboutissant ainsi à un ensemble de HRTF (pour « Head Related Transfer Functions » ou « Fonctions de Transferts Relatives à la Tête ») dans lesquelles aucun effet de salle n'est présent.

25 Ces fonctions de transfert peuvent également être mesurées dans une salle « classique », aboutissant ainsi à un ensemble de BRIR (pour « Binaural Room Impulse Response » ou « Réponse Impulsionnelle Binaurale de salle ») dans lesquelles l'effet de salle, ou réverbération, est présent. L'ensemble des BRIR correspondent donc à un ensemble de fonctions de transfert entre une position donnée et les oreilles d'un auditeur (réel ou tête artificielle) placé dans une salle.

30 La technique habituelle de mesure de BRIR consiste à envoyer successivement dans chacun des haut-parleurs réels, positionnés autour d'une tête (réelle ou artificielle) équipée de microphones dans les oreilles, un signal de test (par exemple un signal sweep, une séquence binaire pseudo-aléatoire ou un bruit blanc). Ce signal de test rend possible, lors d'un traitement hors temps réel, la reconstitution (généralement par déconvolution) de la réponse impulsionnelle entre la position du haut-parleur et chacune des deux oreilles.

La différence entre un ensemble de HRTF et de BRIR se situe principalement dans la longueur de la réponse impulsionnelle, de l'ordre de la milliseconde pour les HRTF, à l'ordre de la seconde pour les BRIR.

5 Le filtrage étant basé sur la convolution entre le signal monophonique et la réponse impulsionnelle, la complexité pour effectuer une binauralisation avec des BRIR (contenant un effet de salle) est nettement plus élevée qu'avec des HRTF.

10 Il est possible par cette technique de simuler au casque ou sur un nombre limité de haut-parleurs une écoute d'un contenu multicanal (L canaux) généré par L haut-parleurs dans une salle. En effet, il suffit de considérer chacun des L haut-parleurs comme une source virtuelle positionnée idéalement par rapport à l'auditeur, de mesurer dans la salle à simuler les fonctions de transfert (pour les oreilles gauche et droite) de chacun de ces L haut-parleurs, puis d'appliquer à chacun des L signaux audio (censés alimenter les L haut-parleurs réels) les filtres BRIR correspondant aux haut-parleurs. Les signaux alimentant chacune des oreilles sont sommés pour fournir un signal binaural alimentant un casque audio.

15 On note $I(l)$ (avec $l=[1,L]$) le signal d'entrée censé alimenter les L haut-parleurs. On note $BRIR^{g/d}(l)$, les BRIR de chacun des haut-parleurs pour chacune des deux oreilles, et on note $O^{g/d}$ le signal binaural de sortie. La binauralisation du signal multicanal s'écrit donc :

$$O^g = \sum_{l=1}^L I(l) * BRIR^g(l)$$

$$O^d = \sum_{l=1}^L I(l) * BRIR^d(l)$$

Où * représente l'opérateur de convolution.

20 Par la suite, l'indice l tel que $l \in [1, L]$ fait référence à un des L haut-parleurs. On a bien une BRIR pour un signal l .

Ainsi, en référence à la figure 1, deux convolutions (une pour chaque oreille) sont présentes pour chaque haut-parleur (étapes S11 à S1L).

25 Pour L haut-parleurs, la binauralisation nécessite donc $2.L$ convolutions. On peut calculer la complexité C_{conv} dans le cas d'une implémentation rapide par bloc. Une implémentation rapide par bloc est par exemple donnée par une transformée de Fourier rapide (FFT pour « Fast Fourier Transform »). Le document « Submission and Evaluation Procedures for 3D Audio » (MPEG 3D Audio) précise une formule possible pour le calcul de C_{conv} :

$$C_{conv} = (L + 2). (nBlocs). (6. \log_2(2Fs / nBlocs))$$

Dans cette équation, L représente le nombre de FFT pour transformer en fréquence les signaux d'entrée (1 FFT par signal d'entrée), le 2 représente le nombre de FFT inverse pour obtenir le signal binaural temporel (2 FFT inverse pour les deux canaux du binaural), le 6 indique un coefficient de complexité par FFT, le second 2 indique un bourrage de zéros nécessaire pour éviter les problèmes dus à la convolution circulaire, F_s indique la taille de chacune des BRIR, et n_{Blocs} représente le fait d'utiliser un traitement par bloc, plus réaliste dans une approche où la latence ne doit pas être excessivement élevée, et \cdot représente la multiplication.

Ainsi pour utilisation typique avec $n_{\text{Blocs}}=10$, $F_s=48000$, $L=22$, la complexité par échantillon de signal multicanal pour une convolution directe basée sur une FFT est de $C_{\text{conv}} = 19049$ multiplications-additions.

Cette complexité est trop élevée pour une implémentation réaliste à l'heure actuelle sur des processeurs courants (mobiles par exemple), il est donc nécessaire de diminuer cette complexité sans dégrader fortement le rendu de la binauralisation.

Pour que la spatialisation soit de bonne qualité, la totalité du signal temporel des BRIR doit être appliqué.

La présente invention vient améliorer la situation.

Elle vise à diminuer fortement la complexité de la binauralisation d'un signal multicanal avec effet de salle tout en conservant au mieux la qualité audio.

La présente invention propose à cet effet un procédé de spatialisation sonore, dans lequel on applique au moins un filtrage, avec sommation, à au moins deux signaux d'entrée ($I(1)$, $I(2)$, ..., $I(L)$), caractérisé en ce que le filtrage comporte:

- l'application d'au moins une première fonction de transfert à effet de salle ($A^k(1)$, $A^k(2)$, ..., $A^k(L)$), cette première fonction de transfert étant spécifique à chaque signal d'entrée,
- et l'application d'au moins une deuxième fonction de transfert à effet de salle (B_{mean}^k), cette deuxième fonction de transfert étant commune à tous les signaux d'entrée.

Les signaux d'entrée correspondent par exemple aux différents canaux d'un signal multicanal. Un tel filtrage peut en particulier délivrer au moins deux signaux de sortie destinés à une restitution spatialisée (en binaural ou en transaural, ou encore en restitution ambiophonique faisant intervenir plus de deux signaux de sortie). Dans un mode de réalisation particulier, le filtrage délivre précisément deux signaux de sortie, le premier signal de sortie étant spatialisé pour l'oreille gauche

et le second signal de sortie étant spatialisé pour l'oreille droite. Cela rend notamment possible la conservation du degré de corrélation naturel qui peut exister entre les oreilles gauche et droite à basses fréquences.

5 Les propriétés physiques (par exemple l'énergie ou la corrélation entre les différentes fonctions de transfert) des fonctions de transfert sur certains intervalles temporels rendent possible des simplifications. Sur ces intervalles, on peut ainsi approximer les fonctions de transferts par un filtre moyen.

10 L'application des fonctions de transfert à effet de salle est donc avantageusement compartimentée sur ces intervalles. Au moins une première fonction de transfert spécifique à chaque signal d'entrée peut être appliquée pour les intervalles où il n'est pas possible de faire d'approximations. Au moins une deuxième fonction de transfert approximée à un filtre moyen peut être appliquée pour les intervalles où il est possible de faire des approximations.

15 L'application d'une seule fonction de transfert commune à chacun des signaux d'entrée réduit sensiblement le nombre de calcul à effectuer pour la spatialisation. La complexité de cette spatialisation est donc avantageusement réduite. Cette simplification réduit ainsi avantageusement le temps de traitement tout en sollicitant moins le ou les processeurs utilisés pour ces calculs.

Dans un mode de réalisation particulier, les première et deuxième fonctions de transfert sont respectivement représentatives:

- de propagations sonores directes et de premières réflexions sonores de ces propagations ; et
- 20 - d'un champ sonore diffus présent après ces premières réflexions,

et le procédé au sens de l'invention comporte en outre:

- l'application de premières fonctions de transfert respectivement spécifiques aux signaux d'entrée, et
- l'application d'une deuxième fonction de transfert, identique pour tous les signaux
- 25 d'entrée, et résultant d'une approximation globale d'un effet de champ sonore diffus.

Ainsi, la complexité du traitement est avantageusement réduite par cette approximation. De plus, l'influence d'une telle approximation sur la qualité du traitement est réduite car cette approximation est liée aux effets de champ sonore diffus et non aux propagations sonores directes. Ces effets de champ sonore diffus sont en effet moins sensibles à des approximations. Les

30 premières réflexions sonores sont typiquement une première succession d'échos de l'onde sonore.

Dans un exemple de réalisation concret, on considère que ces premières réflexions sont au nombre de deux, au plus.

Dans un autre mode de réalisation, une étape préalable de construction des première et deuxième fonctions de transfert à partir de réponses impulsionnelles incorporant un effet de salle, comporte, pour la construction d'une première fonction de transfert, les opérations:

- de détermination d'un instant de début de présence d'ondes sonores directes,

- de détermination d'un instant de début de présence du champ sonore diffus après les premières réflexions, et

- de sélection, dans une réponse impulsionnelle, d'une partie de la réponse qui s'étend temporellement entre l'instant de début de présence d'ondes sonores directes jusqu'à l'instant de début de présence de champ diffus, la partie de réponse sélectionnée correspondant à la première fonction de transfert.

Dans un mode de réalisation particulier, l'instant de début de présence de champ diffus est déterminé à partir de critères prédéterminés. Dans un exemple de réalisation possible, la détection d'une décroissance monotone d'une densité spectrale de puissance sonore dans une salle donnée peut typiquement caractériser le début de présence du champ diffus, et, de là, donner l'instant de début de présence du champ diffus.

Dans une variante, l'instant de début de présence peut être déterminé par une estimation en fonction des caractéristiques de la salle, par exemple simplement à partir du volume de la salle comme on le verra plus loin.

En variante, dans un exemple de réalisation plus simple, on peut considérer que si une réponse impulsionnelle s'étend sur N échantillons, alors l'instant de début de présence du champ diffus intervient par exemple après $N/2$ échantillons de la réponse impulsionnelle. Ainsi, l'instant de début de présence est prédéterminé et correspond donc à une valeur fixe. Typiquement, cette valeur peut correspondre par exemple au 2048^{ème} échantillon sur 48000 échantillons d'une réponse impulsionnelle incorporant un effet de salle.

L'instant de début de présence d'ondes sonores directes, précité, peut correspondre par exemple au début du signal temporel d'une réponse impulsionnelle avec effet de salle.

Dans un mode de réalisation complémentaire, la deuxième fonction de transfert est construite à partir d'un ensemble de parties de réponses impulsionnelles commençant temporellement après l'instant de début de présence du champ diffus.

Dans une variante, la deuxième fonction de transfert peut être déterminée à partir des caractéristiques de la salle, ou de filtres standards prédéterminées.

5 Ainsi, les réponses impulsionnelles incorporant un effet de salle sont avantageusement compartimentées en deux parties séparées par un instant de début de présence. Une telle séparation rend possible un traitement adapté à chacune de ces parties. On peut prendre par exemple une sélection des premiers échantillons (les 2048 premiers) d'une réponse impulsionnelle pour l'utiliser en tant que première fonction de transfert dans le filtrage et ignorer alors les échantillons restants (de 2048 à 48000 par exemple) ou les moyenner avec ceux d'autres réponses impulsionnelles.

10 L'intérêt d'une telle réalisation est alors, de façon particulièrement avantageuse, de simplifier les calculs de filtrage spécifiques aux signaux d'entrée, et de rajouter une forme de bruit issu de la diffusion sonore qui peut être calculé à partir des secondes moitiés de réponses impulsionnelles (sous forme de moyenne par exemple comme on le verra plus loin), ou encore simplement à partir d'une réponse impulsionnelle prédéterminée, estimée simplement en fonction de caractéristiques de la salle donnée (son volume, les revêtements des parois de la salle, ou autres), ou d'une salle
15 standard.

Dans une autre variante, la deuxième fonction de transfert est donnée par application d'une formule du type :

$$B_{mean}^k = \frac{1}{L} \sum_{l=1}^L [B_{norm}^k(l)]$$

avec k l'indice relatif à un signal de sortie,

20 $l \in [1; L]$ l'indice relatif à un signal d'entrée,

L le nombre de signaux d'entrée,

$B_{norm}^k(l)$ une fonction de transfert normalisée obtenue à partir d'un ensemble de parties de réponses impulsionnelles commençant temporellement après l'instant de début de présence du champ diffus.

25 Dans un mode de réalisation, les première et deuxième fonctions de transfert sont obtenues à partir d'une pluralité de réponses impulsionnelles binaurales de salle BRIR.

Dans un autre mode de réalisation, ces première et deuxième fonctions de transfert sont obtenues à partir de valeurs expérimentales issues de mesure de propagations et réverbérations dans une salle donnée. Ainsi, le traitement est effectué à partir de données expérimentales. De telles données
30 traduisent très précisément les effets de salle et garantissent donc un grand réalisme du rendu.

Dans un autre mode de réalisation, les première et deuxième fonctions de transfert sont obtenues à partir de filtres de référence, synthétisés par exemple avec un réseau de retards bouclés.

5 Dans un mode de réalisation, on applique une troncature au début des BRIR. Ainsi, les premiers échantillons de BRIR pour lesquels l'application aux signaux d'entrée n'a pas d'influence sont avantageusement supprimés.

Dans un autre mode de réalisation particulier, on applique un délai de compensation de troncature de début de BRIR. Ce délai de compensation permet de compenser le décalage temporel introduit par la troncature.

10 Dans un autre mode de réalisation, on applique une troncature en fin de BRIR. Ainsi, les derniers échantillons de BRIR pour lesquels l'application aux signaux d'entrée n'a pas d'influence sont avantageusement supprimés.

15 Dans un mode de réalisation, le filtrage comporte l'application d'au moins un délai de compensation correspondant à un écart temporel entre l'instant précité de début d'ondes sonores directes et l'instant de début de présence de champ diffus. Ainsi, les retards pouvant être introduits par l'application de fonctions de transfert temporellement décalées sont avantageusement compensés.

20 Dans un autre mode de réalisation, les première et deuxième fonctions de transfert à effet de salle sont appliquées parallèlement aux signaux d'entrée. De plus, au moins un délai de compensation est appliqué aux signaux d'entrée filtrés par les deuxième fonctions de transfert. Ainsi, un traitement simultané de ces deux fonctions de transfert est possible pour chacun des signaux d'entrée. Un tel traitement réduit avantageusement le temps de traitement pour la mise en œuvre de la présente invention.

25 Dans un autre mode de réalisation, on pondère au moins un signal d'entrée par un poids, ce poids étant spécifique à chacun des signaux d'entrée. Ainsi, les écarts énergétiques entre les différents signaux d'entrées peuvent être pris en compte même si le traitement qui leur est appliqué est en partie approximé par un filtre moyen.

Dans un mode de réalisation, chacun des signaux de sortie est donné par application d'une formule du type :

30

$$O^k = \sum_{l=1}^L (I(l) * A^k(l)) + z^{-iDD} \cdot \sum_{l=1}^L \left(\frac{1}{W^k(l)} \cdot I(l) \right) * B_{mean}^k$$

avec k l'indice relatif à un signal de sortie,

O^k un signal de sortie,

$l \in [1; L]$ l'indice relatif un signal d'entrée parmi les signaux d'entrée,

L le nombre de signaux d'entrée,

5 $I(l)$ un signal d'entrée parmi les signaux d'entrée,

$A^k(l)$ une fonction de transfert avec effet de salle parmi les premières fonctions de transfert avec effet de salle,

B_{mean}^k une fonction de transfert avec effet de salle parmi les deuxièmes fonctions de transfert avec effet de salle,

10 $W^k(l)$ un poids de pondération parmi les poids de pondération,

z^{-iDD} correspond à l'application du délai de compensation,

où \cdot est la multiplication, et

où $*$ est l'opérateur de convolution.

15 Dans un autre mode de réalisation, une étape de décorrélation est appliquée aux signaux d'entrée préalablement à l'application des deuxièmes fonctions de transfert. Dans ce mode de réalisation, les signaux de sortie sont donc obtenus par application d'une formule du type :

$$O^k = \sum_{l=1}^L (I(l) * A^k(l)) + z^{-iDD} \cdot \sum_{l=1}^L \left(\frac{1}{W^k(l)} \cdot I_d(l) \right) * B_{mean}^k$$

avec $I_d(l)$ un signal d'entrée décorrélé parmi lesdits signaux d'entrée, les autres valeurs étant celles définies ci-avant. Ainsi, les écarts d'énergies dus aux différences énergétiques entre les additions de signaux corrélés et les additions de signaux décorrélés peuvent être pris en compte.

20 Dans un mode de réalisation particulier, la décorrélation est appliquée préalablement au filtrage. Ainsi, on peut s'affranchir d'étapes de compensation d'énergie pendant le filtrage.

Dans un autre mode de réalisation, on applique à au moins un signal d'entrée, au moins un gain de compensation d'énergie. Ainsi, l'amplitude de sortie est avantageusement normalisée.

25 Dans un mode de réalisation particulier, le gain de compensation d'énergie est fonction de la corrélation entre les signaux d'entrée. Ainsi, la corrélation entre signaux est avantageusement prise en compte.

Dans un mode de réalisation, les signaux de sortie sont obtenus par application d'une formule du type :

$$O^k = \sum_{l=1}^L (I(l) * A^k(l)) + z^{-iDD} \cdot G(I(l)) \cdot \sum_{l=1}^L \left(\frac{1}{W^k(l)} \cdot I(l) \right) * B_{mean}^k$$

avec $G(I(l))$ le gain de compensation d'énergie déterminé, les autres valeurs étant celles définies ci-avant. Dans une variante, G ne dépend pas de $I(l)$.

- 5 Dans un mode de réalisation, le poids pour la pondération est donné par application d'une formule du type :

$$W^k(l) = \frac{\sqrt{E_{B_{mean}^k}}}{\sqrt{E_{B^k(l)}}}$$

avec k l'indice relatif à un signal de sortie,

$l \in [1; L]$ l'indice relatif un signal d'entrée parmi les signaux d'entrée,

L le nombre de signaux d'entrée,

- 10 avec $E_{B_{mean}^k}$ l'énergie d'une fonction de transfert avec effet de salle parmi les deuxièmes fonctions de transfert avec effet de salle,

$E_{B^k(l)}$ une énergie relative au gain en normalisation.

L'invention vise aussi un programme informatique comportant des instructions pour la mise en œuvre du procédé décrit ci-avant.

- 15 L'invention peut être mise en œuvre par un dispositif de spatialisation sonore, comportant au moins un filtre avec sommation appliqué à au moins deux signaux d'entrée ($I(1)$, $I(2)$, ..., $I(L)$), caractérisé en ce que le filtre utilise:

- au moins une première fonction de transfert à effet de salle ($A^k(1)$, $A^k(2)$, ..., $A^k(L)$), cette première fonction de transfert étant spécifique à chaque signal d'entrée,
- 20 - et au moins une deuxième fonction de transfert à effet de salle (B_{mean}^k), cette deuxième fonction de transfert étant commune à tous les signaux d'entrée.

Un tel dispositif peut prendre la forme matérielle par exemple d'un processeur et éventuellement d'une mémoire de travail, typiquement dans un terminal de communication.

L'invention peut également être mise en œuvre dans un module de décodage de signaux sonores, en tant que signaux d'entrée, comportant le dispositif de spatialisation décrit ci-avant.

D'autres avantages et caractéristiques de l'invention apparaîtront à la lecture de la description détaillée ci-après d'exemples de réalisation de l'invention et à l'examen des dessins sur lesquels :

- 5 - la figure 1 illustre un procédé de spatialisation de l'art antérieur,
- la figure 2 illustre schématiquement les étapes d'un procédé au sens de l'invention, dans un exemple de réalisation,
- la figure 3 représente une réponse impulsionnelle binaurale de salle BRIR,
- la figure 4 illustre schématiquement les étapes d'un procédé au sens de l'invention, dans un
- 10 exemple de réalisation,
- la figure 5 illustre schématiquement les étapes d'un procédé au sens de l'invention, dans un exemple de réalisation,
- la figure 6 représente schématiquement un dispositif comportant des moyens de mise en œuvre du procédé au sens de l'invention.

15

On se réfère à la figure 6 pour illustrer tout d'abord un contexte possible de mise en œuvre de la présente invention dans un dispositif de type terminal connecté TER (par exemple un téléphone, smartphone ou autre, ou encore une tablette connectée, un ordinateur connecté, ou autres). Un tel

20 dispositif TER comporte des moyens de réception (une antenne typiquement) de signaux audio Xc codés en compression, un dispositif de décodage DECOD délivrant des signaux décodés X prêts à être traités par un dispositif de spatialisation avant la restitution des signaux audio (par exemple par

25 en binaural sur un casque à oreillettes CAS). Bien entendu, dans certains cas, il peut être avantageux de garder les signaux partiellement décodés (par exemple dans le domaine des sous-bandes) si le traitement de spatialisation est effectué dans le même domaine (traitement fréquentiel dans le domaine des sous-bandes par exemple).

En référence encore à la figure 6, le dispositif de spatialisation se présente par une combinaison d'éléments :

- hardware comportant typiquement un ou plusieurs circuits CIR coopérant avec une mémoire de travail MEM et un processeur PROC,

- et software, dont les figures 2 et 4 sont des exemples d'organigrammes en illustrant l'algorithme général.

Ici, la coopération entre les éléments hardware et software produit un effet technique procurant notamment une économie de complexité de la spatialisation pour sensiblement un même rendu audio (même sensation pour un auditeur), comme on le verra plus loin.

On se réfère maintenant à la figure 2 pour décrire un traitement au sens de l'invention, ainsi mis en œuvre par des moyens informatiques.

Dans une première étape S21, une préparation des données est effectuée. Cette préparation est optionnelle, les signaux peuvent être traités selon les étapes S22 et suivantes sans ce pré-traitement.

En particulier, cette préparation consiste à effectuer une troncature de chaque BRIR pour ignorer les échantillons inaudibles au début et à la fin de la réponse impulsionnelle.

Cette préparation, pour la troncature en début de réponse impulsionnelle TRONC S, à l'étape S211, consiste à déterminer un instant de début d'ondes sonores directes et peut-être mis en œuvre par les étapes suivantes :

- Une somme cumulée des énergies de chacun des filtres BRIR(l) est calculée. Typiquement, cette énergie est calculée par une somme au carré des amplitudes des échantillons 1 à j, avec j compris dans [1 ; J] avec J le nombre d'échantillon d'un filtre BRIR.

- La valeur d'énergie du filtre d'énergie maximum valMax (parmi les filtres relatifs à l'oreille gauche et à l'oreille droite) est calculée.

- Pour chacun des haut-parleurs l, on calcule l'indice pour lequel l'énergie de chacun des filtres BRIR(l) dépasse un certain seuil en dB calculé par rapport à valMax (e.g. valMax-50dB).

- L'indice de troncature iT retenu pour toutes les BRIR est l'indice minimum parmi tous les indices des BRIR et il est considéré comme l'instant de début d'ondes sonores directes.

L'indice iT obtenu correspond donc au nombre d'échantillons à ignorer pour chacune des BRIR. Une troncature brutale en début de réponse impulsionnelle avec une fenêtre rectangulaire peut mener à des artefacts audibles si elle est appliquée dans une partie trop énergétique. Il peut donc être préférable d'appliquer une fenêtre de fondu d'entrée adaptée, cependant si des précautions ont été prises dans le seuil choisi, ce fenêtrage devient inutile, car inaudible (on coupe juste du signal inaudible).

Le synchronisme entre BRIR rend possible l'application d'un délai constant pour toutes les BRIR dans un souci de simplicité d'implémentation, même si une optimisation de complexité est possible.

5 La troncature de chaque BRIR pour ignorer les échantillons inaudibles à la fin de la réponse impulsionnelle TRONC E, à l'étape S212, peut être réalisée à partir d'étape similaires à celles décrites ci-avant, adaptées pour convenir à la fin de la réponse impulsionnelle. Une troncature brutale en fin de réponse impulsionnelle avec une fenêtre rectangulaire peut mener à des artefacts audibles sur des signaux impulsionnels où la queue de réverbération peut se révéler audible. Ainsi, dans un mode de réalisation, on applique une fenêtre de fondu de sortie adaptée.

10

A l'étape 22, une isolation en synchronisme ISOL A/B est effectuée. Cette isolation en synchronisme consiste à séparer, pour chaque BRIR, la partie « son direct » et « premières réflexions » (ou Direct, noté A) et de la partie « son diffus » (ou Diffus, noté B). En effet, le traitement à effectuer sur la partie « son diffus » peut avantageusement être différent de celui à effectuer sur la partie « son direct » dans la mesure où il est préférable d'avoir une meilleure qualité de traitement sur la partie « son direct » que sur la partie « son diffus ». Ceci rend possible une optimisation du rapport qualité/complexité.

20 En particulier, pour réaliser l'isolation en synchronisme, on détermine un index « iDD » d'échantillon unique et commun à toutes les BRIR (d'où le terme « synchronisme ») à partir duquel on considère que le reste de la réponse impulsionnelle correspond à un champ diffus. On partitionne donc les réponses impulsionnelles BRIR(l) en deux : A(l) et B(l), où la concaténation des deux correspond à BRIR(l).

25 La figure 3 montre l'index de partitionnement iDD à l'échantillon 2000. La partie à gauche de cet indice iDD correspond à la partie A. La partie à droite de cet indice iDD correspond à la partie B. Dans un mode de réalisation, ces deux parties sont isolées, sans fenêtrage, afin de subir des traitements différents. Dans une variante, un fenêtrage entre les parties A(l) et B(l) est appliqué.

L'indice iDD peut-être spécifique à la salle pour laquelle les BRIR ont été déterminés. Le calcul de cet indice peut donc dépendre de l'enveloppe spectrale, de la corrélation des BRIR ou encore de l'échogramme de ces BRIR. Par exemple, iDD peut être déterminé par une formule du type

30 $iDD = \sqrt{V_{salle}}$ avec V_{salle} le volume de la salle de mesure.

Dans un mode de réalisation, iDD est une valeur fixe, typiquement 2000. Dans une variante, iDD varie, avantageusement de manière dynamique, en fonction de l'environnement à partir duquel les signaux d'entrée sont capturés.

Le signal de sortie pour les oreilles gauche (g) et droite (d), représenté par $O^{g/d}$, s'écrit donc :

$$\begin{aligned} O^{g/d} &= \sum_{l=1}^L I(l) * BRIR^{g/d}(l) = O_A^{g/d} + z^{-iDD} \cdot O_B^{g/d} \\ &= \sum_{l=1}^L I(l) * A^{g/d}(l) + z^{-iDD} \cdot \sum_{l=1}^L I(l) * B^{g/d}(l) \end{aligned}$$

où z^{-iDD} correspond au délai de iDD échantillons.

L'application de ce délai aux signaux est effectuée en stockant les valeurs calculées pour $\sum_{l=1}^L I(l) * B^{g/d}(l)$ dans une mémoire temporaire (par exemple dans un buffer) et en les restituant

5

Dans un mode de réalisation, les index d'échantillons choisis pour A et B peuvent également tenir compte des longueurs de trames dans le cas d'intégration dans un codeur audio. En effet, des tailles typiques de trames de 1024 échantillons peut mener à un choix tel que A fasse 1024 et B fasse 2048, en s'assurant que B est bien une zone de champ diffus pour toutes les BRIR.

10 En particulier, il peut être intéressant que la taille de B soit un multiple de la taille de A car si le filtrage est implémenté par blocs de FFT, alors le calcul d'une FFT pour A peut être réutilisé pour B.

15 Un champ diffus se caractérise par le fait qu'il est statistiquement identique en tous les points de la salle. Ainsi, sa réponse en fréquence varie peu en fonction du haut-parleur à simuler. La présente invention exploite cette caractéristique dans le but de remplacer tous les filtres Diffus $D(l)$ de toutes les BRIR par un seul et unique filtre « moyen » B_{mean} afin de diminuer fortement la complexité due aux convolutions multiples. Pour cela, on peut modifier la partie champ diffus B à l'étape S23B, toujours en référence à la figure 2.

20 A l'étape S23B1, on calcule la valeur du filtre moyen B_{mean} . D'abord, il est extrêmement rare que le système complet soit calibré idéalement, on peut donc appliquer un gain de pondération qui sera reporté dans le signal d'entrée afin d'effectuer une unique convolution par oreille pour la partie champs diffus. On décompose donc les BRIR en des filtres normalisés en énergie, et on reporte le gain de normalisation $\sqrt{E_{B^{g/d}(l)}}$ dans le signal d'entrée :

$$\begin{aligned} O_B^{g/d} &= \sum_{l=1}^L [I(l) * B^{g/d}(l)] = \sum_{l=1}^L \left[I(l) * \left(\sqrt{E_{B^{g/d}(l)}} \cdot B_{\text{norm}}^{g/d}(l) \right) \right] \\ &= \sum_{l=1}^L \left[\left(\sqrt{E_{B^{g/d}(l)}} \cdot I(l) \right) * B_{\text{norm}}^{g/d}(l) \right] \end{aligned}$$

avec $B_{norm}^{g/d}(l) = \frac{B^{g/d}(l)}{\sqrt{E_{B^{g/d}(l)}}}$ où $E_{B^{g/d}(l)}$ représente l'énergie de $B^{g/d}(l)$.

Ensuite, on approxime $B_{norm}^{g/d}(l)$ par un seul et unique filtre moyen $B_{mean}^{g/d}$ qui n'est plus fonction du haut-parleur l , mais qu'il est possible de normaliser également en énergie :

$$O_B^{g/d} \approx \hat{O}_B^{g/d} = \sum_{l=1}^L \left[\left(\sqrt{E_{B^{g/d}(l)}} \cdot I(l) \right) * \left(\frac{B_{mean}^{g/d}}{\sqrt{E_{B_{mean}^{g/d}}}} \right) \right]$$

5 avec $B_{mean}^{g/d} = \frac{1}{L} \sum_{l=1}^L [B_{norm}^{g/d}(l)]$

Dans un mode de réalisation, ce filtre moyen peut être obtenu par moyennage des échantillons temporels. Dans une variante, il peut être obtenu par tout autre type de moyennage comme par exemple un moyennage des densités spectrales de puissance.

10 Dans un mode de réalisation, l'énergie du filtre moyen $E_{B_{mean}^{g/d}}$ peut être mesurée directement à partir du filtre construit $B_{mean}^{g/d}$. Dans une variante, il peut également être estimé en prenant en compte l'hypothèse que les filtres $B_{norm}^{g/d}(l)$ sont décorrélés. En effet, dans ce cas, comme on somme des signaux d'énergie unitaire, on a :

$$E_{B_{mean}^{g/d}} = \sum \left(\frac{1}{L} \sum_{l=1}^L [B_{norm}^{g/d}(l)] \right)^2 = \frac{1}{L^2} \cdot (L \cdot E_{B_{norm}^{g/d}}) = \frac{1}{L}$$

L'énergie peut être calculée sur l'ensemble des échantillons correspondants à la partie champ diffus.

15 A l'étape S23B2, on calcule la valeur du facteur de pondération $W^{g/d}(l)$. Un seul facteur de pondération à appliquer au signal d'entrée est calculé, prenant en compte les normalisations des filtres Diffus et du filtre moyen:

$$\hat{O}_B^{g/d} = \sum_{l=1}^L \left[\left(\frac{\sqrt{E_{B^{g/d}(l)}}}{\sqrt{E_{B_{mean}^{g/d}}}} \cdot I(l) \right) * B_{mean}^{g/d} \right] = \sum_{l=1}^L \left[\left(\frac{1}{W^{g/d}(l)} \cdot I(l) \right) * B_{mean}^{g/d} \right]$$

$$\text{Avec } W^{g/d}(l) = \frac{\sqrt{E_{B_{mean}^{g/d}}}}{\sqrt{E_{B^{g/d}(l)}}}$$

Le filtre moyen étant constant, il peut sortir de la somme :

$$\hat{O}_B^{g/d} = \sum_{l=1}^L \left[\left(\frac{1}{W^{g/d}(l)} \cdot I(l) \right) \right] * B_{mean}^{g/d}$$

Ainsi, les L convolutions avec la partie champ diffus sont remplacées par une seule convolution avec un filtre moyen, moyennant une somme pondérée du signal d'entrée.

- 5 A l'étape S23B3, on peut optionnellement calculer un gain G corrigeant le gain du filtre moyen $B_{mean}^{g/d}$. En effet, dans le cas de la convolution entre les signaux d'entrée et les filtres non approximatés, quelles que soient les valeurs de corrélation entre les signaux d'entrée, le filtrage par des filtres décorrelés que sont les $B^{g/d}(l)$ mène à des signaux à sommer qui sont alors eux aussi décorrelés. A l'inverse, dans le cas de la convolution entre les signaux d'entrée et le filtre moyen
- 10 approximaté, l'énergie du signal issu de la sommation des signaux filtrés va dépendre de la valeur de corrélation existant entre les signaux d'entrée.

Par exemple,

* si tous les signaux d'entrée $I(l)$ sont identiques et d'énergie unitaire, et que les filtres $B(l)$ sont tous décorrelés (puisque champs diffus) et d'énergie unitaire, on a :

$$E_{O_B^{g/d}} = \text{energie} \left(\sum_{l=1}^L [I(l) * B_{norm}^{g/d}(l)] \right) = L$$

15

* si tous les signaux d'entrée $I(l)$ sont décorrelés et d'énergie unitaire, et que les filtres $B(l)$ sont tous d'énergie unitaire, mais remplacés par des filtres identiques $\frac{B_{mean}^{g/d}}{\sqrt{E_{B_{mean}^{g/d}}}}$, on a :

$$\begin{aligned} E_{\hat{O}_B^{g/d}} &= \text{energie} \left(\sum_{l=1}^L \left[I(l) * \left(\frac{B_{mean}^{g/d}}{\sqrt{E_{B_{mean}^{g/d}}}} \right) \right] \right) \\ &= \text{energie} \left(\frac{1}{\sqrt{E_{B_{mean}^{g/d}}}} \cdot \sum_{l=1}^L [I(l) * B_{mean}^{g/d}] \right) = \left(\frac{1}{\sqrt{E_{B_{mean}^{g/d}}}} \right)^2 \cdot \left(L \cdot \frac{1}{L} \right) = L \end{aligned}$$

Car les énergies des signaux décorrelés s'ajoutent.

Ce cas est équivalent au précédent dans le sens où les signaux issus du filtrage sont tous décorrélés, grâce aux signaux d'entrée dans le premier cas, et grâce aux filtres dans le second cas.

* si tous les signaux d'entrée $I(l)$ sont identiques et d'énergie unitaire, et que les filtres $B(l)$ sont tous d'énergie unitaire, mais remplacés par des filtres identiques $\frac{B_{mean}^{g/d}}{\sqrt{E_{B_{mean}^{g/d}}}}$, on a:

$$\begin{aligned} E_{\hat{o}_B}^{g/d} &= \text{energie} \left(\sum_{l=1}^L \left[I(l) * \left(\frac{B_{mean}^{g/d}}{\sqrt{E_{B_{mean}^{g/d}}}} \right) \right] \right) \\ &= \text{energie} \left(\frac{1}{\sqrt{E_{B_{mean}^{g/d}}}} \cdot \sum_{l=1}^L [I(l) * B_{mean}^{g/d}] \right) = \left(\frac{1}{\sqrt{E_{B_{mean}^{g/d}}}} \right)^2 \cdot \left(L^2 \cdot \frac{1}{L} \right) \\ &= L^2 \end{aligned}$$

Car les énergies des signaux identiques s'ajoutent en quadrature (car leurs amplitudes s'ajoutent).

Ainsi,

- si deux haut-parleurs sont actifs simultanément, alimentés par des signaux décorrélés, alors aucun gain n'est apporté en appliquant les étapes S23B1 et S23B2 par rapport à la méthode classique.
- si deux haut-parleurs sont actifs simultanément, alimentés par des signaux identiques, alors un gain de $10 \cdot \log_{10}(L^2/L) = 10 \cdot \log_{10}(2^2/2) = 3.01dB$ est apporté en appliquant les étapes S23B1 et S23B2 par rapport à la méthode classique.
- si trois haut-parleurs sont actifs simultanément, alimentés par des signaux identiques, alors un gain de $10 \cdot \log_{10}(L^2/L) = 10 \cdot \log_{10}(3^2/3) = 4.77dB$ est apporté en appliquant les étapes S23B1 et S23B2 par rapport à la méthode classique.

Les cas évoqués ci-avant correspondent aux cas extrêmes de signaux identiques ou décorrélés. Ces cas sont toutefois réalistes : une source positionnée au milieu de deux haut-parleurs, virtuels ou réels, fournira un signal identique à ces deux haut-parleurs (par exemple avec une technique de type VBAP, pour « Vector base amplitude panning »). Dans le cas d'un positionnement dans un système 3D, les 3 haut-parleurs peuvent recevoir le même signal au même niveau.

Ainsi, on peut appliquer une compensation afin de respecter l'énergie des signaux binauralisés.

Idéalement, ce gain de compensation G sera déterminé en fonction du signal d'entrée (soit $G(I(l))$) et sera appliqué à la somme des signaux d'entrée pondérés :

$$\hat{O}_B^{g/d} = G \cdot \sum_{l=1}^L \left[\frac{1}{W^{g/d}(l)} \cdot I(l) \right] * B_{mean}^{g/d}$$

Le gain $G(I(l))$ peut être estimé par un calcul de corrélation entre chacun des signaux. Il peut également être estimé en comparant les énergies des signaux avant et après sommations. Dans ce cas, le gain G peut varier dynamiquement au cours du temps, en fonction par exemple des corrélations entre les signaux d'entrée, qui varient eux-mêmes au cours du temps.

Dans un mode de réalisation simplifié, il est possible de fixer un gain constant, par exemple $G = -3dB = 10^{-3/20}$, ce qui évitera d'avoir à effectuer une estimation de corrélation qui peut être coûteuse. Le gain constant G peut alors être appliqué hors ligne aux facteurs de pondération (donnant ainsi $\frac{G}{W^{g/d}(l)}$), ou au filtre $B_{mean}^{g/d}$, ce qui évitera l'application d'un gain supplémentaire au vol.

Une fois les fonctions de transfert A et B isolées et les filtres $B_{mean}^{g/d}$ (optionnellement les poids $W^{g/d}(l)$ et G) calculés, on applique ces fonctions de transfert et ces filtres aux signaux d'entrée.

Dans un premier mode de réalisation, décrit en référence à la figure 4, le traitement du signal multicanal par application des filtres Direct (A) et Diffus (B) pour chacune des oreilles est effectué de la manière suivante :

- On applique (étapes S4A1 à S4AL) au signal multicanal d'entrée un filtrage efficace (par exemple convolution directe basée-FFT) par les filtres Direct (A), comme décrit dans l'état de l'art. On obtient un signal $\hat{O}_A^{g/d}$
- En fonction des relations entre les signaux d'entrée, notamment en fonction de leur corrélation, on peut optionnellement corriger à l'étape S4B11 le gain du filtre moyen $B_{mean}^{g/d}$ par application du gain G aux signaux de sortie après sommation des signaux d'entrée préalablement pondérés (étapes M4B1 à M4BL).
- On applique au signal multicanal B à l'étape S4B1 un filtrage efficace par le filtre Diffus moyen B_{mean} . Cette étape a bien lieu après sommation des signaux d'entrée préalablement pondérés (étapes M4B1 à M4BL). On obtient le signal $\hat{O}_B^{g/d}$.
- On applique au signal $\hat{O}_B^{g/d}$ un délai iDD afin de compenser le retard introduit lors de l'étape d'isolation du signal B à l'étape S4B2.

- Les signaux $\hat{O}_A^{g/d}$ et $\hat{O}_B^{g/d}$ sont sommés.
- Si une troncature éliminant les échantillons inaudibles au début des réponses impulsionnelles a été réalisée, alors on applique à l'étape S41 au signal d'entrée un délai iT correspondant aux échantillons inaudibles supprimés.

5 Dans une variante, en référence à la figure 5, les signaux ne sont pas seulement calculés pour les oreilles gauches et droites (indices g et d ci-avant) mais pour k dispositifs de restitution (typiquement des haut-parleurs).

Dans un second mode de réalisation, on applique le gain G préalablement à la sommation des signaux d'entrée, c'est-à-dire pendant les étapes de pondération (étapes M4B1 à M4BL).

10 Dans un troisième mode de réalisation, on applique une décorrélation aux signaux d'entrée. Ainsi, les signaux sont décorrélés après convolution par le filtre B_{mean} quelles que soient les corrélations d'origine entre signaux d'entrée. On peut utiliser une implémentation efficace de décorrélation (par exemple en utilisant un réseau de retards bouclés) afin d'éviter l'utilisation de filtres de décorrélation coûteux.

15

Ainsi, en supposant de manière réaliste que des BRIR de longueur 48000 échantillons peuvent être :

- tronquées entre l'échantillon 150 et l'échantillon 3222 par la technique décrite à l'étape S21,
- 20 - décomposées en deux parties : champ direct A de 1024 échantillons, et champ diffus B de 2048 échantillons, par la technique décrite à l'étape S22,

alors la complexité de binauralisation peut être approximativement donnée par :

$$C_{\text{inv}} = C_{\text{invA}} + C_{\text{invB}} = (L+2).(6.\log_2(2.NA)) + (L+2).(6.\log_2(2.NB))$$

Avec NA et NB les tailles en échantillons de A et B

25 Ainsi pour nBlocs=10, Fs=48000, L=22, NA=1024 et NB=2048, la complexité par échantillon de signal multicanal pour une convolution basée-FFT est de $C_{\text{conv}} = 3312$ multiplications-additions.

Ce résultat est cependant logiquement à comparer à une solution simple implémentant seulement la troncature, soit pour nBlocs=10, Fs=3072, L=22 :

$$C_{\text{tronc}} = (L+2).(nBlocs).(6.\log_2(2.Fs/nBlocs)) = 13339$$

Il existe donc un facteur $19049/3312=5.75$ de complexité entre l'état de l'art et la présente invention, et encore un facteur $13339/3312=4$ de complexité entre l'état de l'art bénéficiant de la troncature et la présente invention.

5 Si la taille de B est un multiple de la taille de A, alors si le filtrage est implémenté par blocs de FFT, le calcul d'une FFT pour A peut être réutilisé pour B. On a donc besoin de L FFT sur NA points, qui serviront à la fois au filtrage par A et par B, deux FFT inverse sur NA points pour obtenir le signal binaural temporel, et la multiplication des spectres en fréquence.

Dans ce cas, la complexité peut être approximée (les additions sont négligées, (L+1) correspond à la multiplication des spectres, L pour A et 1 pour B) par :

$$10 \quad C_{inv2} = (L+2).(6.\log_2(2.NA)) + (L+1) = 1607$$

Avec cette approche, on gagne encore un facteur 2, et donc un facteur 12 et 8 par rapport à l'état de l'art non tronquée et tronquée.

L'invention peut trouver une application directe dans la norme MPEG-H 3D Audio.

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

Ainsi par exemple, on a décrit ci-avant un mode de réalisation dans lequel le signal Direct A n'est pas approximé par un filtre moyen. Bien entendu, on peut utiliser filtre moyen de A pour faire les convolutions (étapes S4A1 à S4AL) avec les signaux issus des haut-parleurs.

20 On a décrit ci-avant un mode de réalisation basé sur le traitement d'un contenu multicanal généré pour L haut-parleurs. Bien sûr, le contenu multicanal peut-être généré par tout type de source audio comme par exemple la voix, un instrument de musique, un bruit quelconque, etc.

25 On a décrit ci-avant un mode de réalisation basé sur des formules s'appliquant dans un certain domaine de calcul (par exemple domaine transformé). Bien sûr, la présente invention ne se limite pas à ces formules et ces formules peuvent être modifiées pour être applicables dans d'autres domaines de calcul (par exemple domaine temporel, domaine fréquentiel, domaine temps-fréquence, etc.).

30 On a décrit ci-avant un mode de réalisation basé sur des valeurs de BRIR déterminées dans une salle. Bien sûr, on peut mettre en œuvre la présente invention pour tout type d'environnement extérieur (par exemple salle de concert, plein air, etc.).

On a décrit ci-avant un mode de réalisation basé sur l'application de deux fonctions de transfert. Bien sûr, on peut mettre en œuvre la présente invention avec plus de deux fonctions de transfert. Par exemple, on peut isoler en synchronisme une partie relative aux sons émis directement, une partie relative aux premières réflexions et une partie relative aux sons diffus.

REVENDEICATIONS

1. Procédé de spatialisation sonore, dans lequel on applique au moins un filtrage, avec sommation, à au moins deux signaux d'entrée ($I(1)$, $I(2)$, ..., $I(L)$), caractérisé en ce que le filtrage comporte:

- 5
- l'application d'au moins une première fonction de transfert à effet de salle ($A^k(1)$, $A^k(2)$, ..., $A^k(L)$), ladite première fonction de transfert étant spécifique à chaque signal d'entrée,
 - et l'application d'au moins une deuxième fonction de transfert à effet de salle (B_{mean}^k), ladite deuxième fonction de transfert étant commune à tous les signaux d'entrée.

10 2. Procédé selon la revendication 1, caractérisé en ce que lesdites première et deuxième fonctions de transfert sont respectivement représentatives:

- de propagations sonores directes et de premières réflexions sonores desdites propagations ; et
- d'un champ sonore diffus présent après lesdites premières réflexions,

et en ce qu'il comporte :

- 15
- l'application de premières fonctions de transfert respectivement spécifiques aux signaux d'entrée, et
 - l'application d'une deuxième fonction de transfert, identique pour tous les signaux d'entrée, et résultant d'une approximation globale d'un effet de champ sonore diffus.

20 3. Procédé selon la revendication 2, caractérisé en ce qu'il comporte une étape préalable de construction desdites première et deuxième fonctions de transfert à partir de réponses impulsionnelles incorporant un effet de salle, ladite étape préalable comportant, pour la construction d'une première fonction de transfert, les opérations:

- de détermination d'un instant de début de présence d'ondes sonores directes,

25

- de détermination d'un instant de début de présence dudit champ sonore diffus après les premières réflexions, et

- de sélection, dans une réponse impulsionnelle, d'une partie de la réponse qui s'étend temporellement entre ledit instant de début de présence d'ondes sonores directes jusqu'audit instant

de début de présence de champ diffus, ladite partie de réponse sélectionnée correspondant à ladite première fonction de transfert.

4. Procédé selon revendication 3, caractérisé en ce que la deuxième fonction de transfert est construite à partir d'un ensemble de parties de réponses impulsionnelles commençant temporellement après ledit instant de début de présence du champ diffus.

5. Procédé selon l'une des revendications 3 ou 4, dans lequel ladite deuxième fonction de transfert est donnée par application d'une formule du type :

$$B_{\text{mean}}^k = \frac{1}{L} \sum_{l=1}^L [B_{\text{norm}}^k(l)]$$

avec k l'indice relatif à un signal de sortie,

$l \in [1; L]$ l'indice relatif à un signal d'entrée,

L le nombre de signaux d'entrée,

$B_{\text{norm}}^k(l)$ une fonction de transfert normalisée obtenue à partir d'un ensemble de parties de réponses impulsionnelles commençant temporellement après ledit instant de début de présence du champ diffus.

6. Procédé selon l'une des revendications 3 à 5, caractérisé en ce que ledit filtrage comporte l'application d'au moins un délai de compensation correspondant à un écart temporel entre ledit instant de début d'ondes sonores directes et ledit instant de début de présence de champ diffus.

7. Procédé selon la revendication 6, caractérisé en ce que lesdites première et deuxième fonctions de transfert à effet de salle sont appliquées parallèlement auxdits signaux d'entrée et en ce que ledit au moins un délai de compensation est appliqué aux signaux d'entrée filtrés par lesdites deuxièmes fonctions de transfert.

8. Procédé selon l'une des revendications précédentes, caractérisé en ce que l'on pondère au moins un signal d'entrée par un poids, ledit poids étant spécifique à chacun des signaux d'entrée.

9. Procédé selon la revendication 8, caractérisé en ce que chacun desdits signaux de sortie est donné par application d'une formule du type :

$$O^k = \sum_{l=1}^L (I(l) * A^k(l)) + z^{-iDD} \cdot \sum_{l=1}^L \left(\frac{1}{W^k(l)} \cdot I(l) \right) * B_{mean}^k$$

avec k l'indice relatif à un signal de sortie,

O^k un signal de sortie,

$l \in [1; L]$ l'indice relatif un signal d'entrée parmi lesdits signaux d'entrée,

L le nombre de signaux d'entrée,

10 $I(l)$ un signal d'entrée parmi lesdits signaux d'entrée,

$A^k(l)$ une fonction de transfert avec effet de salle parmi lesdites premières fonctions de transfert avec effet de salle,

B_{mean}^k une fonction de transfert avec effet de salle parmi lesdites deuxièmes fonctions de transfert avec effet de salle,

15 $W^k(l)$ un poids de pondération parmi lesdits poids de pondération,

z^{-iDD} correspond à l'application dudit délai de compensation,

où \cdot est la multiplication, et

où $*$ est l'opérateur de convolution.

20 10. Procédé selon la revendication 8, caractérisé en ce qu'il comprend une étape de décorrélation des signaux d'entrée, préalablement à l'application des deuxièmes fonctions de transfert et en ce que chacun desdits signaux de sortie est donné par application d'une formule du type :

$$O^k = \sum_{l=1}^L (I(l) * A^k(l)) + z^{-iDD} \cdot \sum_{l=1}^L \left(\frac{1}{W^k(l)} \cdot I_d(l) \right) * B_{mean}^k$$

avec k l'indice relatif à un signal de sortie,

O^k un signal de sortie,

$l \in [1; L]$ l'indice relatif un signal d'entrée parmi lesdits signaux d'entrée,

L le nombre de signaux d'entrée,

5 $I(l)$ un signal d'entrée parmi lesdits signaux d'entrée,

$I_d(l)$ un signal d'entrée décorrélé parmi lesdits signaux d'entrée,

$A^k(l)$ une fonction de transfert avec effet de salle parmi lesdites premières fonctions de transfert avec effet de salle,

10 B_{mean}^k une fonction de transfert avec effet de salle parmi lesdites deuxièmes fonctions de transfert avec effet de salle,

$W^k(l)$ un poids de pondération parmi lesdits poids de pondération,

z^{-iDD} correspond à l'application dudit délai de compensation,

où \cdot est la multiplication, et

où $*$ est l'opérateur de convolution.

15

11. Procédé selon la revendication 8, caractérisé en ce qu'il comprend une étape de détermination d'un gain de compensation d'énergie en fonction des signaux d'entrée et en ce que chacun desdits signaux de sortie est donné par application d'une formule du type :

$$O^k = \sum_{l=1}^L (I(l) * A^k(l)) + z^{-iDD} \cdot G(I(l)) \cdot \sum_{l=1}^L \left(\frac{1}{W^k(l)} \cdot I(l) \right) * B_{mean}^k$$

avec k l'indice relatif à un signal de sortie,

20 O^k un signal de sortie,

$l \in [1; L]$ l'indice relatif un signal d'entrée parmi lesdits signaux d'entrée,

L le nombre de signaux d'entrée,

$I(l)$ un signal d'entrée parmi lesdits signaux d'entrée,

$G(I(l))$ ledit gain de compensation d'énergie déterminé,

$A^k(l)$ une fonction de transfert avec effet de salle parmi lesdites premières fonctions de transfert avec effet de salle,

5 E_{mean}^k une fonction de transfert avec effet de salle parmi lesdites deuxièmes fonctions de transfert avec effet de salle,

$W^k(l)$ un poids de pondération parmi lesdits poids de pondération,

z^{-iDD} correspond à l'application dudit délai de compensation,

où \cdot est la multiplication, et

10 où $*$ est l'opérateur de convolution.

12. Procédé selon l'une des revendications 8 à 11, caractérisé en ce que ledit poids est donné par application d'une formule du type :

$$W^k(l) = \frac{\sqrt{E_{mean}^k}}{\sqrt{E_{B^k(l)}}}$$

15 avec k l'indice relatif à un signal de sortie,

$l \in [1; L]$ l'indice relatif un signal d'entrée parmi lesdits signaux d'entrée,

L le nombre de signaux d'entrée,

avec E_{mean}^k l'énergie d'une fonction de transfert avec effet de salle parmi lesdites deuxièmes fonctions de transfert avec effet de salle,

20 $E_{B^k(l)}$ une énergie relative au gain en normalisation.

13. Programme informatique comportant des instructions pour la mise en œuvre du procédé selon l'une des revendications 1 à 12, lorsque ces instructions sont exécutées par un processeur.

14. Dispositif de spatialisation sonore, comportant au moins un filtre avec sommation appliqué à au moins deux signaux d'entrée ($I(1), I(2), \dots, I(L)$), caractérisé en ce que le filtre utilise:

- au moins une première fonction de transfert à effet de salle ($A^k(1), A^k(2), \dots, A^k(L)$), ladite première fonction de transfert étant spécifique à chaque signal d'entrée,
- 5 - et au moins une deuxième fonction de transfert à effet de salle (B_{mean}^k), ladite deuxième fonction de transfert étant commune à tous les signaux d'entrée.

15. Module de décodage de signaux sonores, comportant un dispositif de spatialisation selon la revendication 14, desdits signaux sonores en tant que signaux d'entrée.

1/5

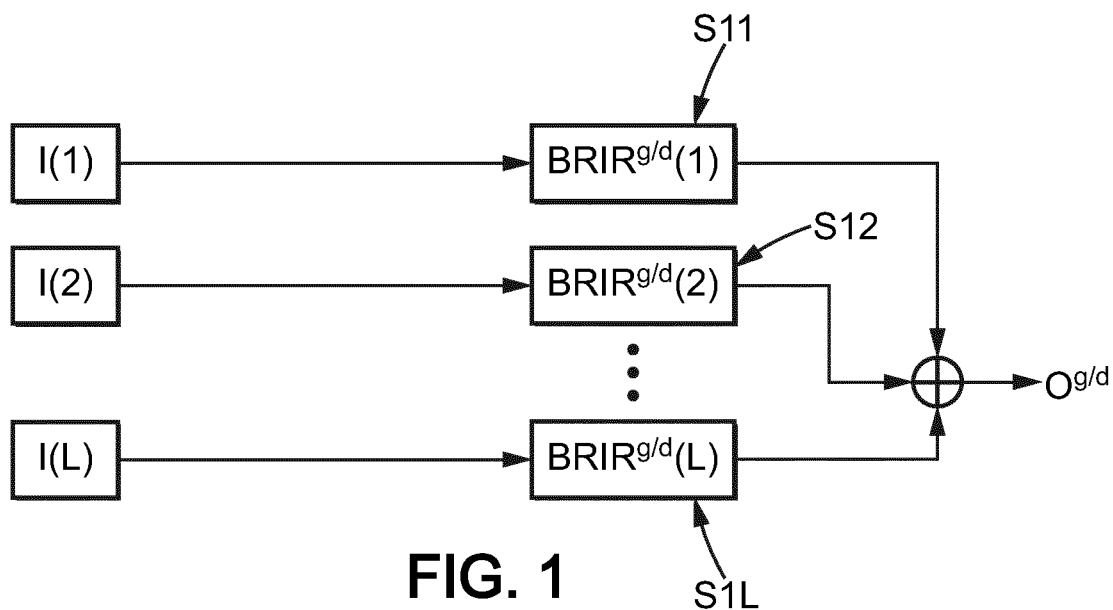


FIG. 1

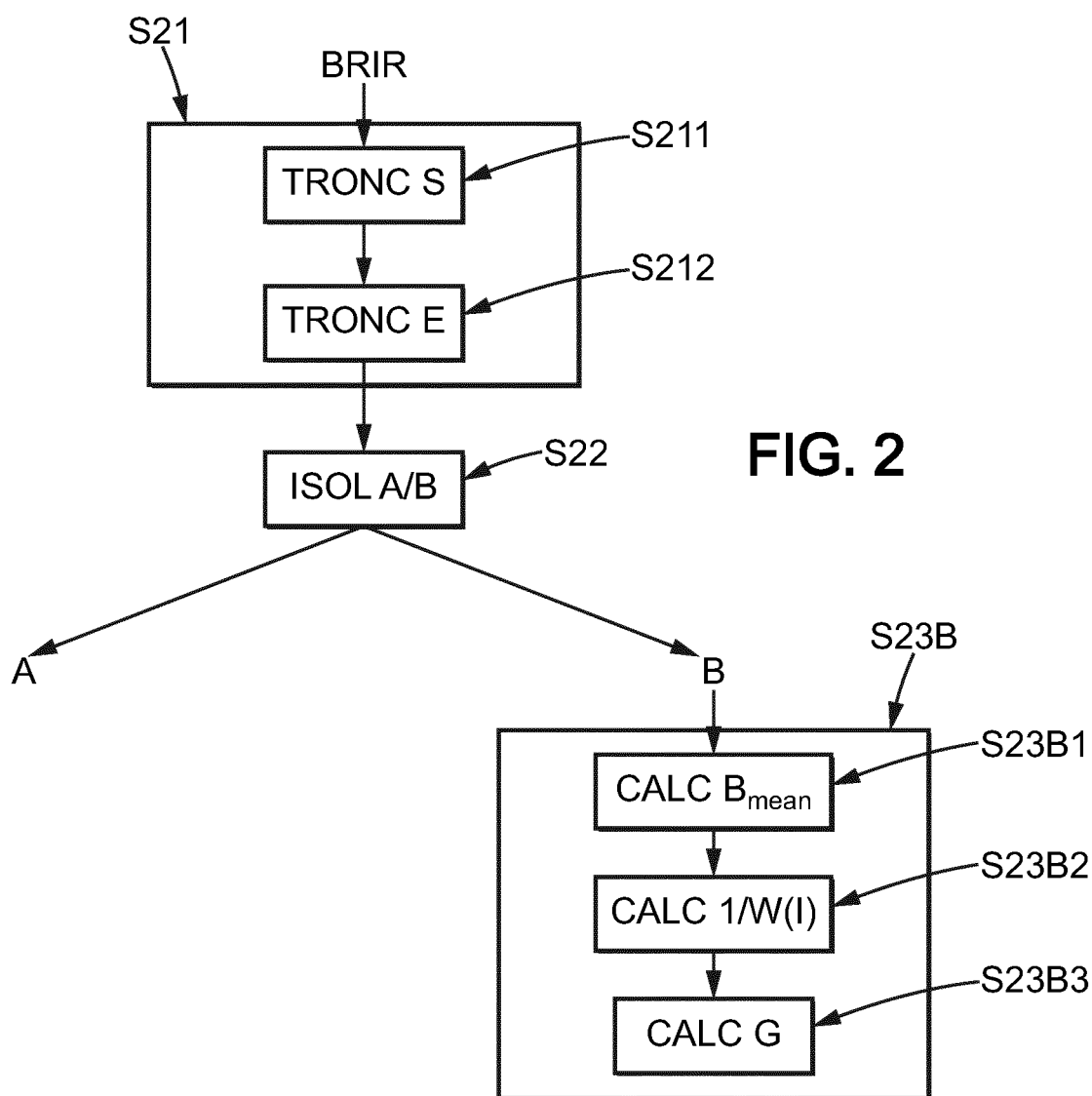
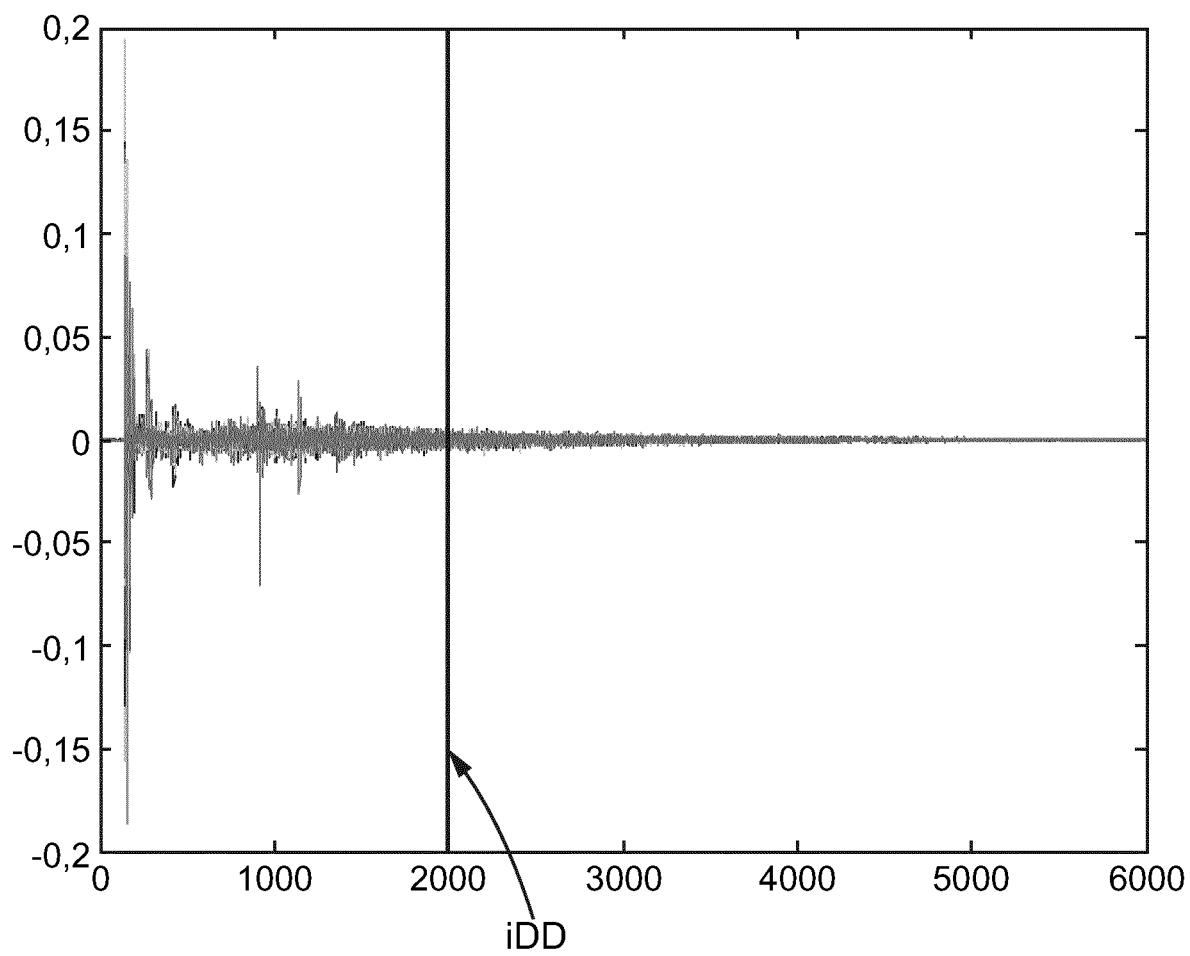
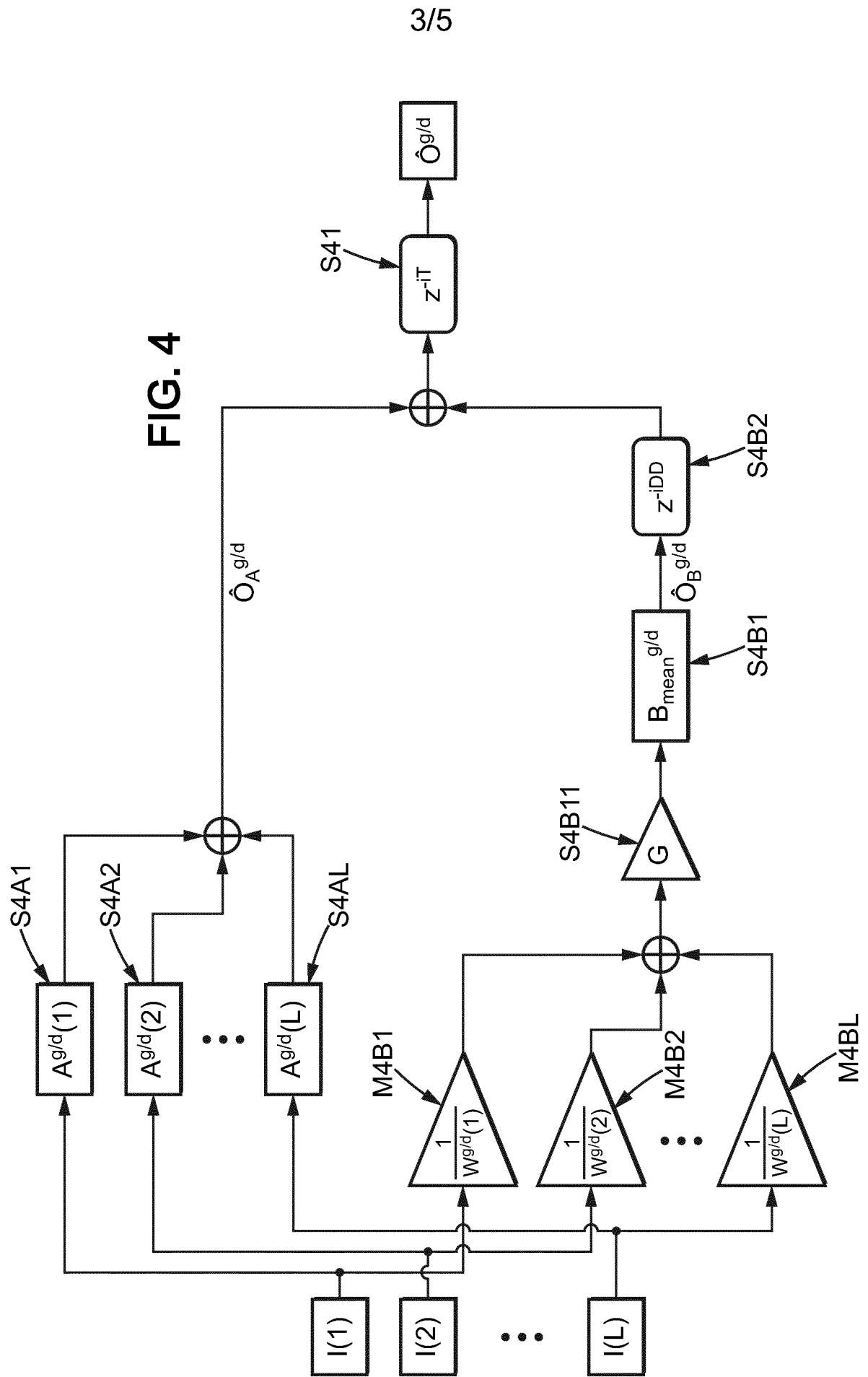
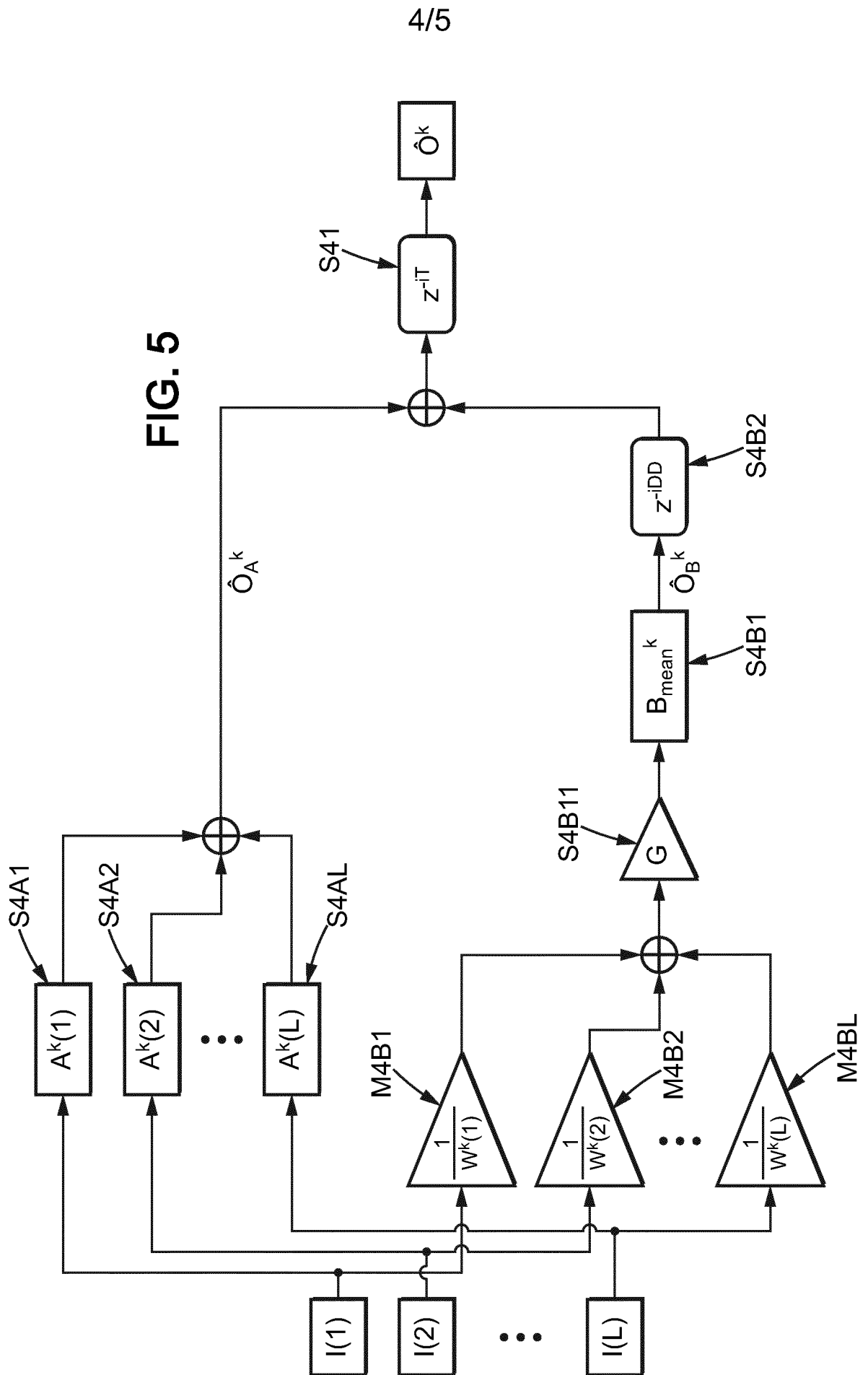


FIG. 2

2/5

**FIG. 3**





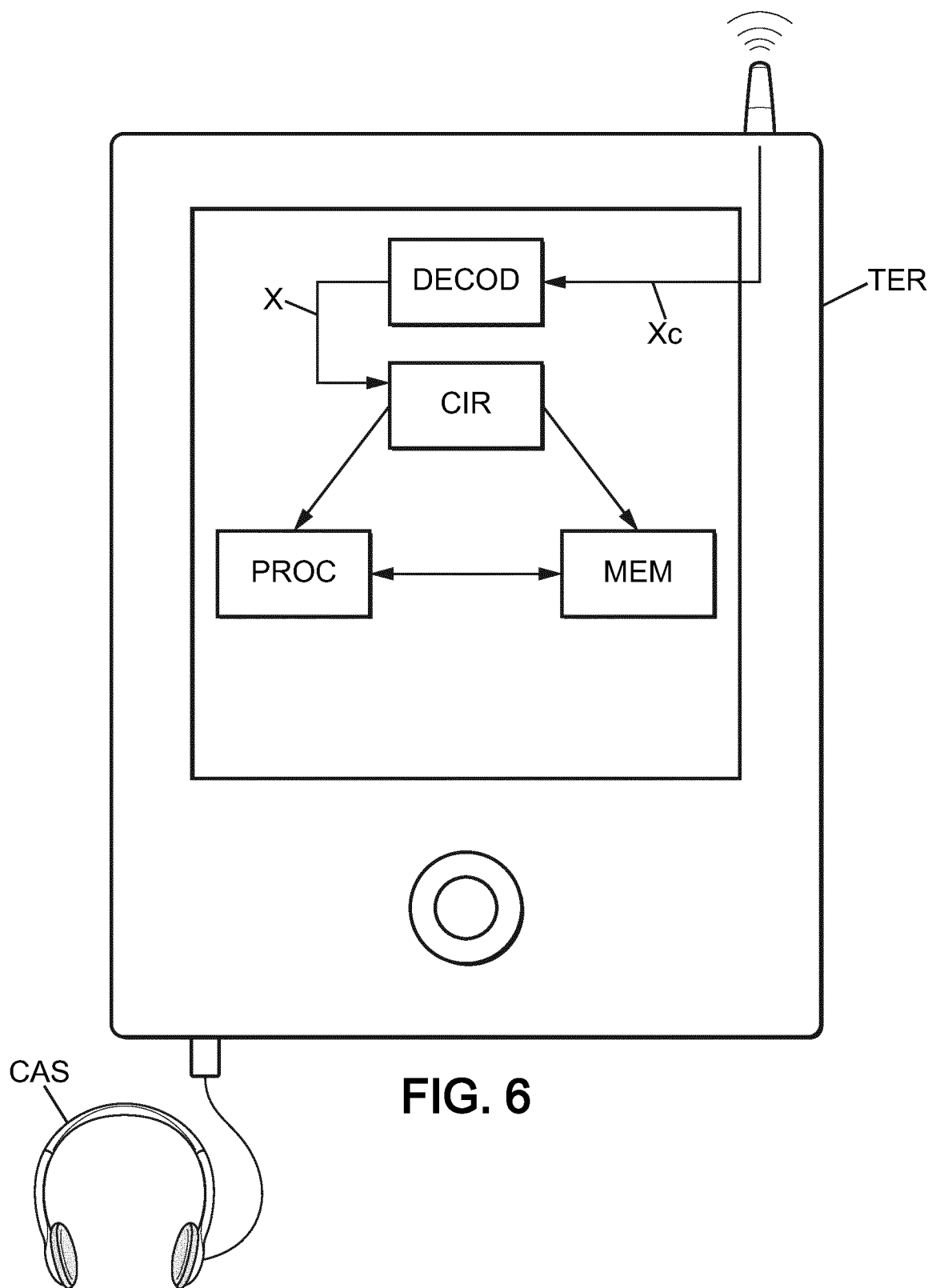


FIG. 6



**RAPPORT DE RECHERCHE
PRÉLIMINAIRE**

N° d'enregistrement
national

établi sur la base des dernières revendications
déposées avant le commencement de la recherche

FA 787170
FR 1357299

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X	US 2006/045294 A1 (SMYTH STEPHEN M [GB] SMYTH STEPHEN MALCOLM [GB]) 2 mars 2006 (2006-03-02)	1-8,10, 11,13-15	H04R5/02 H04S3/00 G10L19/00
Y	* alinéa [0244] - alinéa [0250]; figures 1,28,41,49,50 *	9,12	
Y	* alinéas [0079], [0271] * -----		
Y	STEWART REBECCA ET AL: "Generating a Spatial Average Reverberation Tail Across Multiple Impulse Responses", CONFERENCE: 35TH INTERNATIONAL CONFERENCE: AUDIO FOR GAMES; FEBRUARY 2009, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, 1 février 2009 (2009-02-01), XP040509247, * section 2; page 1, colonne 2 - page 2, colonne 1 * * section 3.1; page 2, colonne 2 - page 3, colonne 1 *	9,12	
A	WO 2007/031906 A2 (KONINKL PHILIPS ELECTRONICS NV [NL]; BREEBAART JEROEN [NL]) 22 mars 2007 (2007-03-22) * page 10, ligne 30 - page 11, ligne 21 *	10	DOMAINES TECHNIQUES RECHERCHÉS (IPC) H04S G10L G10H
A	BREEBAART J ET AL: "Multi-channel goes mobile: MPEG surround binaural rendering", AES INTERNATIONAL CONFERENCE. AUDIO FOR MOBILE AND HANDHELD DEVICES, 2 septembre 2006 (2006-09-02), pages 1-13, XP007902577, * le document en entier *	1-15	
	----- -/--		
Date d'achèvement de la recherche		Examineur	
6 mai 2014		Carrière, Olivier	
CATÉGORIE DES DOCUMENTS CITÉS		T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant	
X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire			

1

EPO FORM 1503 12.99 (P04C14)



**RAPPORT DE RECHERCHE
PRÉLIMINAIRE**

N° d'enregistrement national

établi sur la base des dernières revendications déposées avant le commencement de la recherche

FA 787170
FR 1357299

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X	<p>JEAN-MARC JOT: "Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces", MULTIMEDIA SYSTEMS, vol. 7, no. 1, 1 janvier 1999 (1999-01-01), pages 55-69, XP055116104, ISSN: 0942-4962, DOI: 10.1007/s005300050111 * section 2.1; page 57, colonne 1 - colonne 2; figures 1,3,4 * * section 2.4.1; page 60, colonne 2 * * section 3; page 62, colonne 2 - page 63, colonne 1 *</p>	1,2, 13-15	
A	<p>SAVIOJA L ET AL: "CREATING INTERACTIVE VIRTUAL ACOUSTIC ENVIRONMENTS", JOURNAL OF THE AUDIO ENGINEERING SOCIETY, AUDIO ENGINEERING SOCIETY, NEW YORK, NY, US, vol. 47, no. 9, 1 septembre 1999 (1999-09-01), pages 675-705, XP000927390, ISSN: 1549-4950 * le document en entier *</p>	1-15	DOMAINES TECHNIQUES RECHERCHÉS (IPC)
A	<p>MERIMAA JUHA ET AL: "Spatial Impulse Response Rendering I: Analysis and Synthesis", JAES, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, vol. 53, no. 12, 1 décembre 2005 (2005-12-01), pages 1115-1127, XP040507491, * le document en entier *</p>	1-15	
Date d'achèvement de la recherche		Examineur	
6 mai 2014		Carrière, Olivier	
<p>CATÉGORIE DES DOCUMENTS CITÉS</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant</p>			

1
EPO FORM 1503 12.99 (P04C14)

**ANNEXE AU RAPPORT DE RECHERCHE PRÉLIMINAIRE
RELATIF A LA DEMANDE DE BREVET FRANÇAIS NO. FR 1357299 FA 787170**

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche préliminaire visé ci-dessus.

Les dits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du **06-05-2014**

Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets, ni de l'Administration française

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 2006045294 A1	02-03-2006	CA 2578469 A1	09-03-2006
		CN 101133679 A	27-02-2008
		EP 1787494 A2	23-05-2007
		JP 4990774 B2	01-08-2012
		JP 2008512015 A	17-04-2008
		KR 20070094723 A	21-09-2007
		US 2006045294 A1	02-03-2006
		WO 2006024850 A2	09-03-2006

WO 2007031906 A2	22-03-2007	CN 101263740 A	10-09-2008
		CN 101263742 A	10-09-2008
		CN 102395098 A	28-03-2012
		EP 1927265 A2	04-06-2008
		JP 4938015 B2	23-05-2012
		JP 2009508385 A	26-02-2009
		KR 20080046712 A	27-05-2008
		KR 20130045414 A	03-05-2013
		US 2008304670 A1	11-12-2008
		WO 2007031906 A2	22-03-2007
