



(11) **EP 2 000 002 B1**

(12) **FASCICULE DE BREVET EUROPEEN**

(45) Date de publication et mention
de la délivrance du brevet:
05.08.2009 Bulletin 2009/32

(51) Int Cl.:
H04S 3/02 (2006.01)

(21) Numéro de dépôt: **07731710.5**

(86) Numéro de dépôt international:
PCT/FR2007/050894

(22) Date de dépôt: **08.03.2007**

(87) Numéro de publication internationale:
WO 2007/110519 (04.10.2007 Gazette 2007/40)

(54) **PROCEDE ET DISPOSITIF DE SPATIALISATION SONORE BINAURALE EFFICACE DANS LE
DOMAINE TRANSFORME**

VERFAHREN UND EINRICHTUNG ZUR EFFIZIENTEN BINAURALEN RAUMKLANGERZEUGUNG
IM TRANSFORMIERTEN BEREICH

METHOD AND DEVICE FOR EFFICIENT BINAURAL SOUND SPATIALIZATION IN THE
TRANSFORMED DOMAIN

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

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

(30) Priorité: **28.03.2006 FR 0602685**

(56) Documents cités:
WO-A-20/05094125 FR-A- 2 851 879

(43) Date de publication de la demande:
10.12.2008 Bulletin 2008/50

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

- **KULKARNI A ET AL: "On the minimum-phase approximation of head-related transfer functions" 15 octobre 1995 (1995-10-15), APPLICATIONS OF SIGNAL PROCESSING TO AUDIO AND ACOUSTICS, 1995., IEEE ASSP WORKSHOP ON NEW PALTZ, NY, USA 15-18 OCT. 1995, NEW YORK, NY, USA, IEEE, US, PAGE (S) 84-87 , XP010154639 ISBN: 0-7803-3064-1 cité dans la demande le document en entier**

(72) Inventeurs:

- **EMERIT, Marc**
F-35000 Rennes (FR)
- **PHILIPPE, Pierrick**
F-35520 MELESSE (FR)
- **VIRETTE, David**
F-22560 Pleumeur Bodou (FR)

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

EP 2 000 002 B1

Description

[0001] L'invention est relative à la spatialisation, dite rendu 3D, de signaux audio compressés.

[0002] Une telle opération est par exemple exécutée lors de la décompression d'un signal compressé audio 3D par exemple, 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.

[0003] Ainsi, le terme « binaural » vise la restitution sur un casque stéréophonique d'un signal sonore avec néanmoins des effets de spatialisation. L'invention ne se limite toutefois pas à la technique précitée et s'applique, notamment, à des techniques dérivées du « binaural », telles que les techniques de restitution dites techniques TRANSAURAL®, c'est-à-dire sur des haut-parleurs distants. TRANSAURAL® est une marque de commerce déposée par la société COOPER BAUCK CORPORATION. De telles techniques peuvent alors utiliser une « annulation de diaphonie » (« crosstalk cancellation » en anglais), laquelle consiste à annuler les chemins acoustiques croisés, de manière à ce qu'un son, ainsi traité puis émis par les haut-parleurs, puisse n'être perçu que par une seule des deux oreilles d'un auditeur.

[0004] En conséquence, 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.

[0005] 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.

[0006] Parmi les techniques connues dans le domaine de la spatialisation sonore binaurale, différentes approches ont été proposées.

[0007] Un procédé de spatialisation sonore du genre indiqué dans le préambule de la revendication 1 ci-dessous est décrit dans la demande de brevet FR 2 851 879 A.

[0008] En particulier, la synthèse binaurale bicanale consiste, en référence à la figure 1a, à filtrer le signal des différentes sources sonores S_i que l'on souhaite positionner, à la restitution, à une position dans l'espace, par l'intermédiaire de fonctions de transfert acoustiques gauche HRTF-l et droite HRTF-r dans le domaine fréquentiel correspondant à la direction appropriée, définie en coordonnées polaires (θ_1, φ_1) . Les fonctions de transfert HRTF, pour « Head Related Transfer Functions » en anglais, précitées sont les fonctions de transfert acoustique de la tête de l'auditeur entre les positions de l'espace et le conduit auditif. On désigne en outre par « HRIR » pour « Head Related Impulse Response » leur forme temporelle. Ces fonctions peuvent en outre comporter un effet de salle.

[0009] On obtient, pour chaque source sonore S_i deux signaux gauche et droit qui sont alors additionnés aux signaux gauche et droit issus de la spatialisation des autres sources sonores, pour donner finalement les signaux L et R diffusés aux oreilles gauche et droite de l'auditeur.

[0010] Le nombre de filtres, ou fonctions de transfert, nécessaires est alors $2.N$ pour une synthèse binaurale statique et $4.N$ pour une synthèse binaurale dynamique, N désignant le nombre de sources sonore ou de flux audio à spatialiser.

[0011] Des travaux intitulés « A model of head-related transfer functions based on principal components analysis and minimum - phase reconstruction » conduits par D. Kistler et F.L. Wightman, publiés au J. Acoust. Soc. Am. 91(3) : p 1637-1647 (1992) et par A. Kulkarni 1995 « IEEE ASSP Workshop on Applications of signal Processing to Audio and Acoustics » IEEE catalog number : 95TH8144, ont permis de vérifier que les phases des HRTF peuvent se décomposer en la somme de deux termes, l'un correspondant au retard interaural et l'autre égal à la phase minimale associée au module de la HRTF.

[0012] Ainsi, pour une fonction de transfert HRTF exprimée sous la forme :

$$H(f) = |H(f)| e^{-j\varphi(f)}$$

$$\varphi(f) = \varphi_{\text{retard}}(f) + \varphi_{\text{min}}(f)$$

$\varphi_{\text{retard}}(f) = 2\pi f t$ correspond au retard interaural ;

$\varphi_{\text{min}}(f) = H(\log(|H(f)|))$ est la phase minimale associée au module du filtre H.

[0013] L'implémentation des filtres binauraux se fait, en général, sous la forme de deux filtres à phase minimale et d'un retard pur, correspondant à la différence des retards gauche et droit appliqués à l'oreille la plus éloignée de la source. Ce retard est en général implémenté à l'aide d'une ligne à retard.

[0014] Le filtre à phase minimale est un filtre à réponse impulsionnelle finie et peut être exécuté dans le domaine temporel ou fréquentiel. Des filtres à réponse impulsionnelle infinie peuvent être recherchés pour approximer le module

des filtres HRTF à phase minimale.

[0015] En ce qui concerne la binauralisation, on se place, en référence à la figure 1b, dans le cadré non limitatif d'une scène sonore spatialisée en mode 5.1, en vue de la restitution de celle-ci sur le casque audio d'un être humain HB.

[0016] Cinq haut-parleurs C : Centre, Lf : Left front, Rf : Right front, SI : Surround left, Sr : Surround right, produisent chacun un son qui est perçu par l'être humain HB sur les deux récepteurs que sont ses oreilles. On modélise les transformations subies par le son par une fonction de filtrage représentant la modification que ce son subit lors de sa propagation entre le haut-parleur qui restitue ce son et une oreille donnée.

[0017] En particulier, le son émanant du haut-parleur Lf affecte l'oreille gauche LE au travers d'un filtre HRTF A mais ce même son atteint l'oreille droite RE modifié par un filtre HRTF B.

[0018] La position des haut-parleurs par rapport à l'individu HB précités peut être symétrique ou non.

[0019] Chaque oreille reçoit donc la contribution des 5 haut-parleurs sous la forme modélisée ci-après :

$$\text{Oreille gauche LE : } BI = ALf + CC + BRf + DSI + ESr,$$

$$\text{Oreille droite RE : } Br = ARf + CC + BLf + DSr + ESI,$$

où BI est le signal binauralisé pour l'oreille gauche LE et Br est le signal binauralisé pour l'oreille droite RE.

[0020] Les filtres A, B, C, D et E sont modélisés, le plus souvent, par des filtres numériques linéaires et il faut donc, dans la configuration représentée en figure 1b, 10 fonctions de filtrage à appliquer, lesquelles peuvent être réduites à 5, compte tenu des symétries.

[0021] De manière connue en tant que telle, les opérations de filtrage précitées peuvent être réalisées dans le domaine fréquentiel, par exemple grâce à une convolution rapide exécutée dans le domaine de Fourier. On utilise alors une transformée de Fourier rapide FFT, pour « Fast Fourier Transform » en anglais, pour exécuter la binauralisation de façon efficace.

[0022] Les filtres HRTF A, B, C, D et E peuvent être simplifiés sous la forme d'un égaliseur en fréquence et d'un retard. Le filtre HRTF A peut être réalisé sous la forme d'un simple égaliseur, car il s'agit d'une trajectoire directe, alors que le filtre HRTF B inclut un retard supplémentaire. De manière classique les filtres HRTF peuvent être décomposés en un filtre à phase minimale et un retard pur. Le retard pour l'oreille la plus proche de la source peut être pris égal à zéro.

[0023] L'opération de reconstruction par décodage spatial d'une scène sonore 3D audio, à partir d'un nombre réduit de canaux transmis, telle que représentée en figure 1c, est également connue de l'état de la technique. La configuration représentée en figure 1c est celle relative au décodage d'une voie sonore codée disposant de paramètres de localisation dans le domaine fréquentiel, afin de reconstruire une scène sonore spatialisée 5.1.

[0024] La reconstruction précitée est effectuée par un décodeur spatial par sous-bandes fréquentielles, tel que représenté en figure 1c. Le signal audio codé m subit 5 étapes de traitement de spatialisation, qui sont commandées par des paramètres ou coefficients complexes de spatialisation CLD et ICC calculés par l'encodeur et qui permettent, par le biais d'opérations de décorrélation et de correction de gain, de reconstruire de façon réaliste la scène sonore composée de six canaux, les cinq canaux représentés en figure 1b, auxquels est ajouté un canal d'effet de basse fréquence lfe.

[0025] Lorsque l'on souhaite procéder à une binauralisation des canaux sonores issus d'un décodeur spatial tel que représenté en figure 1c, on est en fait contraint, à l'heure actuelle, de mettre en oeuvre un traitement selon le schéma représenté en figure 1d.

[0026] En référence au schéma précité, il apparaît nécessaire de réaliser la transformation des canaux sonores dont on dispose dans le domaine temporel, avant de procéder à la binauralisation du signal. Cette opération de retour dans le domaine temporel est symbolisée par les blocs synthétiseurs « Synth » qui exécutent l'opération de transformation fréquence-temps pour chacun des canaux issus du décodeur spatial (SD). Le filtrage par filtres HRTF peut ensuite être réalisé par les filtres A, B, C, D, E, avec ou sans application du schéma égalisé, correspondant à un filtrage classique.

[0027] Une variante de binauralisation des canaux sonores d'un décodeur spatial peut consister également, ainsi que représenté en figure 1e, à convertir chaque canal sonore délivré par le décodeur audio dans le domaine temporel par un synthétiseur « Synth » puis à exécuter l'opération de décodage spatial et de binauralisation, ou spatialisation, dans le domaine fréquentiel de Fourier, après transformation par FFT.

[0028] Dans cette hypothèse, chaque module OTT correspondant à une matrice de coefficients de décodage, doit alors être converti dans le domaine de Fourier, au prix d'une approximation, car les opérations ne sont pas effectuées dans le même domaine. En outre, la complexité est encore accrue, car l'opération de synthèse « Synth » est suivie de trois transformations FFT.

[0029] Ainsi, pour binauraliser une scène sonore issue d'un décodeur spatial, il n'existe guère d'autre possibilité que

de réaliser :

- soit 6 transformations temps-fréquence, si l'on veut réaliser la binauralisation en dehors du décodeur spatial ;
- soit une opération de synthèse suivie de 3 transformations de Fourier, FFT, si l'on veut réaliser l'opération dans le domaine FFT.

[0030] A la rigueur, une autre solution peut consister à effectuer le filtrage HRTF directement dans le domaine des sous-bandes, ainsi que représenté en figure 1f.

[0031] Toutefois, dans cette hypothèse, les filtrages HRTF sont complexes à réaliser, car ces derniers imposent l'utilisation de filtres en sous-bandes, dont la longueur minimale est fixée et qui doivent prendre en compte le phénomène de repliement spectral des sous-bandes.

[0032] L'économie introduite par la réduction d'opérations de transformation est compensée négativement par l'explosion du nombre d'opérations nécessaires pour le filtrage, en raison de l'exécution de ces opérations dans le domaine PQMF pour *Pseudo Quadrature Mirror Filter* en anglais.

[0033] La présente invention a pour objectif de remédier aux nombreux inconvénients des techniques antérieures précitées de spatialisation sonore des scènes audio 3 D, notamment de transauralisation ou de binauralisation de scènes audio 3 D.

[0034] En particulier, un objectif de la présente invention est l'exécution d'un filtrage spécifique de signaux ou canaux audio codés spatialement dans le domaine des sous-bandes fréquentielles d'un décodage spatial, afin de limiter le nombre de transformations deux à deux, tout en réduisant les opérations de filtrage au minimum, mais en conservant une bonne qualité de spatialisation source, notamment en transauralisation ou binauralisation.

[0035] Selon un aspect particulièrement remarquable de la présente invention, l'exécution du filtrage spécifique précité s'appuie sur la mise sous forme égaliseur-retard des filtres de spatialisation, transaurale ou binaurale, pour une application directe d'un filtrage par égalisation-retard dans le domaine des sous-bandes.

[0036] Un autre objectif de la présente invention est l'obtention d'une qualité de rendu 3 D très proche de celle obtenue à partir de filtres de modélisation tels que des filtres HRTF d'origine, par la seule adjonction d'un traitement spatial transaural de très basse complexité, suite à un décodage spatial classique dans le domaine transformé.

[0037] Un objectif de la présente invention est enfin une nouvelle technique de spatialisation source applicable non seulement au rendu transaural ou binaural d'un son monophonique, mais également à plusieurs sons monophoniques et notamment aux canaux multiples de sons stéréo 5.1, 6.1, 7.1, 8.1 ou supérieurs.

[0038] La présente invention a ainsi pour objet un procédé de spatialisation sonore d'une scène audio comportant un premier ensemble comprenant un nombre supérieur ou égal à l'unité de canaux audio codés spatialement sur un nombre de sous-bandes de fréquences déterminé, et décodés dans un domaine transformé, en un deuxième ensemble comprenant un nombre supérieur ou égal à deux de canaux sonores de restitution dans le domaine temporel, à partir de filtres de modélisation de la propagation acoustique des signaux audio du premier ensemble de canaux.

[0039] Conformément à l'invention ce procédé est remarquable en ce que, pour chaque filtre de modélisation converti sous forme d'au moins un gain et d'un retard applicables dans le domaine transformé, il consiste à effectuer au moins, pour chaque sous-bande fréquentielle du domaine transformé :

- un filtrage par égalisation-retard du signal en sous-bande, par application d'un gain respectivement d'un retard sur le signal en sous-bande, pour engendrer à partir des canaux codés spatialement, une composante égalisée et retardée d'une valeur déterminée dans la sous-bande fréquentielle considérée,
- une addition d'un sous-ensemble de composantes égalisées et retardées, pour créer un nombre de signaux filtrés dans le domaine transformé correspondant au nombre du deuxième ensemble, supérieur ou égal à deux, de canaux sonores de restitution dans le domaine temporel,
- une synthèse de chacun des signaux filtrés dans le domaine transformé par un filtre de synthèse, pour obtenir le deuxième ensemble de nombre supérieur ou égal à deux de signaux sonores de restitution dans le domaine temporel.

[0040] Le procédé objet de l'invention est également remarquable en ce que le filtrage par égalisation-retard du signal en sous-bande inclut au moins l'application d'un déphasage et le cas échéant d'un retard pur par mémorisation, pour l'une au moins des sous-bandes de fréquences.

[0041] Le procédé objet de l'invention est également remarquable en ce qu'il inclut un filtrage par égalisation-retard dans un domaine transformé hybride, comportant une étape supplémentaire de découpe en fréquence en sous-bandes supplémentaires, avec ou sans décimation.

[0042] Le procédé objet de l'invention est enfin remarquable en ce que pour convertir chaque filtre de modélisation en une valeur de gain respectivement de retard dans le domaine transformé, il consiste au moins à associer comme valeur de gain à chaque sous-bande une valeur réelle définie comme la moyenne du module du filtre de modélisation dans cette sous-bande et à associer comme valeur de retard à chaque sous-bande une valeur de retard correspondant

au retard de réception entre l'oreille gauche et l'oreille droite pour différentes positions.

[0043] La présente invention a corrélativement pour objet un dispositif de spatialisation sonore d'une scène audio comportant un premier ensemble comprenant un nombre, supérieur ou égal à l'unité, de canaux audio codés spatialement sur un nombre de sous-bandes de fréquences déterminé, et décodés dans un domaine transformé, en un deuxième ensemble comportant un nombre supérieur ou égal à deux de canaux sonores de restitution dans le domaine temporel, à partir de filtres de modélisation de la propagation acoustique signaux audio du premier sous-ensemble de canaux.

[0044] Conformément à l'invention ce dispositif est remarquable en ce que, pour chaque sous-bande fréquentielle d'un décodeur spatial dans le domaine transformé, ce dispositif comprend outre ce décodeur spatial :

- un module de filtrage par égalisation-retard du signal en sous-bande par application d'un gain respectivement d'un retard sur le signal en sous-bande, pour engendrer à partir de chacun des canaux audio-codés spatialement une composante égalisée et retardée d'une valeur de retard déterminée dans la sous-bande de fréquences considérée,
- un module d'addition d'un sous-ensemble de composantes égalisées et retardées pour créer un nombre de signaux filtrés dans le domaine transformé correspondant au nombre du deuxième ensemble supérieur ou égal à deux des canaux sonores de restitution dans le domaine temporel,
- un module de synthèse de chacun des signaux filtrés dans le domaine transformé pour obtenir le deuxième ensemble comprenant un nombre supérieur ou égal à deux des canaux sonores de restitution dans le domaine temporel.

[0045] Le procédé et le dispositif objets de l'invention trouvent application à l'industrie électronique des appareils audio et/ou vidéo à haute fidélité, à l'industrie des jeux audio-vidéo exécutés localement ou en ligne.

[0046] Ils seront mieux compris à la lecture de la description et à l'observation des dessins ci-après dans lesquels, outre les figures 1a à 1f relatives à l'art antérieur,

- la figure 2a représente un organigramme illustratif des étapes de mise en oeuvre du procédé de spatialisation sonore objet de l'invention ;
- la figure 2b représente à titre illustratif, une variante de mise en oeuvre du procédé objet de l'invention représenté en figure 2a, obtenu par création de sous-bandes supplémentaires, en l'absence de décimation ;
- la figure 2c représente à titre illustratif, une variante de mise en oeuvre du procédé objet de l'invention représenté en figure 2a obtenu par création de sous-bandes supplémentaires, en présence de décimation ;
- la figure 3a représente, à titre illustratif, un étage, pour une sous-bande de fréquences d'un décodeur spatial, d'un dispositif de spatialisation sonore objets de l'invention ;
- la figure 3b représente, à titre illustratif, un détail de mise en oeuvre d'un filtre par égalisation-retard permettant la mise en oeuvre du dispositif objet de l'invention représenté en figure 3a ;
- la figure 4 représente à titre illustratif, un exemple de mise en oeuvre du dispositif objet de l'invention dans lequel le calcul des filtres d'égalisation retard est délocalisé.

[0047] Une description plus détaillée du procédé de spatialisation sonore d'une scène audio conforme à l'objet de la présente invention sera maintenant donnée en liaison avec la figure 2a et les figures suivantes.

[0048] Le procédé objet de l'invention s'applique à une scène audio telle qu'une scène audio 3 D représentée par un premier ensemble comprenant un nombre N de canaux audio codés spatialement supérieur ou égal à l'unité, $N \geq 1$, sur un nombre de sous-bandes de fréquences déterminé et décodé dans un domaine transformé.

[0049] Le domaine transformé s'entend d'un domaine fréquentiel transformé tel que domaine de Fourier, domaine PQMF ou de tout domaine hybride issu de ces derniers par création de sous-bandes de fréquences supplémentaires, soumises ou non à un processus de décimation temporel.

[0050] En conséquence, les canaux audio codés spatialement constitutifs du premier ensemble N de canaux, sont représentés de manière non limitative par les canaux FI, Fr, Sr, SI, C, lfe précédemment décrits dans la description et correspondant à un mode de décodage d'une scène audio 3 D dans le domaine transformé correspondant, ainsi que décrit précédemment dans la description. Ce mode n'est autre que le mode 5.1 précédemment mentionné.

[0051] En outre, ces signaux sont décodés dans le domaine transformé précité selon un nombre de sous-bandes

déterminé propres au décodage, l'ensemble des sous-bandes étant noté $(SB_k)_{k=1}^{k=K}$
 k désigne le rang de la sous-bande considérée.

[0052] Le procédé objet de l'invention permet de transformer l'ensemble des canaux audio codés spatialement précédemment cités en un deuxième ensemble comportant un nombre, supérieur ou égal à deux, de canaux sonores de restitution dans le domaine temporel, les canaux sonores de restitution étant notés BI et Br pour les canaux binauraux gauche respectivement droit, de manière non limitative dans le cadre de la figure 2a. On comprend, en particulier, qu'en lieu et place de deux canaux binauraux, le procédé objet de l'invention s'applique à tout nombre de canaux supérieur

à deux, permettant par exemple la restitution sonore en temps réel de la scène audio 3D, ainsi que représenté et décrit dans la description en liaison avec la figure 1 b.

[0053] Selon un aspect remarquable du procédé objet de l'invention, celui-ci est mis en oeuvre à partir de filtres de modélisation de la propagation acoustique des signaux audio du premier ensemble de canaux audio codés spatialement, compte tenu d'une conversion sous forme d'au moins un gain et d'un retard applicables dans le domaine transformé, ainsi qu'il sera décrit ultérieurement dans la description. De manière non limitative, les filtres de modélisation seront désignés filtres HRTF dans la suite de la description.

[0054] La conversion précitée est notée pour chaque filtre HRTF considéré pour une sous-bande SB_k de rang k à établir une valeur de gain g_k et de retard d_k correspondant, la conversion précédente étant alors notée, ainsi que représentée en figure 2a HRTF $\Xi(g_k, d_k)$.

[0055] Compte tenu de la conversion précitée, le procédé objet de l'invention consiste, pour chaque sous-bande fréquentielle du domaine transformé de rang k , à effectuer un filtrage à l'étape A par égalisation-retard du signal en sous-bande par application d'un gain g_k respectivement d'un retard d_k sur le signal en sous-bande, pour engendrer à partir des canaux codés spatialement précités, c'est-à-dire les canaux FI, C, Fr, Sr, SI et lfe, une composante égalisée et retardée d'une valeur de retard déterminée dans la sous-bande de fréquence SB_k considérée de rang k .

[0056] Sur la figure 2a, l'opération de filtrage par égalisation-retard est notée de manière symbolique $CED_{kx} = \{FI, C, Fr, Sr, SI, lfe\}(g_{kx}, d_{kx})$.

[0057] Dans la relation symbolique précitée, FEB_{kx} désigne chaque composante égalisée et retardée obtenue par application du gain g_{kx} et du retard d_{kx} sur chacun des canaux audio codés spatialement, c'est-à-dire les canaux FI, C, Fr, Sr, SI, lfe.

[0058] En conséquence et dans la relation symbolique précitée, x , pour la sous-bande de rang k correspondant, peut prendre en fait les valeurs FI, C, Fr, Sr, SI, lfe.

[0059] L'étape A est alors suivie dans le domaine transformé d'une étape B d'addition d'un sous-ensemble de composantes égalisées et retardées pour créer un nombre de signaux filtrés dans le domaine transformé correspondant au nombre N' du deuxième ensemble, supérieur ou égal à 2, de canaux sonores de restitution dans le domaine temporel.

[0060] A l'étape B de la figure 2a, l'opération d'addition est donnée par la relation symbolique :

$$F\{FI, C, Fr, Sr, SI, lfe\} = \sum CED_{kx}.$$

[0061] Dans la relation symbolique précitée, $F\{FI, C, Fr, Sr, SI, lfe\}$ désigne le sous-ensemble des signaux filtrés dans le domaine transformé obtenu par sommation d'un sous-ensemble de composantes égalisées et retardées CED_{kx} .

[0062] A titre d'exemple non limitatif et pour fixer les idées, pour un premier ensemble comportant un nombre de canaux audio codés spatialement $N = 6$, correspondant à un mode 5.1, le sous-ensemble de composantes égalisées et retardées peut consister à additionner cinq de ces composantes égalisées et retardées pour chaque oreille pour obtenir le nombre N' égal à 2 de signaux filtrés dans le domaine transformé, ainsi qu'il sera décrit de manière plus détaillée ultérieurement dans la description.

[0063] L'étape d'addition B précitée est alors suivie d'une étape C de synthèse de chacun des signaux filtrés dans le domaine transformé par un filtre de synthèse pour obtenir le deuxième ensemble de nombre N' supérieur ou égal à deux de signaux sonores de restitution dans le domaine temporel.

[0064] A l'étape C de la figure 2a, l'opération correspondante de synthèse est représentée par la relation symbolique :

$$BI, Br = \text{Synth}(F\{FI, C, Fr, Sr, SI, lfe\})$$

[0065] D'une manière générale, on indique que le procédé objet de l'invention peut être appliqué à toute scène 3D audio composée de N variant de 1 à l'infini de voies ou canaux audio codés de façon spatiale vers N' variant de 2 à l'infini de canaux sonores de restitution.

[0066] En ce qui concerne l'étape de sommation représentée à l'étape B de la figure 2a, on indique que celle-ci consiste de manière plus spécifique à additionner un sous-ensemble de composantes retardées de façon différente par les différents retards pour engendrer les N' composantes pour chaque sous-bande.

[0067] De manière plus spécifique, on indique que le filtrage par égalisation-retard du signal en sous-bande inclut au moins l'application d'un déphasage complété le cas échéant par un retard pur par mémorisation, pour l'une au moins des sous-bandes de fréquence.

[0068] La notion d'application d'un retard pur est symbolisée à l'étape A de la figure 2a par la relation $g_{Ex} = 1$, laquelle représente l'absence d'égalisation pour l'ensemble des canaux audio d'indice x dans la sous-bande de rang $k = E$, la

valeur 1 indiquant une transmission sans modification de l'amplitude de chacun des canaux audio codés spatialement.

[0069] Le domaine transformé peut, ainsi que mentionné précédemment dans la description, correspondre à un domaine transformé hybride ainsi qu'il sera décrit en liaison avec la figure 2b dans le cas où aucune décimation en fréquence n'est appliquée dans la sous-bande correspondante.

[0070] En référence à la figure 2b précitée, le filtrage par égalisation retard représenté à l'étape A de la figure 2a est alors exécuté en trois sous-étapes A1, A2, A3 représentées à la figure 2b.

[0071] Dans ces conditions, l'étape A comporte une étape supplémentaire de découpe en fréquence en sous-bandes supplémentaires sans décimation, pour augmenter le nombre de valeurs de gain appliquées et ainsi la précision en fréquence, suivie d'une étape de regroupement de sous-bandes supplémentaires, auxquelles ont été appliquées les valeurs de gain précitées.

[0072] Les opérations de découpe en fréquence puis de regroupement sont représentées aux sous-étapes A₁ et A₂ de la figure 2b.

[0073] L'étape des découpes en fréquence est représentée à la sous-étape A₁ par la relation :

$$HRTF \equiv \{g_{kz}, dkz\}_{z=1}^{z=Z}.$$

[0074] L'étape de regroupement est représentée à la sous-étape A₂ par la relation :

$$[GCEB_{kz}]_1^z x = \{FI, C, FI, Sr, SI, lfe\} (g_{kz})$$

[0075] A la sous-étape A₁, on comprend que les valeurs de gain et de retard pour la sous-bande de rang k considérée sont subdivisées en Z valeurs de gain correspondantes, une valeur de gain g_{kz} pour chaque sous-bande supplémentaire et à la sous-étape 1₂ on comprend que le regroupement des sous-bandes supplémentaires est effectué à partir des canaux audio codés correspondants pour l'indice x correspondant auquel a été appliqué la valeur de gain g_{kz} dans la sous-bande supplémentaire considérée.

[0076] Dans la relation précédente $[GCEB_{kz}]_{z=1}^{z=Z} x$ désigne le regroupement des sous-bandes supplémentaires auxquelles ont été appliquées les valeurs de gain pour les sous-bandes supplémentaires considérées.

[0077] La sous-étape A₂ est alors suivie d'une sous-étape A₃ consistant à appliquer le retard aux sous-bandes supplémentaires regroupées et en particulier aux canaux audio codés spatialement d'indice x correspondant par l'intermédiaire du retard d_{kx} de manière semblable à l'étape A de la figure 2a.

[0078] L'opération correspondante est notée par la relation :

$$CED_{kz}x = [GCEB_{kz}]_{z=1}^{z=Z} x (dk_x).$$

[0079] En outre, le procédé objet de l'invention peut consister également à effectuer un filtrage par égalisation-retard dans un domaine transformé hybride comportant une étape supplémentaire de découpe de fréquence en sous-bandes supplémentaires avec décimation, ainsi que représentée en figure 2c.

[0080] Dans cette hypothèse, l'étape A'₁ de la figure 2c est identique à l'étape A₁ de la figure 2b, pour exécuter la création des sous-bandes supplémentaires avec décimation.

[0081] Dans cette hypothèse, l'opération de décimation à l'étape A'₁ de la figure 2c est exécutée dans le domaine temporel.

[0082] L'étape A'₁ est alors suivie d'une étape A'₂ correspondant à un regroupement des sous-bandes supplémentaires auxquelles ont été appliquées les valeurs de gain précitées compte tenu de la décimation.

[0083] L'étape A'₂ de regroupement est elle-même précédée ou suivie de l'application du retard dkx ainsi représentée par la double flèche d'intervention des étapes A'₂ et A'₃.

[0084] On comprend, en particulier, que lorsque l'application du retard est effectuée antérieurement au regroupement, le retard est appliqué directement sur les signaux des sous-bandes supplémentaires antérieurement au regroupement.

[0085] En ce qui concerne la conversion de chaque filtre HRTF en une valeur de gain et de retard dans le domaine transformé, cette opération peut consister, avantageusement, à associer, comme valeur de gain à chaque sous-bande de rang k, une valeur réelle définie comme la moyenne du module du filtre HRTF correspondant et à associer, comme

valeur de retard à chaque sous-bande de rang k , une valeur de retard correspondant au retard de propagation entre l'oreille gauche et l'oreille droite d'un auditeur pour différentes positions.

[0086] Ainsi, à partir d'un filtre HRTF, il est possible de calculer de façon automatique les gains et les délais de retard appliqués en sous-bande. A partir de la résolution fréquentielle du banc de filtre HRTF, on associe à chacune des sous-

[0087] Ainsi, à partir d'un filtre HRTF, on peut calculer de façon automatique les gains et les délais de retard à appliquer en sous-bande.

[0088] A partir de la résolution fréquentielle du banc de filtre, on associe à chacune des bandes une valeur réelle. A titre d'exemple non limitatif, il est possible à partir du module du filtre HRTF, de calculer, pour chaque sous-bande, la moyenne du module du filtre HRTF précité. Une telle opération est similaire à une analyse en bande d'octave ou de Bark des filtres HRTF. De même, on détermine le retard à appliquer pour les canaux indirects, c'est-à-dire les valeurs de retard qui sont applicables plus particulièrement aux canaux dont le retard n'est pas minimum. Il existe de nombreuses méthodes pour déterminer de manière automatique les retards interauraux encore désignés ITD pour « *Interaural Time Difference* » et qui correspondent aux retards entre l'oreille gauche et l'oreille droite, pour différentes positions de l'auditeur. On peut utiliser, à titre d'exemple non limitatif, la méthode du seuil décrite par S. Busson dans la thèse de doctorat de l'Université de la Méditerranée Est-Marseille II, 2006, intitulée « *Individualisation d'indices acoustiques pour la synthèse binaurale* ». Le principe des méthodes d'estimation du retard interaural de type seuil est de déterminer le temps d'arrivée, ou encore le retard initial de l'onde sur l'oreille droite T_d et sur l'oreille gauche T_g . Le retard interaural est donné par la relation $ITD_{\text{seuil}} = T_d - T_g$.

[0089] La méthode la plus courante estime le temps d'arrivée comme l'instant où le filtre temporel HRIR dépasse un seuil donné. Par exemple le temps d'arrivée peut correspondre au temps pour lequel la réponse du filtre HRIR atteint 10 % de son maximum.

[0090] Un exemple de mise en oeuvre spécifique dans le domaine transformé PQMF sera maintenant donné ci-après.

[0091] D'une manière générale, on indique que l'application d'un gain dans le domaine PQMF complexe consiste à multiplier la valeur de chaque échantillon du signal en sous-bande, représenté par une valeur complexe, par la valeur de gain formée par un nombre réel.

[0092] En effet, il est bien connu que l'usage d'un domaine transformé PQMF complexe, permet d'appliquer les gains en s'affranchissant des problèmes de repliement de spectre engendrés par le sous-échantillonnage inhérent aux bancs de filtres. Chaque sous-bande SB_k de chaque canal se voit ainsi affectée d'un gain déterminé.

[0093] En outre, l'application d'un retard dans le domaine transformé PQMF consiste au moins, pour chaque échantillon du signal en sous-bande, représenté par une valeur complexe, à introduire une rotation dans le plan complexe par multiplication de cet échantillon par une valeur exponentielle complexe fonction du rang de la sous-bande considérée, du taux de sous-échantillonnage dans la sous-bande considérée et d'un paramètre de retard lié à la différence de retard interaural d'un auditeur.

[0094] La rotation dans le plan complexe est alors suivie d'un retard temporel pur de l'échantillon après rotation. Ce retard temporel pur est une fonction de la différence du retard interaural d'un auditeur et du taux de sous échantillonnage dans la sous-bande considérée.

[0095] De manière pratique, on indique que les retards précités sont appliqués sur les signaux résultants c'est-à-dire les signaux égalisés et en particulier sur les sous-ensembles de ces signaux ou canaux qui ne bénéficient pas d'une trajectoire directe.

[0096] En particulier, la rotation est effectuée sous la forme d'une multiplication complexe par une valeur exponentielle de la forme :

$$\exp(-j \cdot \pi \cdot (k + 0,5) \cdot d/M)$$

et par un retard pur implémenté par une ligne à retard, par exemple réalisant l'opération :

$$y(k,n) = x(k, n-D)$$

[0097] Dans les relations précédentes :

- \exp est la fonction exponentielle ;
- j est tel que $j^2 = -1$;

- k le rang de la sous-bande SB_k considérée ;
- M est le taux de sous-échantillonnage dans la sous-bande considérée, M veut être pris égal à 64, par exemple ;
- $y(k, n)$ est la valeur de l'échantillon de sortie après application du retard pur sur l'échantillon temporel de rang n de la sous-bande SB_k de rang k, c'est-à-dire l'échantillon $x(k, n)$ auquel est appliqué le retard B.
- d et D dans les relations précédentes sont tels qu'ils correspondent à l'application d'un retard de $D \cdot M + d$ dans le domaine temporel non sous-échantillonné. Le retard $D \cdot M + d$ correspond au retard interaural calculé précédemment. d peut prendre des valeurs négatives ce qui permet de simuler une avance de phase en lieu et place d'un retard.

[0098] L'opération ainsi réalisée induit une approximation qui est convenable pour l'effet recherché.

[0099] En terme d'opérations de calcul, le traitement mis en oeuvre consiste donc à réaliser une multiplication complexe entre une exponentielle complexe et d'un échantillon en sous-bande formé par une valeur complexe.

[0100] Un retard éventuel, si le retard total à appliquer est supérieur à la valeur M, est à insérer, mais cette opération ne comporte pas d'opération arithmétique.

[0101] Le procédé objet de l'invention peut également être mis en oeuvre dans un domaine transformé hybride. Ce domaine transformé hybride est un domaine fréquentiel dans lequel les bandes PQMF sont avantageusement redécoupées par un banc de filtres décimé ou non.

[0102] Si le banc de filtres est décimé, la décimation s'entendant d'une décimation en temps, alors l'introduction d'un retard suit avantageusement la procédure incluant un retard pur et un déphaseur.

[0103] Si le banc de filtre n'est pas décimé, alors le retard peut n'être appliqué qu'une seule fois lors de la synthèse. Il est en effet inutile d'appliquer le même retard sur chacune des branches car la synthèse est une opération linéaire, sans sous-échantillonneur.

[0104] L'application des gains reste identique, ceux-ci étant simplement plus nombreux, ainsi que décrit précédemment en liaison avec la figure 2b par exemple, et permettent donc de suivre la découpe plus précise en fréquence. Un gain réel est alors appliqué par sous-bande supplémentaire.

[0105] Enfin, selon une variante de mise en oeuvre, l'on réitère le procédé selon l'invention pour au moins deux couples égalisation-retard et l'on somme les signaux obtenus pour obtenir les canaux sonores dans le domaine temporel.

[0106] Une description plus détaillée d'un dispositif de spatialisation sonore d'une scène audio comportant un premier ensemble comprenant un nombre supérieur ou égal à l'unité de canaux audio codés spatialement sur un nombre de sous-bandes de fréquence déterminé et décodé dans un domaine transformé, en un deuxième ensemble comprenant un nombre supérieur ou égal à 2 de canaux sonores de restitution dans le domaine temporel, conforme à l'objet de la présente invention, sera maintenant décrit en liaison avec les figures 3a et 3b.

[0107] Ainsi que mentionné précédemment, le dispositif objet de l'invention est basé sur le principe de la conversion sous forme d'au moins un gain et d'un retard applicable dans le domaine transformé de filtres de modélisation de la propagation acoustique des signaux audio du premier ensemble de canaux précité. Le dispositif objet de l'invention permet la spatialisation sonore d'une scène audio, telle qu'une scène audio 3D, en un deuxième ensemble comportant un nombre, supérieur ou égal à deux, de canaux sonores de restitution dans le domaine temporel.

[0108] Le dispositif objet de l'invention représenté en figure 3a concerne en étage de ce dispositif spécifique à chaque sous-bande SB_k de rang k de décodage dans le domaine transformé.

[0109] On comprend en particulier que l'étage, pour chaque sous-bande de rang k représenté en figure 3a, est en fait répliqué pour chacune des sous-bandes pour constituer finalement le dispositif de spatialisation sonore conforme à l'objet de la présente invention.

[0110] Par convention, l'étage représenté en figure 3a sera désigné ci-après dispositif de spatialisation sonore objet de l'invention.

[0111] En référence à la figure précitée, le dispositif objet de l'invention tel que représenté sur la figure 3a comporte, outre le décodeur spatial représenté, comportant les modules OTT_0 à OTT_4 correspondant sensiblement à un décodeur spatial SD de l'art antérieur tel que représenté en figure 1c, mais dans lequel on procède en outre, de manière connue en tant que telle de l'état de la technique, à une sommation du canal frontal C et du canal à fréquence basse lfe par un sommateur S, un module 1 de filtrage par égalisation-retard du signal en sous-bande par application d'un gain respectivement d'un retard sur le signal en sous-bande.

[0112] Sur la figure 3a, l'application d'un gain est représenté sur chacun des canaux audio codés spatialement, représentés par des amplificateurs 1_0 à 1_8 , ces derniers engendrant une composante égalisée laquelle peut être soumise ou non à un retard par l'intermédiaire d'éléments de retard notés 1_9 à 1_{12} pour engendrer à partir de chacun des canaux audio codés spatialement une composante égalisée et retardée d'une valeur de retard déterminé dans la sous-bande de fréquence SB_k .

[0113] En référence à la figure 3a, les gains des amplificateurs 1_0 à 1_8 ont des valeurs arbitraires A, B, B.A, C, D, E, E, D respectivement. En outre les valeurs de retard appliquées par les modules de retard 1_9 à 1_{12} ont pour valeurs Df, Bf, Ds, Ds. Sur la figure précitée, la structure des gains et retards introduits est symétrique. Une structure non symétrique peut être mise en oeuvre sans sortir du cadre de l'objet de l'invention.

[0114] Le dispositif objet de l'invention comporte également un module 2 d'addition d'un sous-ensemble de composantes égalisées et retardées pour créer un nombre de signaux filtrés dans le domaine transformé correspondant au nombre N' du deuxième ensemble supérieur ou égal à deux de canaux sonores de restitution dans le domaine temporel.

[0115] Enfin le dispositif objet de l'invention comporte un module 3 de synthèse de chacun des signaux filtrés dans le domaine transformé pour obtenir le deuxième ensemble comprenant un nombre N' supérieur ou égal à deux de signaux sonores de restitution dans le domaine temporel. Le module de synthèse 3 comporte ainsi, dans le mode de réalisation de la figure 3a, un synthétiseur 3_0 et 3_1 lesquels permettent chacun de délivrer un signal sonore de restitution dans le domaine temporel B_1 pour signal binaural gauche, respectivement B_r pour signal binaural droit.

[0116] Les composantes égalisées et retardées dans le mode de réalisation de la figure 3a sont obtenues de la manière ci-après avec :

- $A[k]$ désignant le gain des amplificateurs $1_0, 1_3$ pour la sous-bande SB_k de rang k ,
- $B[k]$ désigne le gain de l'amplificateur $1_1, 1_2$ représenté en figure 3a,
- $C[k]$ désigne le gain de l'amplificateur 1_4 ,
- $D[k]$ désigne le gain des amplificateurs $1_5, 1_8$,
- $E[k]$ désigne le gain des amplificateurs $1_6, 1_7$.

[0117] en ce qui concerne les canaux audio codés spatialement et en particulier ces canaux FI, Fr, Clfe, SI et Sr pour la sous-bande SB_k , on désigne par $FI[k][n]$, $Fr[k][n]$, $Fc[k][n]$, $lfe[k][n]$, $SI[k][n]$, $Sr[k][n]$, le n ème échantillon de la sous-bande SB_k . Ainsi chaque amplificateur, 1_0 à 1_8 délivre les composantes égalisées suivantes successivement :

- $A[k]*FI[k][n]$,
- $B[k]*FI[k][n]$,
- $B[k]*Fr[k][n]$,
- $A[k]*Fr[k][n]$,
- $C[k]*Fc[k][n]$,
- $D[k]*SI[k][n]$,
- $E[k]*SI[k][n]$,
- $E[k]*Sr[k][n]$,
- $D[k]*Sr[k][n]$.

[0118] Les opérations précédentes, ainsi que mentionné précédemment dans la description, sont réalisées sous la forme d'une multiplication réelle agissant dans ce cas sur des nombres complexes.

[0119] Les retards introduits par les éléments de retard $1_9, 1_{10}, 1_{11}$ et 1_{12} sont appliqués sur les composantes égalisées précitées pour engendrer les composantes égalisées et retardées.

[0120] Dans l'exemple représenté en figure 3a, ces retards sont appliqués sur le sous-ensemble qui ne bénéficie pas d'une trajectoire directe. Ce sont, dans la description de la figure 3a, les signaux qui ont subi les multiplications par les gains $B[k]$ et $E[k]$ appliquées par les amplificateurs ou multiplicateurs $1_1, 1_2$ et 1_6 et 1_7 .

[0121] Une description plus détaillée d'un filtre ou élément de filtrage par égalisation-retard constitué par exemple par un amplificateur multiplicateur 1_1 et un élément retardateur 1_9 sera maintenant donnée en liaison avec la figure 3b.

[0122] En ce qui concerne l'application du gain, on indique que l'élément de filtrage, correspondant, représenté en figure 3b, comporte un multiplicateur numérique, c'est-à-dire l'un des multiplicateurs ou amplificateurs 1_0 à 1_8 et représenté par la valeur de gain g_{kx} à la figure 3b, ce multiplicateur permettant la multiplication de tout échantillon complexe de chaque canal audio codé d'indice x correspondant aux canaux FI, Fr, Clfe, SI, ou Sr par une valeur réelle, c'est-à-dire la valeur de gain précédemment mentionnée dans la description.

[0123] En outre, l'élément de filtrage représenté en figure 3b comporte au moins un multiplicateur numérique complexe permettant d'introduire une rotation dans le plan complexe de tout échantillon du signal en sous-bande par une valeur exponentielle complexe, la valeur $\exp(-j \varphi(k, SS_k))$ où $\varphi(k, SS_k)$ désigne une valeur de phase fonction du taux de sous-échantillonnage de la sous-bande considérée et du rang de la sous-bande considérée k .

[0124] Dans un mode de réalisation $\varphi(k, SS_k) = \varphi^*(k+0.5)*d/M$.

[0125] Le multiplicateur numérique complexe est suivi d'une ligne à retard notée L.A.R. introduisant un retard pur de chaque échantillon après rotation, permettant d'introduire un retard temporel pur fonction de la différence du retard interaural d'un auditeur et du taux de sous-échantillonnage M dans la sous-bande SB_k considérée.

[0126] Ainsi, la ligne à retard L.A.R. permet d'introduire le retard sur l'échantillon complexe après rotation de la forme $y(k, n) = x(k, n-D)$.

[0127] Enfin, on indique que les valeurs de d et D sont telles que ces valeurs correspondent à l'application d'un retard $D*M+d$ dans le domaine temporel non échantillonné et que le retard $D*M+d$ correspond au retard interaural précédemment mentionné.

[0128] Pour la mise en oeuvre du dispositif objet de l'invention, tel que représenté en figure 3a, on peut observer que le signal $Fr[k][n]$ est multiplié par le gain $B[k]$ puis retardé, ce qui, conformément à l'un des aspects remarquable de l'objet de l'invention, revient à multiplier ce signal par un gain complexe. Le produit du gain $B[k]$ et de l'exponentielle complexe peut être réalisé une fois pour toute évitant ainsi une opération complémentaire pour chaque échantillon $Fr[k][n]$ successif. Les composantes égalisées et retardées gauches sont référencés L_0 à L_4 et droites R_0 à R_4 et représentées au dessin regroupées par les modules somateurs 2_0 respectivement 2_1 , vérifient alors les relations ci-après :

Tableau T

$L0[k][n] =$	$A[k]F1[k][n]$
$R0[k][n] =$	$B[k]F1[k][n]$ retardé de D_f échantillons
$R1[k][n] =$	$A[k]Fr[k][n]$
$L1[k][n] =$	$B[k]Fr[k][n]$ retardé de D_f échantillons
$L2[k][n] =$	$R2[k][n] = C[k] (Fc[k][n] + 1fe[k][n])$
$L3[k][n] =$	$D[k]S1[k][n]$
$R3[k][n] =$	$E[k]S1[k][n]$ retardé de D_s échantillons
$R4[k][n] =$	$D[k]Sr[k][n]$
$L4[k][n] =$	$E[k]Sr[k][n]$ retardé de D_s échantillons

[0129] Pour obtenir les canaux sonores de restitution dans le domaine temporel, à savoir les canaux B_l gauche respectivement B_r droit représentés en figure 3a c'est-à-dire des signaux binauralisés dans le mode de réalisation de la figure 3a, on additionne pour chaque échantillon de rang n les composantes égalisées et retardées spatiales c'est-à-dire l'addition des composantes :

$L0[k][n] + L1[k][n] + L2[k][n] + L3[k][n] + L4[k][n]$ pour le module sommateur 2_0 , et
 $R0[k][n] + R1[k][n] + R2[k][n] + R3[k][n] + R4[k][n]$ pour le module sommateur 2_1 .

[0130] Les signaux résultants délivrés par les modules de sommation 2_0 et 2_1 sont ensuite passés dans les bancs de filtres de synthèse 3_0 respectivement 3_1 afin d'obtenir les signaux binauralisés dans le domaine temporel B_l respectivement B_r .

[0131] Les signaux précités peuvent ensuite alimenter un convertisseur numérique-analogique, afin de permettre l'écoute des sons gauche B_l et droit B_r sur un casque d'écoute audio par exemple.

[0132] L'opération de synthèse réalisée par les modules de synthèse 3_0 et 3_1 inclut, le cas échéant, l'opération de synthèse hybride telle que décrite précédemment dans la description.

[0133] Le procédé objet de l'invention peut avantageusement consister à dissocier les opérations d'égalisation et de retard, lesquelles peuvent porter sur des sous-bandes de fréquence en nombre différent. En variante, l'égalisation peut par exemple être effectuée dans le domaine hybride et le retard dans le domaine PQMF.

[0134] On comprend que le procédé et le dispositif objets de l'invention bien que décrits pour la binauralisation de six canaux vers un casque d'écoute peuvent également s'appliquer pour effectuer la transauralisation, c'est-à-dire la restitution d'un champ sonore 3D sur une paire de hauts parleurs ou pour convertir de façon peu complexe une représentation de N canaux audio ou sources sonores issus d'un décodeur spatial ou de plusieurs décodeurs monophoniques vers N' canaux audio disponibles au niveau de la restitution. Les opérations de filtrages peuvent alors être à multiplier le cas échéant.

[0135] A titre d'exemple complémentaire non limitatif, le procédé et le dispositif objets de l'invention peuvent être appliqués au cas d'un jeu 3D interactif dans les sons émis par les différents objets ou sources sonores, lesquels peuvent alors être spatialisés en fonction de leur position relative par rapport à l'auditeur. Des échantillons sonores sont alors compressés et stockés dans différents fichiers ou différentes zones mémoires. Pour être joués et spatialisés, ils sont partiellement décodés afin de rester dans le domaine codé et sont filtrés dans le domaine codé par des filtres binauraux adéquats de manière avantageuse en utilisant le procédé d'écrit conformément à l'objet de la présente invention.

[0136] En effet, en regroupant les opérations de décodage et de spatialisation, la complexité globale du processus est fortement réduite sans toutefois entraîner de perte de qualité.

[0137] L'invention couvre enfin un programme d'ordinateur comportant une suite d'instructions mémorisées sur un support de mémorisation pour exécution par un ordinateur ou un dispositif dédié de spatialisation sonore, lequel lors de cette exécution, exécute les étapes de filtrage d'addition et de synthèse telles que décrite en liaison avec les figures 2a à 2c et 3a, 3b précédemment dans la description.

[0138] On comprend en particulier que les opérations représentées aux figures précitées peuvent avantageusement être mises en oeuvre sur des échantillons numériques complexes par l'intermédiaire d'une unité centrale de traitement,

d'une mémoire de travail et d'une mémoire de programme, non représentées au dessin de la figure 3a.

[0139] Enfin, le calcul des gains et des retards constituant les filtres d'égalisation-retard peut être exécuté de manière externe au dispositif objet de l'invention représenté en figure 3a et 3b, ainsi qu'il sera décrit ci-après en liaison avec la figure 4.

[0140] En référence à la figure précitée, on considère une première unité de codage spatial et de codage à réduction de débit I, incluant un dispositif objet de l'invention tel que représenté en figure 3a, 3b, permettant d'opérer le codage spatial précité à partir d'une scène audio en mode 5.1 par exemple et la transmission audio codé, d'une part, et de paramètres spatiaux, d'autre part, vers une unité de décodage et de décodage spatial II.

[0141] Le calcul des filtres d'égalisation retard peut alors être effectué par une unité distincte III, laquelle à partir des filtres de modélisation, filtres HRTF, calcule les valeurs d'égalisation de gain et de retard et les transmet à l'unité I de codage spatial et à l'unité II de décodage spatial.

[0142] Le codage spatial peut ainsi prendre en compte les HRTF qui seront appliquées pour corriger ses paramètres spatiaux et améliorer le rendu 3D. De même le codeur à réduction de débit pourra se servir de ces HRTF pour mesurer les effets perceptifs d'une quantification en fréquence.

[0143] Côté décodage ce sont les HRTF transmises qui seront appliquées dans le décodeur spatial, et permettront le cas échéant de reconstruire les voies restituées.

[0144] Comme dans les exemples précédents, ce sont 2 voies à partir de 5 qui seront restituées, mais d'autres cas peuvent inclure la construction de 5 voies à partir de 3 comme illustré ci-dessus. Le procédé de décodage spatial procédera alors comme suit :

- projection des 3 canaux reçus sur un ensemble de canaux virtuels (supérieur aux 5 de sortie) en utilisant les informations spatiales (upmix) ;
- réduction des canaux virtuels aux 5 canaux de sortie en utilisant les HRTF.

[0145] Si les HRTF ont été appliquées au codeur, alors on pourra éventuellement supprimer leur contribution avant upmix pour réaliser le schéma ci-dessus.

[0146] Les HRTF après conversion sous leur forme gain / retard, peuvent être quantifiées de façon privilégiées sous la forme suivante : codage en différentiel de leurs valeurs puis quantification de leurs différences : si on appelle $G[k]$ les valeurs des gains de l'égaliseur, alors on transmettra les valeurs quantifiées :

$$e[k]=G[k+1]-G[k],$$

linéairement ou logarithmiquement.

[0147] De manière plus spécifique en référence à la figure 4 précitée le processus mis en oeuvre par le dispositif et le procédé objets de l'invention permet ainsi d'exécuter une spatialisation sonore d'une scène audio dans laquelle le premier ensemble comporte un nombre déterminé de canaux audio codés spatialement et, le deuxième ensemble comporte un nombre inférieur de canaux sonores de restitution dans le domaine temporel. Il permet en outre au décodage d'effectuer une transformation inverse d'un nombre de canaux audio codés spatialement vers un ensemble comportant un nombre supérieur ou égal de canaux sonores de restitution dans le domaine temporel.

Revendications

1. Procédé de spatialisation sonore d'une scène audio comportant un premier ensemble comprenant un nombre, supérieur ou égal à l'unité, de canaux audio codés spatialement sur un nombre de sous-bandes de fréquences déterminé, et décodés dans un domaine transformé, en un deuxième ensemble comprenant un nombre supérieur ou égal à deux de canaux sonores de restitution dans le domaine temporel, à partir de filtres de modélisation de la propagation acoustique des signaux audio dudit premier ensemble de canaux, **caractérisé en ce que**, pour chaque filtre de modélisation converti sous forme d'au moins un gain et d'un retard applicables dans ledit domaine transformé, ledit procédé inclut au moins, pour chaque sous-bande fréquentielle dudit domaine transformé :

- le filtrage par égalisation-retard du signal en sous-bande par application d'un gain respectivement d'un retard sur ledit signal en sous-bande, pour engendrer, à partir des canaux codés spatialement, une composante égalisée et retardée d'une valeur de retard déterminé dans la sous-bande de fréquences considérée ;
- l'addition d'un sous-ensemble de composantes égalisées et retardées, pour créer un nombre de signaux filtrés dans le domaine transformé correspondant au nombre dudit deuxième ensemble supérieur ou égal à deux de

canaux sonores de restitution dans le domaine temporel ;

- la synthèse de chacun des signaux filtrés dans le domaine transformé par un filtre de synthèse, pour obtenir ledit deuxième ensemble de nombre supérieur ou égal à deux de canaux sonores de restitution dans le domaine temporel.

5

2. Procédé selon la revendication 1, **caractérisé en ce que** ledit filtrage par égalisation-retard du signal en sous-bande inclut au moins l'application d'un déphasage pour l'une au moins des sous-bandes de fréquences.

10

3. Procédé selon la revendication 2, **caractérisé en ce que** ledit filtrage par égalisation retard inclut en outre un retard pur par mémorisation pour l'une au moins des sous-bandes de fréquences.

15

4. Procédé selon l'une des revendications 1 à 3, **caractérisé en ce que** ledit filtrage par égalisation-retard dans un domaine transformé hybride, comporte une étape supplémentaire de découpe en fréquence en sous-bandes supplémentaires sans décimation, pour augmenter le nombre de valeurs de gain appliquées, suivie d'une étape de regroupement desdites sous-bandes supplémentaires auxquelles ont été appliquées lesdites valeurs de gain, puis d'application dudit retard.

20

5. Procédé selon l'une des revendications 1 à 3, **caractérisé en ce que** ledit filtrage par égalisation-retard dans un domaine transformé hybride comporte une étape supplémentaire de découpe en fréquence en sous-bandes supplémentaires avec décimation, pour augmenter le nombre de valeurs de gain appliquées, suivie d'une étape de regroupement desdites sous-bandes supplémentaires auxquelles ont été appliquées lesdites valeurs de gain, ladite étape de regroupement étant elle-même précédée ou suivie de l'application dudit retard.

25

6. Procédé selon l'une des revendications précédentes, **caractérisé en ce que**, pour convertir chaque filtre de modélisation en une valeur de gain respectivement de retard dans le domaine transformé, celui-ci consiste au moins à :

- associer comme valeur de gain à chaque sous-bande une valeur réelle définie comme la moyenne du module du filtre de modélisation ;

30

- associer comme valeur de retard à chaque sous-bande une valeur de retard correspondant au retard de propagation entre l'oreille gauche et l'oreille droite pour différentes positions.

35

7. Procédé selon l'une des revendications 1 à 3 ou 6, à l'exclusion des revendications 4 ou 5, **caractérisé en ce que** l'application d'un gain dans le domaine PQMF consiste à multiplier la valeur de chaque échantillon du signal en sous-bande, représenté par une valeur complexe, par la valeur de gain formée par un nombre réel.

40

8. Procédé selon l'une des revendications 1 à 3 ou 6 ou 7, à l'exclusion des revendications 4 ou 5, **caractérisé en ce que** l'application d'un retard dans le domaine transformé PQMF consiste au moins, pour chaque échantillon du signal en sous-bande, représenté par une valeur complexe, à :

- introduire une rotation dans le plan complexe par multiplication de cet échantillon par une valeur exponentielle complexe fonction du rang de la sous-bande considérée, du taux de sous échantillonnage dans la sous-bande considérée, et d'un paramètre de retard lié à la différence de retard interaural d'un auditeur ;

45

- introduire un retard temporel pur de l'échantillon après rotation, ledit retard temporel pur étant une fonction de la différence du retard interaural d'un auditeur et du taux de sous-échantillonnage dans la sous-bande considérée.

50

9. Procédé selon l'une des revendications 1 à 8, **caractérisé en ce que** pour une spatialisation sonore binaurale d'une scène audio dans laquelle le premier ensemble comporte un nombre de canaux audio codés spatialement égal à $N=6$, en mode 5.1, ledit deuxième ensemble comporte deux canaux sonores de restitution dans le domaine temporel, pour une restitution par un casque d'écoute audio.

55

10. Procédé selon l'une des revendications 1 à 9, **caractérisé en ce que** l'on réitère le procédé pour au moins deux couples égalisation-retard et l'on somme les signaux obtenus pour obtenir les canaux sonores dans le domaine temporel.

11. Procédé selon l'une des revendications 1 à 9, **caractérisé en ce que** pour une spatialisation sonore d'une scène audio dans laquelle le premier ensemble comporte un nombre déterminé de canaux audio codés spatialement et de le deuxième ensemble comporte un nombre inférieur de canaux sonores de restitution dans le domaine temporel,

ce procédé consiste, au décodage, à effectuer une transformation inverse d'un nombre de canaux audio codés spatialement vers un ensemble comportant un nombre supérieur ou égal de canaux sonores de restitution dans le domaine temporel.

5 12. Procédé selon l'une des revendications précédentes, **caractérisé en ce que** les valeurs de gain et de retard associées au filtre de modélisation sont transmises sous forme quantifiée.

10 13. Dispositif de spatialisation sonore d'une scène audio comportant un premier ensemble comprenant un nombre, supérieur ou égal à l'unité, de canaux audio codés spatialement sur un nombre de sous-bandes de fréquences déterminé, et décodés dans un domaine transformé, en un deuxième ensemble comprenant un nombre supérieur ou égal à deux de canaux sonores de restitution dans le domaine temporel, à partir de filtres de modélisation de la propagation acoustique des signaux audio dudit premier ensemble de canaux, **caractérisé en ce que**, pour chaque sous-bande fréquentielle d'un décodeur spatial, dans le domaine transformé, ledit dispositif comprend, outre ce décodeur spatial :

15 - des moyens de filtrage par égalisation-retard du signal en sous-bande par application d'au moins un gain respectivement d'un retard sur ledit signal en sous-bande, pour engendrer, à partir de chacun des canaux audio codés spatialement une composante égalisée et retardée d'une valeur de retard déterminé dans la sous-bande de fréquences considérée ;

20 - des moyens d'addition d'un sous-ensemble de composantes égalisées et retardées, pour créer un nombre de signaux filtrés dans le domaine transformé correspondant au nombre dudit deuxième ensemble supérieur ou égal à deux de canaux sonores de restitution dans le domaine temporel ;

25 - des moyens de synthèse de chacun des signaux filtrés dans le domaine transformé, pour obtenir ledit deuxième ensemble comprenant un nombre supérieur ou égal à deux de signaux sonores de restitution dans le domaine temporel.

30 14. Dispositif selon la revendication 13, **caractérisé en ce que** lesdits moyens de filtrage par application d'un gain comportent un multiplicateur numérique de tout échantillon complexe de chaque canal audio codé spatialement par une valeur réelle.

35 15. Dispositif selon la revendication 13 ou 14, **caractérisé en ce que** lesdits moyens de filtrage par application d'un retard comportent au moins un multiplicateur numérique complexe, permettant d'introduire une rotation dans le plan complexe de tout échantillon du signal en sous-bande par une valeur exponentielle complexe, fonction du rang de la sous-bande considérée, du taux de sous-échantillonnage dans la sous-bande considérée et d'un paramètre de retard lié à la différence de retard interaural d'un auditeur.

40 16. Dispositif selon la revendication 15, **caractérisé en ce que** lesdits moyens de filtrage comportent en outre une ligne à retard pur de chaque échantillon après rotation, permettant d'introduire un retard temporel pur fonction de la différence du retard interaural d'un auditeur et du taux de sous-échantillonnage dans la sous-bande considérée.

45 17. Programme d'ordinateur comportant une suite d'instructions mémorisées sur un support de mémorisation pour exécution par un ordinateur ou un dispositif dédié, **caractérisé en ce que** lors de cette exécution, ledit programme exécute les étapes de filtrage, d'addition et de synthèse selon l'une des revendications 1 à 12.

Claims

50 1. Method of sound spatialization of an audio scene comprising a first set, having a number, greater than or equal to unity, of audio channels spatially coded on a determined number of frequency sub-bands and decoded in a transformed domain, into a second set having a number greater than or equal to two of sound reproduction channels in the time domain, on the basis of filters for modelling the acoustic propagation of the audio signals of said first set of channels, **characterized in that**, for each modelling filter converted into the form of at least one gain and one delay which are applicable in said transformed domain, said method includes at least, for each frequency sub-band of said said transformed domain:

55 - the filtering by equalization-delay of the sub-band signal by applying a gain respectively a delay to said sub-band signal, so as to produce, on the basis of the spatially coded channels, an equalized component delayed by a determined delay value in the frequency sub-band considered;

- the addition of a subset of equalized and delayed components, so as to create a number of filtered signals in the transformed domain corresponding to the number of said second set greater than or equal to two of sound reproduction channels in the time domain;
- the synthesis of each of the filtered signals in the transformed domain by a synthesis filter, so as to obtain said second set in number greater than or equal to two of sound reproduction channels in the time domain.

2. Method according to Claim 1, **characterized in that** said filtering by equalization-delay of the sub-band signal includes at least the application of a phase shift for one at least of the frequency sub-bands.

3. Method according to Claim 2, **characterized in that** said filtering by equalization-delay furthermore includes a pure delay by storage for one at least of the frequency sub-bands.

4. Method according to one of Claims 1 to 3, **characterized in that** said filtering by equalization-delay in a hybrid transformed domain, comprises an additional step of frequency splitting into additional sub-bands without decimation, so as to increase the number of gain values applied, followed by a step of grouping said additional sub-bands to which said gain values have been applied, and then of applying said delay.

5. Method according to one of Claims 1 to 3, **characterized in that** said filtering by equalization-delay in a hybrid transformed domain comprises an additional step of frequency splitting into additional sub-bands with decimation, so as to increase the number of gain values applied, followed by a step of grouping said additional sub-bands to which said gain values have been applied, said grouping step itself being preceded or followed by the application of said delay.

6. Method according to one of the preceding claims, **characterized in that**, to convert each modelling filter into a value of gain respectively of delay in the transformed domain, the latter consists at least in:

- associating as gain value with each sub-band a real value defined as the mean of the modulus of the modelling filter;
- associating as delay value with each sub-band a delay value corresponding to the propagation delay between the left ear and the right ear for various positions.

7. Method according to one of Claims 1 to 3 or 6, with the exclusion of Claims 4 or 5, **characterized in that** the application of a gain in the PQMF domain consists in multiplying the value of each sample of the sub-band signal, represented by a complex value, by the gain value formed by a real number.

8. Method according to one of Claims 1 to 3 or 6 or 7, with the exclusion of Claims 4 or 5, **characterized in that** the application of a delay in the PQMF transformed domain consists at least, for each sample of the sub-band signal, represented by a complex value, in:

- introducing a rotation in the complex plane by multiplying this sample by a complex exponential value dependent on the rank of the sub-band considered, on the rate of sub-sampling in the sub-band considered, and on a delay parameter related to the difference in interaural delay of a listener;
- introducing a pure time delay of the sample after rotation, said pure time delay being a function of the difference of the interaural delay of a listener and of the rate of sub-sampling in the sub-band considered.

9. Method according to one of Claims 1 to 8, **characterized in that** for a binaural sound spatialization of an audio scene in which the first set comprises a number of spatially coded audio channels equal to N=6, in 5.1 mode, said second set comprises two sound reproduction channels in the time domain, for playback by an audio headset.

10. Method according to one of Claims 1 to 9, **characterized in that** the method is repeated for at least two equalization-delay pairs and the signals obtained are summed so as to obtain the sound channels in the time domain.

11. Method according to one of Claims 1 to 9, **characterized in that** for a sound spatialization of an audio scene in which the first set comprises a determined number of spatially coded audio channels and the second set comprises a lesser number of sound reproduction channels in the time domain, this method consists, on decoding, in performing an inverse transformation of a number of spatially coded audio channels to a set comprising a higher or equal number of sound reproduction channels in the time domain.

12. Method according to one of the preceding claims, **characterized in that** the gain and delay values associated with the modelling filter are transmitted in quantized form.

13. Device for the sound spatialization of an audio scene comprising a first set, having a number, greater than or equal to unity, of audio channels spatially coded on a determined number of frequency sub-bands and decoded in a transformed domain, into a second set having a number greater than or equal to two of sound reproduction channels in the time domain, on the basis of filters for modelling the acoustic propagation of the audio signals of said first set of channels, **characterized in that**, for each frequency sub-band of a spatial decoder, in the transformed domain, said device comprises, in addition to this spatial decoder:

- means for the filtering by equalization-delay of the sub-band signal by applying at least one gain respectively one delay to said sub-band signal, so as to produce, on the basis of each of the spatially coded audio channels an equalized component delayed by a determined delay value in the frequency sub-band considered;
- means for adding a subset of equalized and delayed components, so as to create a number of filtered signals in the transformed domain corresponding to the number of said second set greater than or equal to two of sound reproduction channels in the time domain;
- means for the synthesis of each of the filtered signals in the transformed domain, so as to obtain said second set having a number greater than or equal to two of sound playback signals in the time domain.

14. Device according to Claim 13, **characterized in that** said means for filtering by applying a gain comprise a digital multiplier of any complex sample of each spatially coded audio channel by a real value.

15. Device according to Claim 13 or 14, **characterized in that** said means for filtering by applying a delay comprise at least one complex digital multiplier, making it possible to introduce a rotation in the complex plane of any sample of the sub-band signal by a complex exponential value, dependent on the rank of the sub-band considered, on the rate of sub-sampling in the sub-band considered and on a delay parameter related to the difference in interaural delay of a listener.

16. Device according to Claim 15, **characterized in that** said filtering means furthermore comprise a pure delay line of each sample after rotation, making it possible to introduce a pure time delay dependent on the difference of the interaural delay of a listener and of the sub-sampling rate in the sub-band considered.

17. Computer program comprising a series of instructions stored on a storage medium for execution by a computer or a dedicated device, **characterized in that** during this execution, said program executes the filtering, addition and synthesis steps according to one of Claims 1 to 12.

Patentansprüche

1. Verfahren zur akustischen Verräumlichung einer Audioszene, die eine erste Einheit aufweist, welche eine Anzahl, größer als der oder gleich dem Einheitswert, von Audiokanälen enthält, die räumlich über eine bestimmte Anzahl von Frequenz-Unterbändern codiert sind und in einem transformierten Bereich decodiert werden, in eine zweiten Einheit, die eine Anzahl größer als oder gleich zwei von Wiedergabe-Tonkanälen im Zeitbereich enthält, ausgehend von Modellierungsfiltren der akustischen Ausbreitung der Audiosignale der ersten Einheit von Kanälen, **dadurch gekennzeichnet, dass** für jedes Modellierungsfiltren, das in Form mindestens einer Verstärkung und einer Verzögerung konvertiert ist, die im transformierten Bereich anwendbar sind, das Verfahren für jedes Frequenz-Unterband des transformierten Bereichs mindestens umfasst:

- die Filterung durch Entzerrung-Verzögerung des Unterband-Signals durch Anwendung einer Verstärkung bzw. einer Verzögerung an das Unterband-Signal, um ausgehend von den räumlich codierten Kanälen eine Komponente zu erzeugen, die entzerrt und um einen Verzögerungswert verzögert ist, der im betrachteten Frequenz-Unterband bestimmt wird;
- die Hinzufügung einer Untereinheit von entzerrten und verzögerten Komponenten, um eine Anzahl von gefilterten Signalen im transformierten Bereich zu erzeugen, die der Anzahl der zweiten Einheit größer als oder gleich zwei von Wiedergabe-Tonkanälen im Zeitbereich entspricht;
- die Synthese jedes der gefilterten Signale im transformierten Bereich durch ein Synthesefilter, um die zweite Einheit einer Anzahl größer als oder gleich zwei von Wiedergabe-Tonkanälen im Zeitbereich zu erhalten.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** die Filterung durch Entzerrung-Verzögerung des Unterband-Signals mindestens die Anwendung einer Phasenverschiebung für mindestens eines der Frequenz-Unterbänder umfasst.
- 5 3. Verfahren nach Anspruch 2, **dadurch gekennzeichnet, dass** die Filterung durch Entzerrung-Verzögerung außerdem eine reine Verzögerung durch Speicherung für mindestens eines der Frequenz-Unterbänder umfasst.
4. Verfahren nach einem der Ansprüche 1 bis 3, **dadurch gekennzeichnet, dass** die Filterung durch Entzerrung-Verzögerung in einem hybriden transformierten Bereich einen zusätzlichen Schritt der Frequenzzzerlegung in zusätzliche Unterbänder ohne Frequenzherabsetzung aufweist, um die Anzahl von angewendeten Verstärkungswerten zu erhöhen, gefolgt von einem Schritt der Zusammenfassung der zusätzlichen Unterbänder, an die die Verstärkungswerte angewendet wurden, dann Anwendung der Verzögerung.
- 10 5. Verfahren nach einem der Ansprüche 1 bis 3, **dadurch gekennzeichnet, dass** die Filterung durch Entzerrung-Verzögerung in einem hybriden transformierten Bereich einen zusätzlichen Schritt der Frequenzzzerlegung in zusätzliche Unterbänder mit Frequenzherabsetzung aufweist, um die Anzahl von angewendeten Verstärkungswerten zu erhöhen, gefolgt von einem Schritt der Zusammenfassung der zusätzlichen Unterbänder, an die die Verstärkungswerte angewendet wurden, wobei die Anwendung der Verzögerung vor oder nach dem Schritt der Zusammenfassung selbst liegt.
- 15 6. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass**, um jedes Modellierungsfiler in einen Verstärkungs- bzw. Verzögerungswert im transformierten Bereich zu konvertieren, dieses mindestens darin besteht:
 - 25 - als Verstärkungswert jedem Unterband einen realen Wert zuzuordnen, der als der Mittelwert des Moduls des Modellierungsfilters definiert wird;
 - als Verzögerungswert jedem Unterband einen Verzögerungswert entsprechend der Ausbreitungsverzögerung zwischen dem linken Ohr und dem rechten Ohr für verschiedene Stellungen zuzuordnen.
- 30 7. Verfahren nach einem der Ansprüche 1 bis 3 oder 6, ausschließlich der Ansprüche 4 oder 5, **dadurch gekennzeichnet, dass** die Anwendung einer Verstärkung im PQMF-Bereich darin besteht, den Wert jeder Tastprobe des Unterband-Signals, dargestellt durch einen komplexen Wert, mit dem von einer realen Zahl geformten Verstärkungswert zu multiplizieren.
- 35 8. Verfahren nach einem der Ansprüche 1 bis 3 oder 6 oder 7, ausschließlich der Ansprüche 4 oder 5, **dadurch gekennzeichnet, dass** die Anwendung einer Verzögerung im transformierten PQMF-Bereich für jede Tastprobe des Unterband-Signals, dargestellt durch einen komplexen Wert, mindestens darin besteht:
 - 40 - eine Rotation in der komplexen Ebene durch Multiplizieren dieser Tastprobe mit einem komplexen Exponentialwert abhängig vom Rang des betrachteten Unterbands, vom Unterabtastungsgrad im betrachteten Unterband und von einem Verzögerungsparameter verbunden mit der interauralen Verzögerungsdifferenz eines Hörers einzuführen;
 - eine reine Zeitverzögerung der Tastprobe nach Rotation einzuführen, wobei die reine Zeitverzögerung eine Funktion der Differenz der interauralen Verzögerung eines Hörers und des Unterabtastungsgrads im betrachteten Unterband ist.
- 45 9. Verfahren nach einem der Ansprüche 1 bis 8, **dadurch gekennzeichnet, dass** für eine binaurale akustische Verräumlichung einer Audioszene, bei der die erste Einheit eine Anzahl von räumlich codierten Kanälen gleich $N=6$ im Modus 5.1 aufweist, die zweite Einheit zwei Wiedergabe-Tonkanäle im Zeitbereich für eine Wiedergabe durch einen Audio-Kopfhörer aufweist.
- 50 10. Verfahren nach einem der Ansprüche 1 bis 9, **dadurch gekennzeichnet, dass** das Verfahren für mindestens zwei Entzerrungs-Verzögerungs-Paare wiederholt wird und die erhaltenen Signale summiert werden, um die Tonkanäle im Zeitbereich zu erhalten.
- 55 11. Verfahren nach einem der Ansprüche 1 bis 9, **dadurch gekennzeichnet, dass** für eine akustische Verräumlichung einer Audioszene, bei der die erste Einheit eine bestimmte Anzahl von räumlich codierten Audiokanälen und die zweite Einheit eine geringere Anzahl von Wiedergabe-Tonkanälen im Zeitbereich aufweist, dieses Verfahren beim

Decodieren darin besteht, eine umgekehrte Transformation einer Anzahl von räumlich codierten Tonkanälen in eine Einheit durchzuführen, die eine höhere oder gleiche Anzahl von Wiedergabe-Tonkanälen im Zeitbereich aufweist.

12. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die dem Modellierungsfiler zugeordneten Verstärkungs- und Verzögerungswerte in quantifizierter Form übertragen werden.

13. Vorrichtung zur akustischen Verräumlichung einer Audioszene, die eine erste Einheit aufweist, welche eine Anzahl, größer als der oder gleich dem Einheitswert, von Audiokanälen aufweist, die räumlich auf eine bestimmte Anzahl von Frequenz-Unterbändern codiert sind und in einem transformierten Bereich decodiert werden, in eine zweiten Einheit, die eine Anzahl größer als oder gleich zwei von Wiedergabe-Tonkanälen im Zeitbereich enthält, ausgehend von Modellierungsfiltren der akustischen Ausbreitung der Audiosignale der ersten Einheit von Kanälen, **dadurch gekennzeichnet, dass** für jedes Frequenz-Unterband eines räumlichen Decodierers im transformierten Bereich die Vorrichtung außer diesem räumlichen Decodierer enthält:

- Einrichtungen zur Filterung durch Entzerrung-Verzögerung des Unterband-Signals durch Anwendung mindestens einer Verstärkung bzw. einer Verzögerung an das Unterband-Signal, um ausgehend von jedem der räumlich codierten Audiokanäle eine entzerrte und verzögerte Komponente eines bestimmten Verzögerungswerts im betrachteten Frequenz-Unterband zu erzeugen;
- Einrichtungen zum Hinzufügen einer Untereinheit von entzerrten und verzögerten Komponenten, um eine Anzahl von gefilterten Signalen im transformierten Bereich zu erzeugen, die der Anzahl der zweiten Einheit größer als oder gleich zwei von Wiedergabe-Tonkanälen im Zeitbereich entspricht;
- Einrichtungen zur Synthese jedes der gefilterten Signale im transformierten Bereich, um die zweite Einheit zu erhalten, die eine Anzahl höher als oder gleich zwei Wiedergabe-Tonsignalen im Zeitbereich enthält.

14. Vorrichtung nach Anspruch 13, **dadurch gekennzeichnet, dass** die Einrichtungen zur Filterung durch Anwendung einer Verstärkung einen digitalen Multiplikator jeder komplexen Tastprobe jedes räumlich codierten Audiokanals mit einem realen Wert aufweisen.

15. Vorrichtung nach Anspruch 13 oder 14, **dadurch gekennzeichnet, dass** die Einrichtungen zur Filterung durch Anwendung einer Verzögerung mindestens einen komplexen digitalen Multiplikator aufweisen, der es ermöglicht, eine Rotation in der komplexen Ebene jeder Tastprobe des Unterband-Signals um einen komplexen Exponentialwert, abhängig vom Rang des betrachteten Unterbands, vom Unterabtastungsgrad im betrachteten Unterband und von einem Verzögerungsparameter, der mit der interauralen Verzögerungsdifferenz eines Hörers verbunden ist, einzuführen.

16. Vorrichtung nach Anspruch 15, **dadurch gekennzeichnet, dass** die Filtereinrichtungen außerdem eine reine Verzögerungsleitung jeder Tastprobe nach Rotation aufweisen, die es ermöglicht, eine reine Zeitverzögerung abhängig von der Differenz der interauralen Verzögerung eines Hörers und vom Unterabtastungsgrad im betrachteten Unterband einzuführen.

17. Computerprogramm, das eine Folge von auf einem Speicherträger gespeicherten Anweisungen für die Ausführung durch einen Computer oder eine dedizierte Vorrichtung enthält, **dadurch gekennzeichnet, dass** bei dieser Ausführung das Programm die Schritte der Filterung, des Hinzufügens und der Synthese nach einem der Ansprüche 1 bis 12 durchführt.

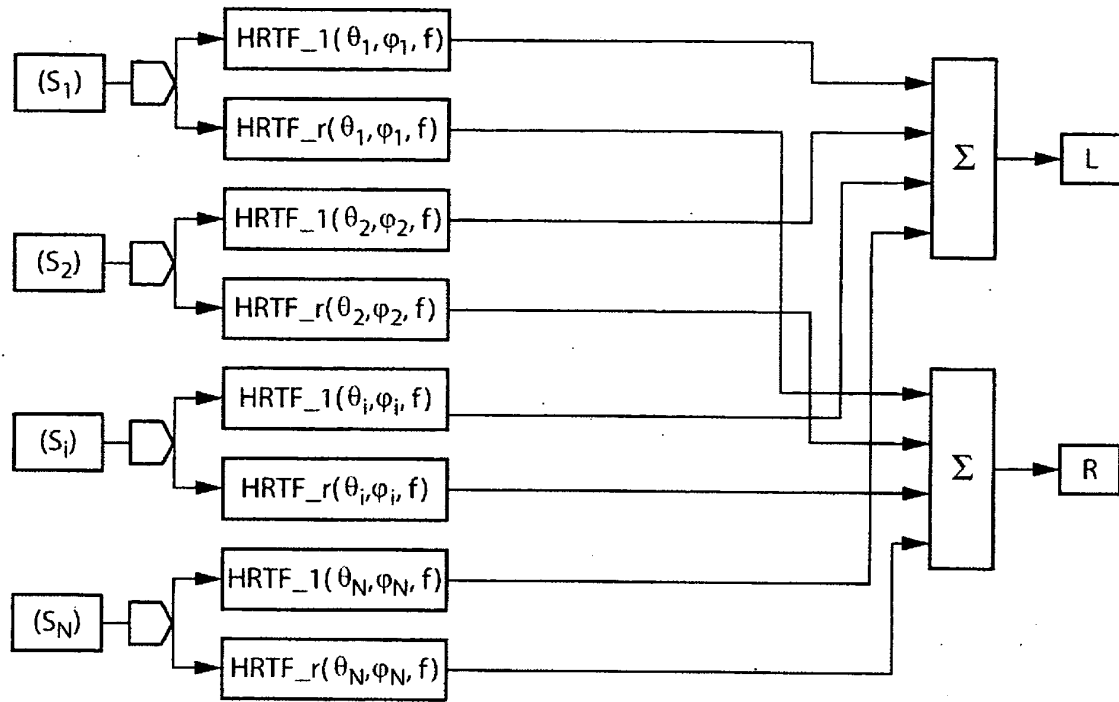


FIG. 1a
(ART ANTÉRIEUR)

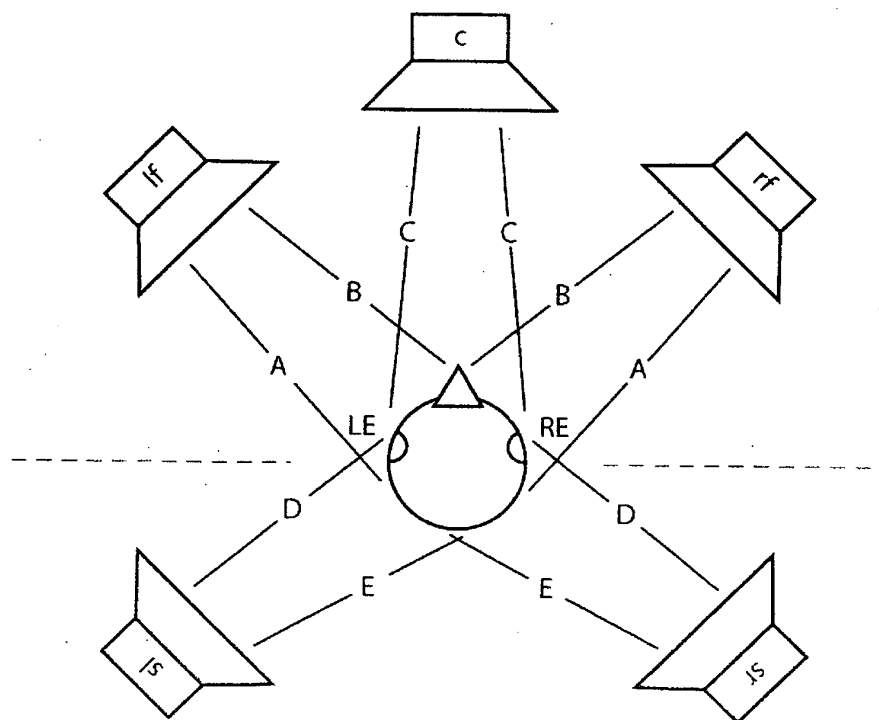


FIG. 1b
(ART ANTÉRIEUR)

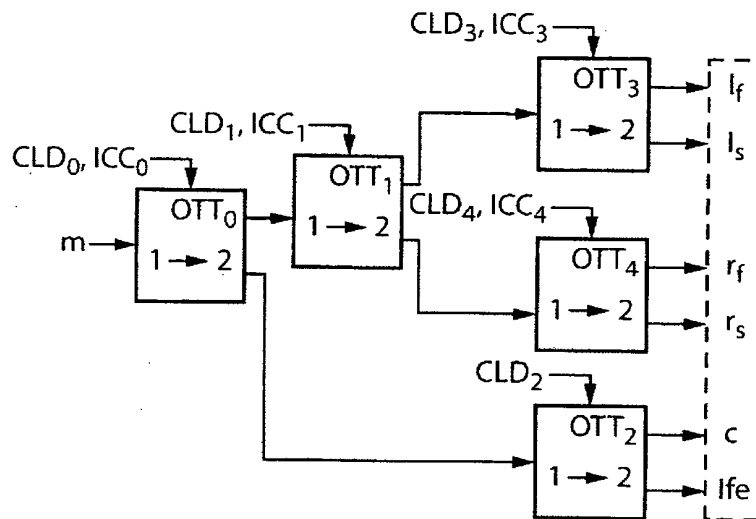


FIG. 1c
(ART ANTÉRIEUR)

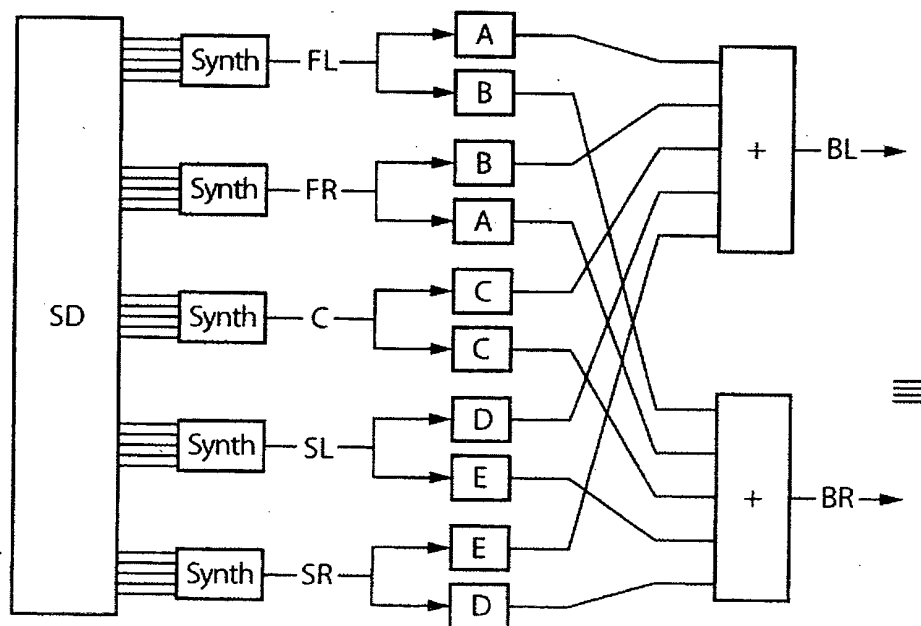


FIG. 1d
(ART ANTÉRIEUR)

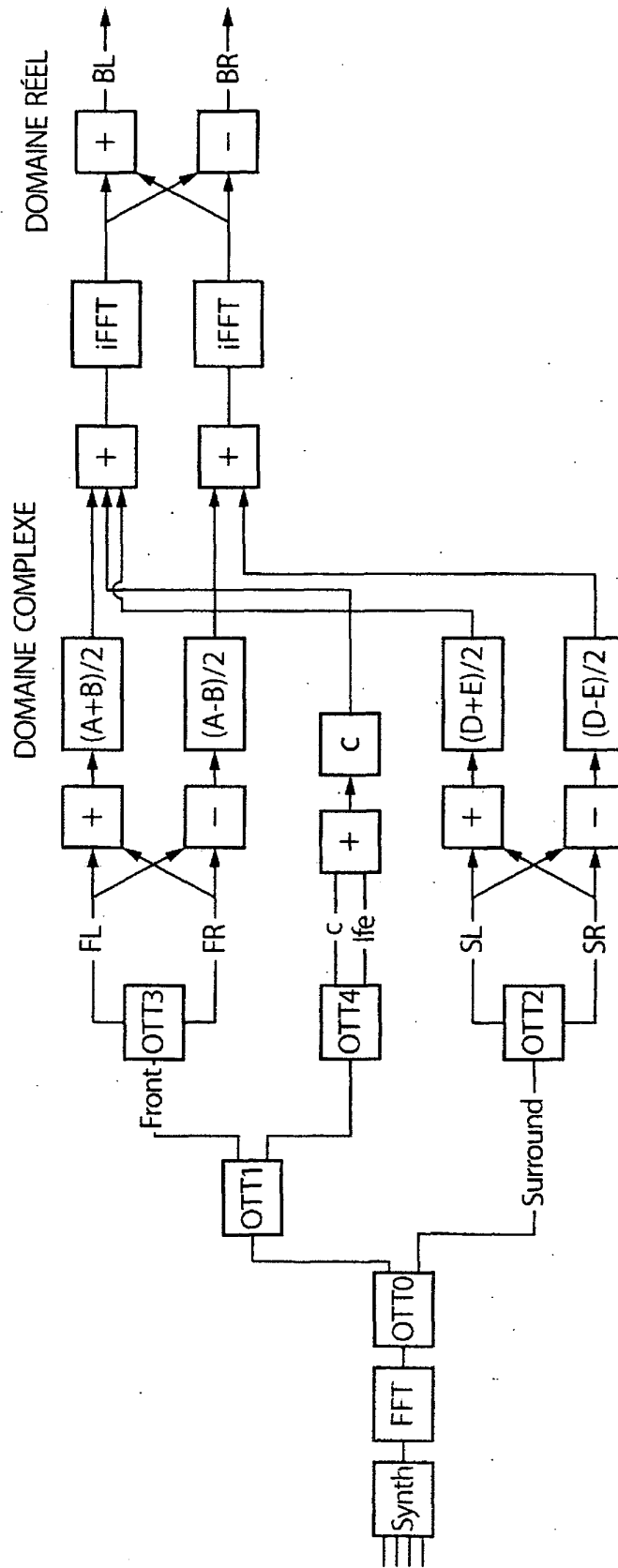


FIG. 1e
(ART ANTÉRIEUR)

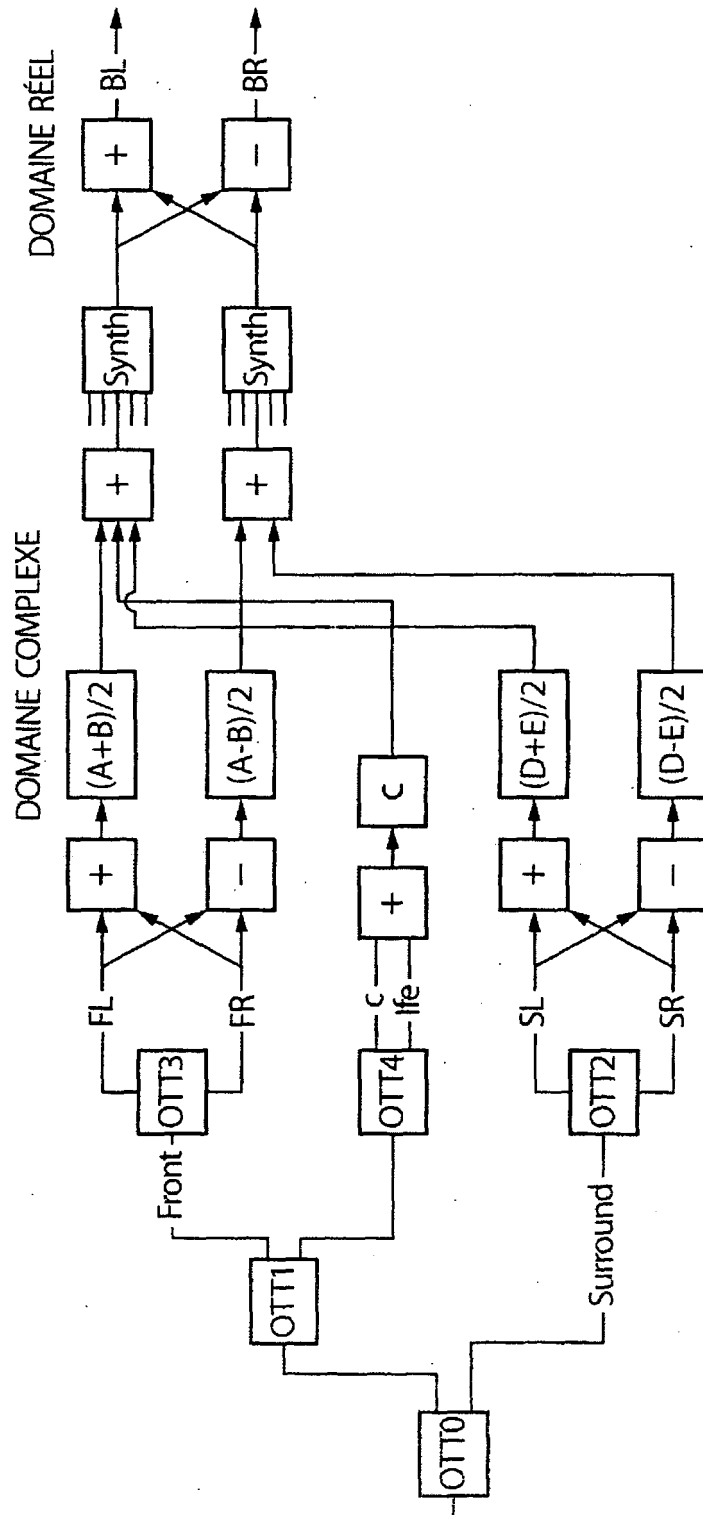


FIG. 1f
(ART ANTÉRIEUR)

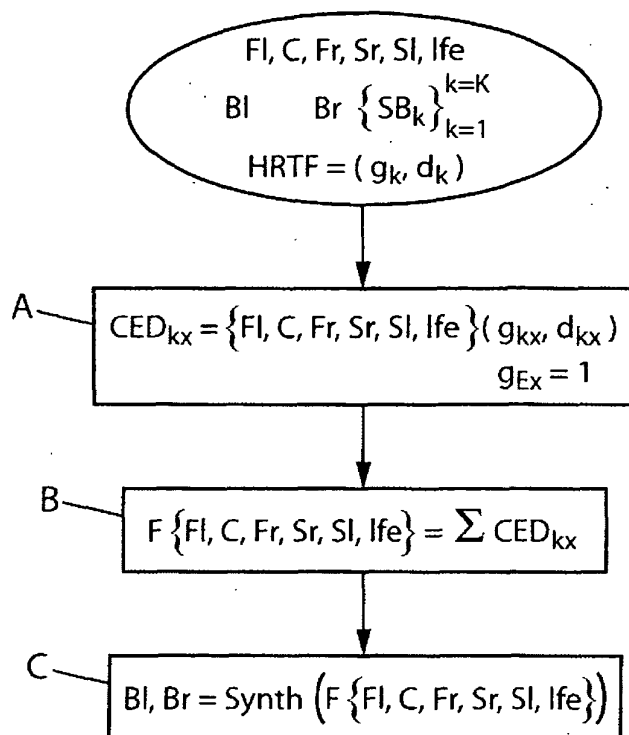


FIG. 2a

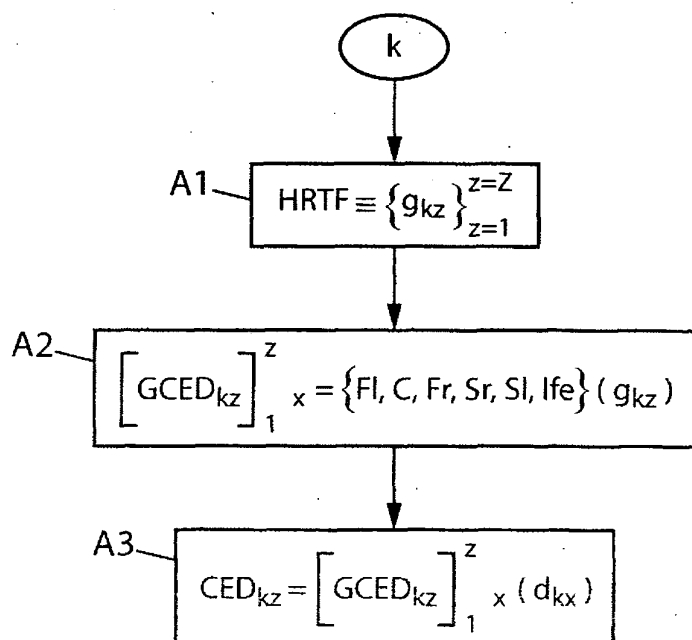


FIG. 2b

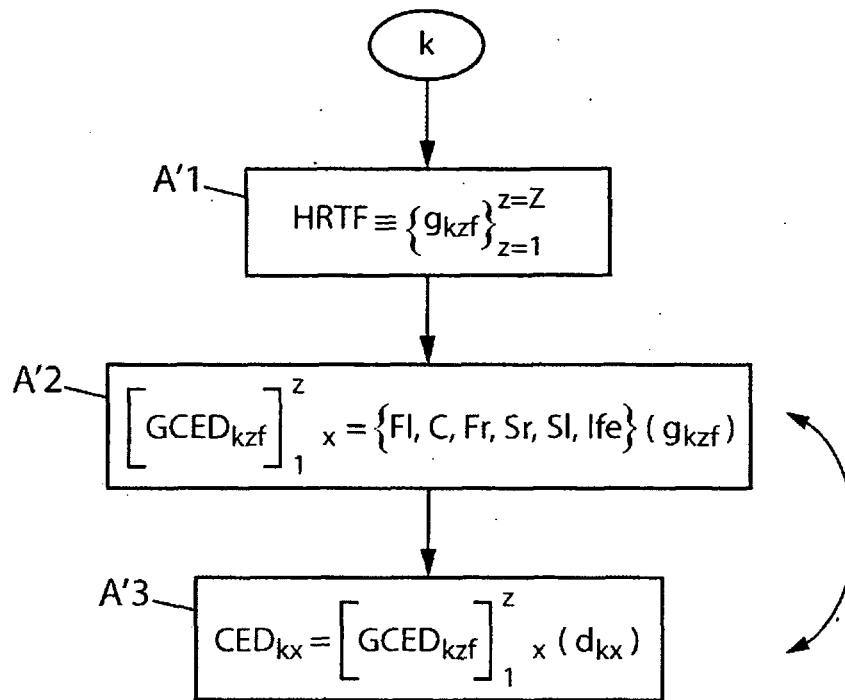


FIG. 2c

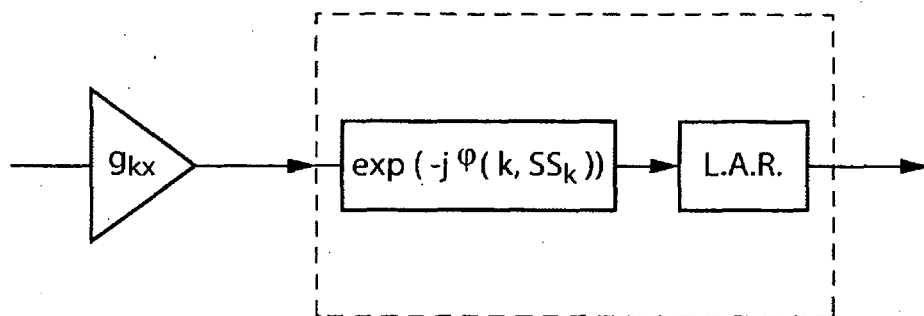


FIG. 3b

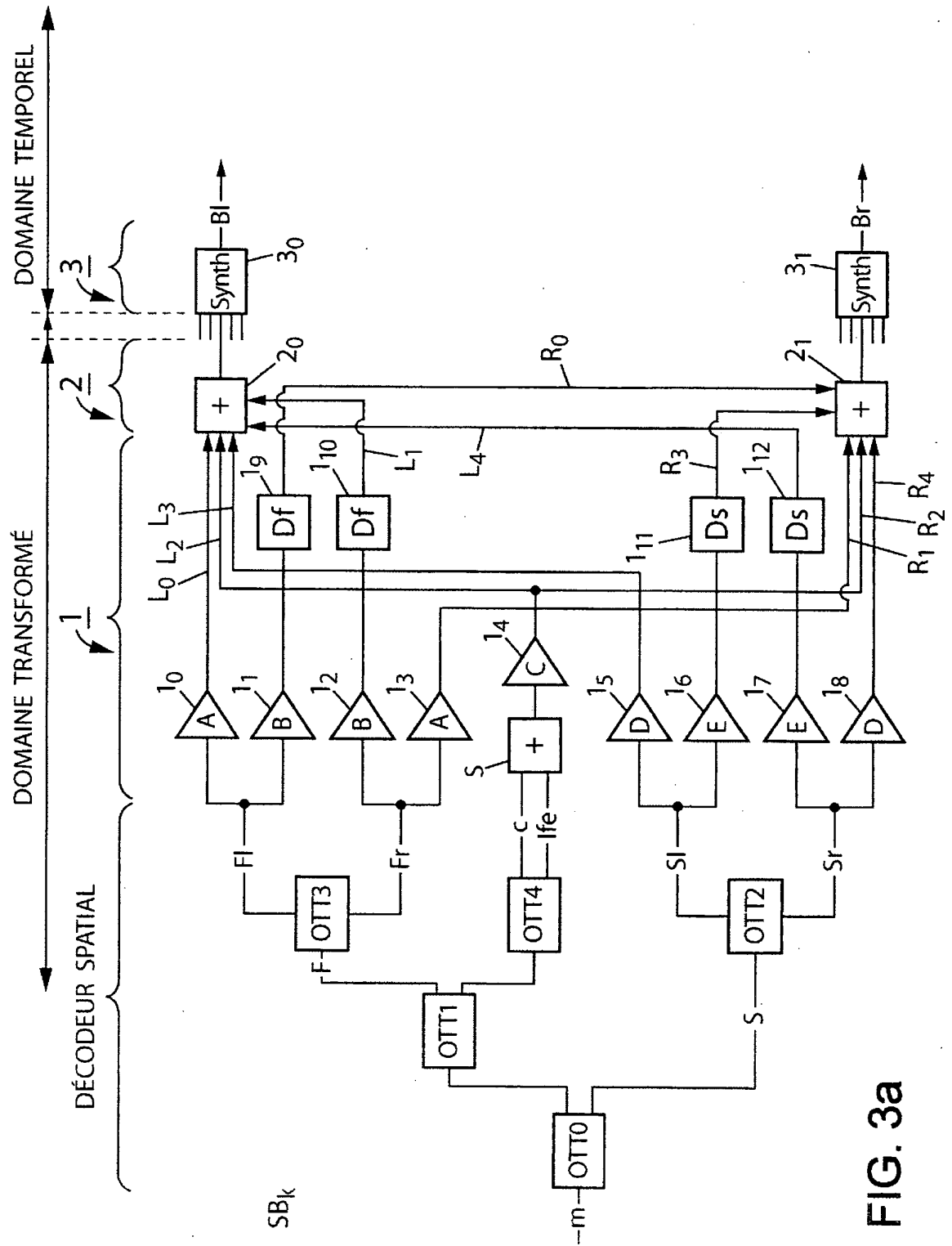


FIG. 3a

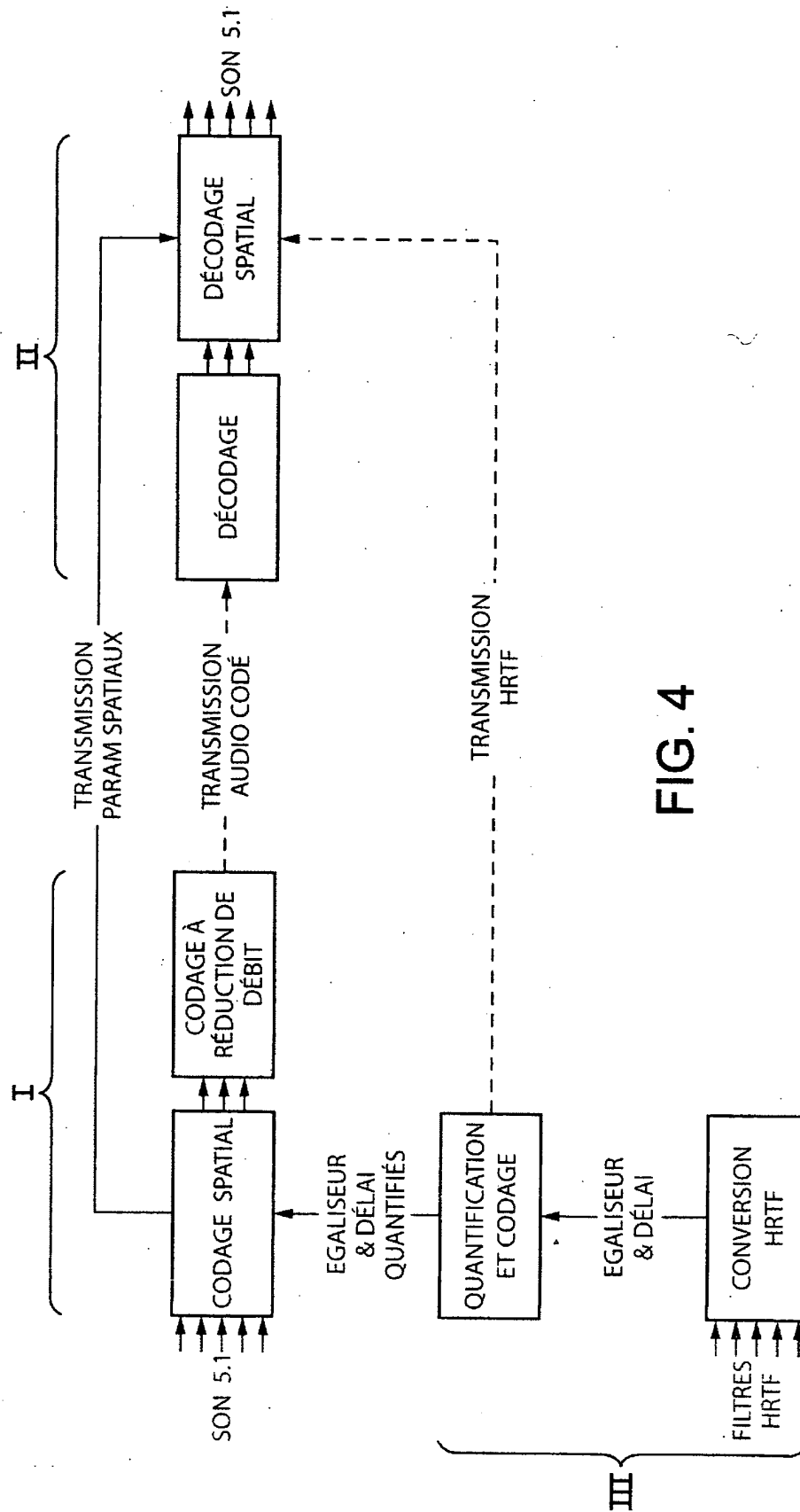


FIG. 4

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

- FR 2851879 A [0007]

Littérature non-brevet citée dans la description

- **D. Kistler ; F.L. Wightman.** A model of head-related transfer functions based on principal components analysis and minimum - phase reconstruction. *J. Acoust. Soc. Am.*, 1992, vol. 91 (3), 1637-1647 [0011]
- **A. Kulkarni.** IEEE ASSP Workshop on Applications of signal Processing to Audio and Acoustics. *IEEE catalog number : 95TH8144*, 1995 [0011]