

(19)



(11)

EP 2 042 001 B1

(12)

FASCICULE DE BREVET EUROPEEN

(45) Date de publication et mention
de la délivrance du brevet:
21.10.2009 Bulletin 2009/43

(51) Int Cl.:
H04S 1/00 (2006.01)

(21) Numéro de dépôt: **07803885.8**

(86) Numéro de dépôt international:
PCT/FR2007/051457

(22) Date de dépôt: **19.06.2007**

(87) Numéro de publication internationale:
WO 2008/003881 (10.01.2008 Gazette 2008/02)

(54) **SPATIALISATION BINAURALE DE DONNEES SONORES ENCODEES EN COMPRESSION**

BINAURALE SPATIALISIERUNG KOMPRESSIONSVERSCHLÜSSELTER TONDATEN

BINAURAL SPATIALIZATION OF COMPRESSION-ENCODED SOUND DATA

(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**

(30) Priorité: **07.07.2006 FR 0606212**

(43) Date de publication de la demande:
01.04.2009 Bulletin 2009/14

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

(72) Inventeurs:
• **VIRETTE, David**
22560 Pleumeur Bodou (FR)
• **GUERIN, Alexandre**
35000 Rennes (FR)

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

(56) Documents cités:
WO-A-20/05101370 WO-A2-20/04028204
US-A1- 2005 047 618

EP 2 042 001 B1

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

Description

[0001] L'invention concerne le traitement de données sonores, en vue d'une restitution spatialisée.

[0002] La spatialisation sonore tridimensionnelle (dite « rendu 3D ») de signaux audio compressés intervient notamment lors de la décompression d'un signal audio 3D, par exemple codé en compression et représenté sur un certain nombre de canaux, vers un nombre de canaux différents (deux par exemple pour permettre la restitution des effets audio 3D sur un casque d'écoute).

[0003] 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 (marque déposée), c'est-à-dire sur des haut-parleurs distants. De telles techniques peuvent alors utiliser une « annulation de diaphonie » (ou « cross-talk 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. Par la suite, on désignera communément ces deux techniques de restitution, binaurale et transaurale, par les mêmes termes « restitution binaurale ».

[0004] Ainsi, plus généralement, l'invention est relative à la transmission de signaux audio multicanaux et à leur conversion pour une restitution spatialisée (avec rendu 3D) sur deux voies. Le dispositif de restitution (simple casque à oreillettes par exemple) est le plus souvent imposé par l'équipement d'un utilisateur. La conversion peut viser par exemple le cas d'une restitution d'une scène sonore initialement au format multicanal 5.1 (ou 7.1, ou autre) par un simple casque d'écoute audio (en technique binaurale).

[0005] Bien entendu, l'invention concerne aussi la restitution, dans le cadre d'un jeu ou d'un enregistrement vidéo par exemple, d'un ou plusieurs échantillons sonores stockés dans des fichiers, en vue de leur spatialisation.

[0006] Concernant l'art antérieur, référence est faite au document US2005/047618, qui montre un procédé de traitement de données sonores pour une restitution spatialisée en trois dimensions sur deux voies de restitution pour l'oreille gauche et droite d'un auditeur utilisant une fonction de transfert.

[0007] Parmi les techniques connues dans le domaine de la spatialisation sonore binaurale, différentes approches ont été proposées.

[0008] En particulier, la synthèse binaurale bicanale consiste, en référence à la figure 1 relative à l'art antérieur, à :

- associer à chaque source sonore S_i (ou chaque canal du signal multicanal) une position dans l'espace,
- filtrer ces sources dans le domaine fréquentiel par les fonctions de transfert acoustiques gauche HRTF-l et droite HRTF-r correspondant à la direction (ou à la position) choisie, et définies par leurs coordonnées polaires (θ_1 , φ_1).

[0009] Ces fonctions de transfert, communément dites fonctions « HRTF » (pour « Head Related Transfer Functions » en anglais), représentent les fonctions de transfert acoustique entre les positions de l'espace et le conduit auditif de chaque oreille de l'auditeur. On désigne par « HRIR » (pour « Head Related Impulse Response » en anglais) leur forme temporelle ou réponse impulsionnelle. Ces fonctions HRIR peuvent inclure en outre un effet de salle.

[0010] On obtient, pour chaque source sonore S_i deux signaux (gauche et droit) qui sont alors additionnés aux signaux de gauche et de droite issus de la spatialisation de toutes les autres sources sonores, pour donner finalement les signaux L et R qui seront diffusés dans les oreilles gauche et droite de l'auditeur à travers deux haut-parleurs respectifs (oreillettes d'un casque en technique binaurale ou haut-parleurs distants en technique transaurale).

[0011] Si N désigne le nombre de sources sonores ou de flux audio incidents à spatialiser, le nombre de filtres, ou fonctions de transfert, nécessaire à la synthèse binaurale est $2 \times N$ pour un rendu en spatialisation binaurale statique et $4 \times N$ pour un rendu en spatialisation binaurale dynamique (avec transitions des fonctions de transfert).

[0012] Le traitement décrit ci-avant en référence à la figure 1 et mettant en jeu les fonctions de transfert HRTF est classique. Il est souvent utilisé pour un rendu 3D à partir de deux haut-parleurs. Il pourra être à la base d'une réalisation mise en oeuvre par la présente invention, comme on le verra plus loin. C'est à ce titre qu'il est introduit ici.

[0013] Néanmoins, l'invention part d'un autre type d'art antérieur.

[0014] Il existe des techniques de compression, souvent dans un domaine transformé, de signaux dans un format multicanal pour pouvoir véhiculer ces signaux, notamment à travers des réseaux de télécommunication, sur un nombre restreint de canaux, par exemple un ou deux canaux seulement. Ainsi, pour l'émission d'un signal dans un format multicanal comportant plus de deux canaux (par exemple 5.1, 7.1 ou autre), un codeur compresse le signal multicanal sur uniquement un ou deux canaux (typiquement selon le débit offert sur le réseau de télécommunication) et délivre en outre des informations de spatialisation. Cette réalisation est illustrée sur la figure 2A où, à titre d'exemple pour un signal dans un format multicanal 5.1, cinq canaux (C pour un haut-parleur central, FL pour un haut-parleur avant gauche, FR pour un haut-parleur avant droit, BL pour un haut-parleur arrière gauche et BR pour un haut-parleur arrière droit) sont encodés en compression par un module ENCOD propre à délivrer deux canaux compressés L et R, ainsi que des informations de spatialisation SPAT. Les canaux compressés L et R, ainsi que les informations de spatialisation SPAT

sont véhiculés ensuite à travers un ou plusieurs réseaux de télécommunication RES, sur un ou deux canaux selon le débit offert (figure 2B).

[0015] En référence à la figure 2C, à la réception du signal compressé sur les deux canaux L et R, un décodeur (DECOD) reconstitue le signal original dans le format multicanal initial grâce aux informations de spatialisation SPAT délivrées par le codeur et, dans l'exemple des figures 2A et 2C, on retrouve encore cinq canaux, après décodage, alimentant cinq haut-parleurs (HP-FL, HP-FR, HP-C, HP-BL et HP-BR) pour une restitution au format 5.1.

[0016] De nombreux types de codeurs/décodeurs paramétriques, notamment normalisés, offrent de telles possibilités.

[0017] Les codeurs audio (AAC, MP3) utilisent des représentations temps-fréquence des signaux pour compresser les informations. Ces représentations sont basées sur une analyse par bancs de filtres ou par transformation temps-fréquence de type MDCT (pour « Modified Discrete Cosine Transform »). Dans le cas où une spatialisation binaurale doit être effectuée après un décodage audio, les opérations de filtrage sont avantageusement réalisées d'emblée dans le domaine transformé.

[0018] Des travaux récents sur le filtrage dans le domaine transformé de sous-bandes ont permis de formaliser l'architecture de filtrage pour un banc de filtres couramment utilisé dans les codeurs audio. On pourra se référer utilement au document :

"A Generic Framework for Filtering in Subband Domain", A.Benjelloun Touimi, Proceeding IEEE - 9th Workshop on Digital Signal Processing, Hunt, Texas, USA, Octobre 2000.

[0019] Une technique plus récente de filtrage dans le domaine transformé des QMF complexes (pour « Quadrature Mirror Filters ») a été proposée dans la norme « MPEG Surround ». Cette technique vise la conversion de la réponse impulsionnelle (finie) du filtre temporel noté $h(v)$ en un ensemble de M filtres complexes notés $h_m(l)$, où M est le nombre de sous-bandes de fréquences. La conversion est réalisée par analyse du filtre temporel $h(v)$ par un banc de filtres complexes similaire au banc de filtres QMF utilisé pour l'analyse du signal. Dans un exemple de réalisation, le filtre prototype $q(v)$ utilisé pour générer le banc de filtre de conversion peut être de longueur 192. On définit une extension avec des zéros du filtre temporel par la formule suivante:

$$\tilde{h}(v) = \begin{cases} h(v), & v = 0, 1, \dots, N_h - 1; \\ 0, & \text{sinon,} \end{cases},$$

où :

- N_h est la longueur du filtre dans le domaine temporel,
- L_q/K_h+2 , avec $k_h = \lceil N_h/64 \rceil$, la longueur des filtres en sous-bandes (pour 64 sous-bandes).

[0020] La conversion est donc donnée par la formule suivante :

$$h_m(l) = \sum_{v=0}^{191} \tilde{h}(v + 64(l-2)) q(v) \exp\left(-j \frac{\pi}{64} \left(m + \frac{1}{2}\right) (v - 95)\right),$$

avec :

- $m = 0, 1, \dots, 63$, correspondant à l'indice de la sous-bande
- $l = 0, 1, \dots, K_h+1$, correspondant à l'indice temporel dans le domaine décimé des sous-bandes.

[0021] En termes plus génériques, on comprendra qu'un tel traitement, directement dans le domaine transformé, permet de passer d'une représentation du signal compressé sur deux canaux L, R vers une représentation du signal sur deux voies de restitution L-BIN, R-BIN (figure 3) avec un élargissement binaural ou transaural. A cet effet, on prévoit un transcodage (module DECOD BIN de la figure 3) qui se base sur une approche consistant à reconstituer, à partir des signaux compressés L, R et des informations de spatialisation SPAT, les fonctions de transfert, de type HRTF, entre une oreille d'un auditeur et chaque haut-parleur (virtuel) qui aurait été alimenté par un canal donné du format multicanal initial.

[0022] Ainsi, en référence maintenant à la figure 4 illustrant une restitution « virtuelle » au format 5.1, donc à partir de cinq haut-parleurs, le transcodage mis en oeuvre par le module DECOD BIN de la figure 3 doit considérer dix fonctions de transfert :

- une pour le chemin A entre le haut-parleur avant gauche HP-FL et l'oreille gauche OL de l'auditeur AU,
- une pour le chemin B entre le haut-parleur avant gauche HP-FL et l'oreille droite OR de l'auditeur AU,
- une pour le chemin C entre le haut-parleur arrière gauche HP-BL et l'oreille gauche OL de l'auditeur AU,
- une pour le chemin D entre le haut-parleur arrière gauche HP-BL et l'oreille droite OR de l'auditeur AU,
- une pour le chemin G entre le haut-parleur avant droit HP-FR et l'oreille droite OR de l'auditeur AU,
- une pour le chemin H entre le haut-parleur avant droit HP-FR et l'oreille gauche OL de l'auditeur AU,
- une pour le chemin F entre le haut-parleur arrière droit HP-BR et l'oreille droite OR de l'auditeur AU,
- une pour le chemin E entre le haut-parleur arrière droit HP-BR et l'oreille gauche OL de l'auditeur AU,
- une pour le chemin J entre le haut-parleur du centre HP-C et l'oreille gauche OL de l'auditeur AU, et
- une pour le chemin I entre le haut-parleur du centre HP-C et l'oreille droite OR.

[0023] Ainsi, les filtres en sous-bandes dans le domaine transformé sont calculés pour chaque oreille et pour chacune des cinq positions des haut-parleurs. Cette technique est souvent appelée « technique des haut-parleurs virtuels ».

[0024] A l'aide de la représentation en sous-bandes des filtres binauraux déterminés comme décrit ci-avant à partir des fonctions de transfert HRTF, la spatialisation binaurale peut alors être avantageusement réalisée en appliquant ces filtres binauraux directement dans le domaine transformé au coeur du décodeur audio DECOD BIN comme représenté sur la figure 3.

[0025] Ainsi, ce type de décodeur DECOD BIN utilise une représentation monophonique ou stéréophonique (voies compressées L, R) de la scène audio multicanal, représentation à laquelle sont associés des paramètres de spatialisation SPAT (qui peuvent consister par exemple en des différences d'énergies entre canaux et indices de corrélation entre canaux). Ces paramètres SPAT sont utilisés au décodage pour reproduire au mieux la scène sonore multicanal originale.

[0026] En outre, lorsque le signal original est encodé par un codeur paramétrique (par exemple au sens de travaux récents dans la norme « MPEG Surround »), en plus du signal monophonique ou stéréophonique transmis et des informations de spatialisation, le décodage peut utiliser des représentations décorrélées de ces signaux L, R (qui sont obtenues, par exemple, par l'application de filtres, de décorrélation passe-tout ou de filtres de réverbération). Ces signaux sont ensuite ajustés en énergie grâce aux différences d'énergie intercanal, puis recombinaison pour obtenir le signal multicanal en vue de la restitution.

[0027] En particulier, l'encodeur paramétrique (ENCOD - figure 2A) du format multicanal vers deux voies compressées (stéréo ou mono) selon le projet de norme « MPEG Surround » délivre une information de décorrélation entre canaux dans le format multicanal initial et cette information de décorrélation peut être reprise par le décodeur paramétrique homologue (DECOD - figure 2C) lors de la restitution au format multicanal initial.

[0028] Une description de travaux préparatoires à cette norme est donnée à l'adresse URL :

« <http://www.chiariglione.org/mpeg/technologies/mpd-mps/index.htm> »

et des précisions quant à un tel codeur selon ce projet de norme peuvent être trouvées dans :

"MPEG Spatial Audio Coding / MPEG Surround: Overview and Current Status", J.Breebaart et al., dans 119th Conv. Aud. Eng. Soc (AES), New York, NY, USA, Octobre 2005.

[0029] Dans le cas d'un décodeur audio paramétrique pour une restitution binaurale (DECOD BIN - figure 3), il est avantageusement possible de simplifier les opérations de filtrage en combinant les filtres avant et arrière correspondant aux différents haut-parleurs gauches (un traitement équivalent étant porté aussi pour les haut-parleurs droits). Cette combinaison est réalisée en fonction des énergies cibles des canaux audio données par les paramètres de spatialisation. Cette combinaison, pour l'oreille gauche et les canaux avant gauche et arrière gauche, est réalisée dans le domaine transformé selon une expression (1) du type :

$$h_{L,L} = g_{L,L} \sigma_{FL} \exp(-j\phi_{FL,BL}^L \sigma_{BL}^2) h_{L,FL} + g_{L,L} \sigma_{BL} \exp(j\phi_{FL,BL}^L \sigma_{FL}^2) h_{L,BL}$$

[0030] Dans cette expression :

- $h_{L,L}$ est le filtre correspondant aux contributions des canaux avant et arrière gauche,
- $g_{L,L}$ est le gain associé à l'ensemble des canaux gauches,
- σ_{FL}^2 et σ_{BL}^2 sont les énergies utiles des canaux respectivement avant et arrière gauches,
- $h_{L,FL}$ et $h_{L,BL}$ sont les fonctions de transfert dans le domaine des sous-bandes entre l'oreille gauche et les haut-parleurs respectivement avant et arrière gauche (chemins A et C de la figure 4),
- $\Phi_{FL,BL}^L$ est le déphasage correspondant au retard entre les filtres temporels avant et arrière gauche $h_{L,FL}$ et $h_{L,BL}$.

[0031] Le but de cette compensation de phase, fonction de l'énergie cible des canaux, est d'éviter un effet dit « de coloration » issu de l'addition de deux filtres décalés dans le temps (filtrage en peigne).

[0032] En référence à la figure 5A, un décodeur reçoit les paramètres de spatialisation SPAT accompagnant les signaux compressés sur deux voies L et R dans l'exemple représenté et on a illustré, sur cette même figure 5A, comment le filtre précité $h_{L,L}$ s'applique au canal compressé L pour former une composante du signal L-BIN, destiné à la restitution binaurale. Néanmoins, comme représenté encore sur la figure 5A, il convient de tenir compte aussi du signal compressé sur le canal R qui doit, quant à lui, être filtré par un filtre faisant intervenir des fonctions de transfert HRTF (notées $H_{L,FR}$ et $H_{L,BR}$) relatives aux chemins croisés H et E de la figure 4, toujours vers l'oreille gauche. Le filtre correspondant à ces chemins croisés (noté $h_{L,R}$) se calcule en fonction des gains, énergies cibles et déphasages, tirés des paramètres de spatialisation SPAT, à l'aide d'une expression équivalente à la relation (1) donnée précédemment. Ce filtre $h_{L,R}$ est finalement appliqué au signal compressé sur le canal R. Il convient de tenir compte en outre de la « contribution » du haut-parleur central dans la construction du signal destiné à la restitution binaurale L-BIN et, pour ce faire, on applique un filtre $h_{L,C}$ (figure 5A) à une combinaison (par exemple par addition) des signaux compressés des deux voies L et R pour tenir compte ici du chemin J vers l'oreille gauche OL de la figure 4.

[0033] En référence encore à la figure 5A, on prévoit un traitement équivalent pour la construction du signal R-BIN destiné à une restitution binaurale pour l'oreille droite OD, avec trois contributions données par :

- le signal compressé sur le canal R filtré par le filtre $h_{R,R}$ représentant les fonctions HRTF des haut-parleurs de droite (chemins directs G et F de la figure 4) ;
- le signal compressé sur le canal L filtré par le filtre $h_{R,L}$ représentant les fonctions HRTF des haut-parleurs de gauche (chemins croisés B et D de la figure 4) ; et
- une combinaison des signaux compressés L et R filtrée par le filtre $h_{R,C}$ représentant les fonctions HRTF du haut-parleur de centre (chemin direct I de la figure 4).

[0034] Sur la figure 5B, on a représenté un autre exemple, dans lequel un décodeur reçoit le signal compressé sur un seul canal M, accompagnant les paramètres de spatialisation SPAT. Dans l'exemple représenté, on duplique le canal M en deux canaux L et R et la suite du traitement est strictement équivalente au traitement représenté sur la figure 5A.

[0035] Les deux signaux L-BIN et R-BIN résultant de ces filtrages peuvent être appliqués ensuite à deux haut-parleurs destinés respectivement à l'oreille gauche et à l'oreille droite de l'auditeur après un passage du domaine transformé au domaine temporel.

[0036] Toutefois, un problème lié à cette combinaison de filtres pour une restitution binaurale est qu'elle ne prend pas en compte une éventuelle décorrélation entre les canaux avant et arrière. Cette information, pourtant utilisée dans le décodage d'une scène 5.1 d'un encodeur selon le projet précité de la norme MPEG Surround, n'est pas exploitée dans la technique de décodage binaural. Ainsi, lorsque la scène sonore comporte des effets de décorrélation entre les canaux avant et arrière (par exemple pour des signaux réverbérés), cette information n'est pas utilisée dans la combinaison des filtres HRTF, ce qui entraîne une dégradation de la qualité de spatialisation et notamment de l'effet d'enveloppement de la scène audio 3D. Ainsi, la restitution au format binaural n'est pas optimale.

[0037] La présente invention vient améliorer la situation.

[0038] Elle vise tout d'abord un procédé de traitement de données sonores pour une restitution spatialisée en trois dimensions sur deux voies de restitution pour les oreilles respectives d'un auditeur, les données sonores étant initialement représentées dans un format multicanal puis encodées en compression sur un nombre réduit de canaux (par exemple un ou deux canaux),

ledit format multicanal initial consistant à prévoir plus de deux canaux susceptibles d'alimenter des haut-parleurs respectifs,

le procédé comportant les étapes :

- obtenir, avec les données compressées sur ledit nombre réduit de canaux, des paramètres de spatialisation,
- pour chaque voie de restitution associée à une oreille de l'auditeur, former, à partir desdits paramètres de spatialisation, une combinaison de filtres représentatifs chacun de fonctions de transfert entre cette oreille de l'auditeur et des haut-parleurs susceptibles d'être alimentés par des canaux respectifs du format multicanal initial, et
- appliquer aux données compressées la combinaison de filtres associée à chaque voie de restitution.

[0039] Le procédé au sens de l'invention comporte en outre les étapes :

- pour chaque voie de restitution associée à une oreille de l'auditeur, déterminer à partir desdits paramètres de spatialisation au moins une fonction de transfert d'un haut-parleur situé à l'arrière de l'oreille de l'auditeur et représentative d'une décorrélation entre les canaux du format multicanal respectivement associés au haut-parleur arrière et à au moins un haut-parleur situé devant l'oreille de l'auditeur, et
- pour chaque voie de restitution, intégrer ladite fonction de transfert représentative d'une décorrélation dans ladite combinaison de filtres associée à cette voie de restitution.

[0040] La restitution spatialisée sur deux voies, au sens de l'invention, peut être aussi bien au format binaural ou transaural. Le format multicanal, initial peut être de type ambiophonique (ou « ambisonic » en anglais et visant la décomposition du signal sonore dans une base d'harmoniques sphérique). En variante, il peut s'agir d'un format de type 5.1 ou 7.1, ou encore 10.2. On comprendra alors que pour ces derniers types de format mettant en oeuvre des canaux destinés à alimenter respectivement au moins des couples de haut-parleurs avant-gauche/arrière-gauche, d'une part, et avant-droit/arrière-droit, d'autre part, l'information de décorrélation peut viser les canaux respectifs des haut-parleurs avant/arrière préférentiellement associés à une même oreille (gauche ou droite).

[0041] Selon un avantage que procure l'invention, du fait que cette information de décorrélation à l'arrière d'une scène 3D soit représentée dans la restitution binaurale ou transaurale, on obtient une meilleure représentation des ambiances, par exemple des bruits de foule ou une réverbération à l'arrière d'une scène, ou autre, contrairement aux réalisations de l'art antérieur.

[0042] Dans un mode de réalisation particulier, la combinaison de filtres comporte une pondération, selon un coefficient choisi, entre :

- une fonction de transfert brute du haut-parleur situé à l'arrière, et
- une version de la fonction de transfert de ce haut-parleur, représentative de la décorrélation.

[0043] Cette pondération permet avantageusement de favoriser la fonction de transfert brute de ce haut-parleur arrière, ou la version décorrélée de cette fonction de transfert brute, selon que le signal dans le canal arrière du format multicanal initial est corrélé ou non avec au moins un signal de l'un des canaux avant.

[0044] Par ailleurs, dans une réalisation particulière, la combinaison de filtres associée à une voie de restitution comporte au moins un regroupement formant filtre à partir de :

- la fonction de transfert d'un haut-parleur avant,
- la fonction de transfert d'un haut-parleur arrière, et
- la fonction de transfert représentative d'une décorrélation entre canaux,

et ces haut-parleurs avant et arrière sont situés sur un même côté par rapport à l'auditeur. Il peut s'agir par exemple des haut-parleurs avant et arrière situés tous les deux à gauche (ou tous les deux à droite) de l'auditeur au format 5.1 (comme illustré sur la figure 4). Dans une telle réalisation, lorsque la pondération entre la version décorrélée et la version brute des fonctions de transfert est prévue, il pourra être avantageux de privilégier la version décorrélée dans la combinaison de filtres des haut-parleurs de gauche pour la voie de restitution sur l'oreille droite (et inversement) et privilégier la version brute (non décorrélée) dans la combinaison de filtres des haut-parleurs de droite (gauche) pour la voie de restitution sur l'oreille droite (gauche).

[0045] Avantageusement, l'encodage en compression met en oeuvre un codeur paramétrique délivrant, dans le flux compressé incluant les paramètres de spatialisation, une information de décorrélation entre canaux du format multicanal, à partir de laquelle peut être déterminée, de façon dynamique, la pondération précitée.

[0046] Ainsi, dans cette réalisation, pour un transcodage entre un format multicanal vers un format binaural, la combinaison précitée de fonctions de transfert tire parti des informations déjà présentes concernant la corrélation entre signaux de canaux au format multicanal, ces informations étant simplement fournies par le codeur paramétrique, avec les paramètres de spatialisation précités.

[0047] A titre d'exemple, on rappelle que le décodeur paramétrique selon le projet de norme MPEG Surround délivre une telle information de décorrélation entre canaux au format multicanal 5.1.

[0048] D'autres avantages et caractéristiques de l'invention apparaîtront à la lecture de la description détaillée donnée ci-après à titre d'exemple, et à l'observation des dessins annexés, dans lesquels, outre les figures 1, 2A, 2B, 2C, 3 et 4, 5A et 5B commentées ci-avant :

- les figures 6A et 6B illustrent à titre d'exemple un traitement par filtrage de données compressées (sur deux canaux dans l'exemple représenté), le filtrage étant déterminé par la mise en oeuvre du procédé au sens de l'invention pour

délivrer des signaux L-BIN et R-BIN destinés à alimenter respectivement les voies gauche et droite d'un dispositif de restitution binaurale tel qu'un casque à deux oreillettes, en tenant compte d'une décorrélation avant/arrière, et

- la figure 7 illustre schématiquement la structure d'un module mettant en oeuvre le procédé au sens de l'invention.

[0049] En référence à la figure 6A, on récupère tout d'abord le signal compressé, souvent dans le domaine transformé, sur deux canaux L et R dans l'exemple représenté, ainsi que les paramètres de spatialisation SPAT qu'a fourni un encodeur tel que le module ENCOD de la figure 2A décrite précédemment. A partir des paramètres de spatialisation SPAT, on détermine des fonctions de transfert pour construire une combinaison de filtres (signe « + » de la figure 6A), chaque filtre étant à appliquer à un canal L (filtre $h_{L,L}$ de la figure 5A), ou R (filtre $h_{L,R}$ de la figure 5A), ou à une combinaison de ces canaux (filtre $h_{L,C}$ de la figure 5A) pour construire un signal alimentant l'une des deux voies de restitution binaurale L-BIN. Ces fonctions de transfert, de type HRTF, sont représentatives des perturbations que subit une onde acoustique sur un chemin entre un haut-parleur qui aurait été alimenté par un canal du format multicanal initial et une oreille de l'auditeur. Par exemple si le contenu audio est initialement au format 5.1, comme décrit ci-avant en référence à la figure 4, on détermine en tout dix fonctions de transfert HRTF, cinq fonctions HRTF pour l'oreille droite (sur les chemins B, D, G, F et I de la figure 4) et cinq fonctions HRTF pour l'oreille gauche (sur les chemins A, C, H, E et J). On indique que le haut-parleur du centre est traité séparément dans la spatialisation binaurale et l'obtention du filtre correspondant $h_{L,C}$ ou hic ne sera pas décrite ici, étant entendu qu'elle ne fait pas intervenir, a priori dans l'exemple décrit, l'objet de l'invention.

[0050] Ainsi, en termes généraux, les fonctions HRTF de haut-parleurs avant et arrière, situés sur un même côté de l'auditeur sont donc regroupées pour construire chaque filtre d'une combinaison de filtres propre à une voie de restitution sur une oreille d'un auditeur. Un regroupement de fonctions HRTF pour construire un filtre est par exemple une addition, moyennant des coefficients multiplicatifs dont un exemple sera décrit plus loin.

[0051] Au sens de l'invention, à partir des paramètres de spatialisation SPAT récupérés, on détermine en outre une version décorrélée des fonctions HRTF des haut-parleurs situés à l'arrière de l'auditeur (chemins C, D, E et F de la figure 4) et on intègre cette version décorrélée à chaque regroupement pour former un filtre à appliquer à un canal compressé.

[0052] A titre d'exemple purement illustratif, les données sonores initiales peuvent être au format multicanal 5.1 et, en référence à la figure 6A, un premier regroupement comporte :

- la fonction HRTF-A (pour le haut-parleur avant gauche selon un chemin direct vers l'oreille gauche OL de la figure 4),
- la fonction HRTF-C (pour le haut-parleur arrière gauche selon un chemin direct vers l'oreille gauche),
- et la version décorrélée de cette fonction HRTF-C, notée HRTF-C*, pour former le filtre à appliquer au canal compressé L.

Un second regroupement comporte :

- la fonction HRTF-H (pour le haut-parleur avant droit selon un chemin croisé vers l'oreille gauche),
- la fonction HRTF-E (pour le haut-parleur arrière droit selon un chemin croisé),
- et la version décorrélée de cette fonction HRTF-E, notée HRTF-E*, pour former le filtre à appliquer au canal compressé R.

L'addition des deux signaux résultant de tels filtrages sera une composante du signal alimentant la voie de restitution binaurale L-BIN associée à l'oreille gauche.

[0053] Un traitement similaire est prévu pour construire le signal destiné à alimenter l'autre voie de restitution binaurale R-BIN de la figure 6B. Ici, on tient compte des fonctions HRTF des chemins menant à l'oreille droite OD de l'auditeur AU (figure 4). Un premier regroupement comporte les fonctions HRTF-G (pour le haut-parleur avant droit selon un chemin direct), HRTF-F (pour le haut-parleur arrière droit selon un chemin direct) et la version décorrélée notée HRTF-F* de la fonction HRTF-F pour former le filtre à appliquer au canal compressé R. Un second regroupement comporte la fonction HRTF-B (pour le haut-parleur avant gauche selon un chemin croisé), la fonction HRTF-D (pour le haut-parleur arrière gauche selon un chemin croisé) et la version décorrélée, notée HRTF-D*, de la fonction HRTF-D, pour former le filtre à appliquer au canal compressé L.

[0054] Finalement, les combinaisons de filtres intégrant les versions décorrélées des fonctions HRTF des haut-parleurs arrière sont appliquées aux canaux compressés L et R pour délivrer les voies de restitution L-BIN et R-BIN, pour une restitution binaurale spatialisée avec rendu 3D.

[0055] Dans l'exemple représenté sur les figures 6A et 6B, les données sonores reçues sont encodées en compression sur deux canaux stéréophoniques L et R comme illustré dans l'exemple de la figure 5A. En variante, elles pourraient être encodées en compression sur un canal unique M, monophonique, comme illustré sur la figure 5B, auquel cas les combinaisons de filtres sont appliquées au canal monophonique (dupliqué) comme illustré sur la figure 5B, pour délivrer

encore deux signaux alimentant respectivement les deux voies de restitution L-BIN, R-BIN.

[0056] Dans une réalisation avantageuse, les données sonores initiales sont au format multicanal 5.1 et sont encodées en compression par un codeur paramétrique selon le projet de norme précité MPEG Surround. Plus particulièrement, lors d'un tel encodage, il est possible d'obtenir, parmi les paramètres de spatialisation fournis, une information de

décorrélacion entre le canal arrière droit et le canal avant droit (haut-parleurs respectifs HP-BR et HP-FR de la figure 4), ainsi qu'une information de décorrélacion homologue entre le canal arrière gauche et le canal avant gauche (haut-parleurs respectifs HP-FR et HP-BR de la figure 4).

[0057] Ces informations de décorrélacion, dans un format 5.1, visent à rendre la restitution des haut-parleurs arrière la plus indépendante possible de la restitution des haut-parleurs avant, pour enrichir, au format 5.1, l'effet d'enveloppement par des bruits de réverbération ou de public pour des enregistrements de concerts par exemple. On rappelle que cet enrichissement de l'enveloppement 3D n'a pas été proposé en restitution binaurale et un avantage de l'invention est de tirer profit de la disponibilité des informations de décorrélacion parmi les paramètres de spatialisation SPAT pour construire des versions décorrélées des fonctions HRTF qui s'intègrent avantageusement aux combinaisons de filtres pour une restitution binaurale.

[0058] Selon un autre avantage, ces combinaisons de filtres peuvent être calculées directement dans le domaine transformé, par exemple dans le domaine des sous-bandes, et les filtres représentant les versions décorrélées des fonctions HRTF des haut-parleurs arrière peuvent être obtenues par exemple en appliquant aux fonctions HRTF initiales un déphasage fonction de la sous-bande de fréquences considérée.

[0059] Plus généralement, les filtres de décorrélacion peuvent être des filtres de réverbération dite « naturelle » (enregistrée dans un environnement acoustique particulier comme une salle de concert par exemple), ou « synthétique » (créée par sommation de réflexions multiples d'amplitudes décroissantes dans le temps). L'application d'un filtre décorrélé peut donc revenir à appliquer au signal décomposé en sous-bandes de fréquences un déphasage différent dans chacune des sous-bandes, combiné avec l'ajout d'un délai global. Dans le cas d'un décodeur paramétrique du type précité (formule (1) donnée précédemment en description de l'art antérieur), cela revient à multiplier chaque sous-bande de fréquences par une exponentielle complexe, de phase différente dans chaque sous-bande. Ces filtres de décorrélacion peuvent donc correspondre à des synthèses de filtres passe-tout déphaseur.

[0060] On prévoit avantageusement une pondération entre la fonction de transfert d'un haut-parleur arrière et sa version décorrélée dans un même regroupement formant un filtre. Ainsi, en reprenant la formule (1) donnée précédemment pour le calcul d'un filtre par exemple $h_{L,L}$ pour l'oreille gauche, on introduit des coefficients de pondération α et $(1-\alpha)$ et la version décorrélée d'une fonction de transfert comme suit :

$$h_{L,L} = g_{L,L} \sigma_{FL} \exp(-j\phi_{FL,BL}^L \sigma_{BL}^2) h_{L,FL} + g_{L,L} \sigma_{BL} \exp(j\phi_{FL,BL}^L \sigma_{FL}^2) (\alpha h_{L,BL} + (1-\alpha) h_{L,BL}^{Decorr})$$

avec les mêmes notations explicitées précédemment et où $h_{L,BL}^{Decorr}$ représente la version décorrélée de la fonction de transfert du haut-parleur arrière gauche. Bien entendu, on prévoit un même type de relations donnant les autres filtres $h_{L,R}$, $h_{R,R}$ et $h_{R,L}$ (figures 5A et 5B).

Par exemple pour le filtre $h_{L,R}$ propre aux chemins croisés vers l'oreille gauche, on a :

$$h_{L,R} = g_{L,R} \sigma_{FR} \exp(-j\phi_{FR,BR}^L \sigma_{BR}^2) h_{L,FR} + g_{L,R} \sigma_{BR} \exp(j\phi_{FR,BR}^L \sigma_{FR}^2) (\alpha h_{L,BR} + (1-\alpha) h_{L,BR}^{Decorr})$$

[0061] Plus spécifiquement, on prévoit une pondération par des coefficients différents α_1 , $(1-\alpha_1)$ et α_2 , $(1-\alpha_2)$ selon le fait que le haut-parleur arrière est du même côté que l'oreille considérée ($\alpha=\alpha_1$ donnant les filtres $h_{L,L}$ et $h_{R,R}$) ou non ($\alpha=\alpha_2$ donnant les filtres $h_{L,R}$ et $h_{R,L}$). Préférentiellement, on privilégie la version décorrélée pour des chemins croisés (haut-parleur arrière droit pour l'oreille gauche et haut-parleur arrière gauche pour l'oreille droite), de sorte qu'en général, le coefficient α_1 , pourra être souvent supérieur au coefficient α_2 .

[0062] En pratique, les coefficients α (α_1 ou α_2) sont donnés par des fonctions de pondération variables de manière à favoriser dynamiquement la version brute de la fonction HRTF du haut-parleur arrière ou sa version décorrélée selon que le signal arrière est corrélé ou non avec le signal avant. On obtient ainsi une meilleure représentation des ambiances (bruits de foule, réverbération, ou autre) dans le rendu 3D.

La fonction de pondération α peut être définie dynamiquement grâce à l'information de décorrélacion fournie avec les paramètres de spatialisation, à titre d'exemple non limitatif, de la façon suivante :

$$\alpha = \text{sqrt}(\text{abs}(\text{ICC}_L)), \text{ si } \text{abs}(\text{ICC}_L) > \sigma_{BL}^2$$

5

$$\alpha = \text{sqrt}(\sigma_{BL}^2), \text{ sinon,}$$

10

15

où la notation « sqrt » vise la fonction « racine carrée », la notation « abs » vise la fonction « valeur absolue » et le terme ICC_L représente l'information de décorrélation (appelée autrement « indice de corrélation ») entre le canal avant et le canal arrière du même côté gauche et fait partie des paramètres de spatialisation transmis par l'encodeur selon le projet de norme MPEG Surround indiqué ci-avant. Comme décrit ci-avant, le terme σ_{BL} représente l'énergie cible du canal arrière gauche lorsqu'il s'agit de déterminer le coefficient α pour calculer le filtre $h_{L,L}$ ($\alpha = \alpha_1$). Bien entendu, une expression équivalente peut être appliquée pour calculer le coefficient de pondération α intervenant dans le filtre homologue $h_{R,R}$ propre aux chemins acoustiques directs vers l'oreille droite. Néanmoins, pour les filtres $h_{L,R}$ et $h_{R,L}$ propres aux chemins croisés, par exemple pour le filtre $h_{L,R}$ propre aux chemins croisés vers l'oreille gauche, le coefficient $\alpha = \alpha_2$ peut préférentiellement s'écrire :

20

$$\alpha_2 = \text{abs}(\text{ICC}_R), \text{ si } \text{abs}(\text{ICC}_R) > \sigma_{BR}^2,$$

25

$$\alpha_2 = \sigma_{BR}^2, \text{ sinon,}$$

30

le terme σ_{BR} représentant l'énergie cible du canal arrière droit et le terme ICC_R représentant l'indice de corrélation entre le canal avant droit et le canal arrière droit.

On relèvera que la fonction « sqrt » ne s'applique plus pour les chemins croisés et pour le calcul du coefficient correspondant α_2 , dans l'exemple décrit. En effet, les énergies cibles et les indices de corrélation sont des termes compris entre 0 et 1 de sorte que le coefficient α_2 est généralement inférieur au coefficient α_1 .

35

[0063] La combinaison de filtres globale, pour la voie L-BIN, comporte bien des regroupements de fonctions HRTF formant des filtres $h_{L,L}$ et $h_{L,R}$ obtenus par les formules données précédemment, et, dans chaque regroupement, interviennent bien la fonction HRTF d'un haut-parleur avant, la fonction HRTF d'un haut-parleur arrière et une version décorrélée de cette dernière fonction HRTF, ce qui permet de représenter une décorrélation entre les canaux avant et arrière directement dans la combinaison de filtres, et donc directement dans la synthèse binaurale.

40

[0064] On rappelle que, les données sonores L, R (ou M) étant encodées en compression dans un domaine transformé, la combinaison de filtres peut être appliquée directement dans le domaine transformé en fonction des énergies cibles (σ_{FL} , σ_{BL} , σ_{FR} , σ_{BR}) associées aux canaux du format multicanal, ces énergies cibles étant déterminées à partir des paramètres de spatialisation SPAT. Dans cette réalisation, bien entendu, on prévoit ensuite le passage à nouveau du domaine transformé au domaine temporel pour la restitution proprement dite en contexte binaural (modules TRANS des figures 6A et 6B).

45

[0065] La présente invention vise aussi un module de décodage DECOD BIN tel que représenté à titre d'exemple sur la figure 7, pour une restitution spatialisée en trois dimensions sur deux voies de restitution L-BIN et R-BIN, et comportant en particulier des moyens de traitement de données sonores (canaux compressés L, éventuellement R en mode stéréophonique et les paramètres de spatialisation SPAT) pour la mise en oeuvre du procédé décrit ci-avant. Ces moyens peuvent typiquement comporter :

50

- une entrée E pour recevoir les canaux compressés et les paramètres de spatialisation,
- une mémoire de travail MEM et un processeur PROC pour construire les combinaisons de filtres à partir des paramètres SPAT et appliquer ces combinaisons respectivement aux canaux compressés L et R,
- et une sortie S pour délivrer les signaux compressés et filtrés pour une restitution binaurale spatialisée sur les voies de restitution respectives L-BIN et R-BIN.

55

[0066] La présente invention vise aussi un programme informatique, destiné à être stocké dans une mémoire d'un module de décodage, telle que la mémoire MEM du module DECOD-BIN de la figure 7, pour une restitution spatialisée en trois dimensions sur deux voies de restitution L-BIN et R-BIN. Le programme comporte alors des instructions pour l'exécution du procédé selon l'invention et, en particulier, pour construire les combinaisons de filtres intégrant les versions

décorrélées comme illustré sur les figures 6A et 6B décrites ci-avant. A ce titre, l'une ou l'autre de ces figures peut constituer un organigramme représentant l'algorithme à la base du programme.

5 Revendications

1. Procédé de traitement de données sonores pour une restitution spatialisée en trois dimensions sur deux voies de restitution pour les oreilles respectives d'un auditeur, les données sonores étant initialement représentées dans un format multicanal puis encodées en compression (ENCOD) sur un nombre réduit de canaux (L,R), ledit format multicanal consistant à prévoir plus de deux canaux susceptibles d'alimenter des haut-parleurs respectifs, le procédé comportant les étapes :

- obtenir, avec les données compressées sur ledit nombre réduit de canaux, des paramètres de spatialisation (SPAT),
- pour chaque voie de restitution associée à une oreille de l'auditeur, former, à partir desdits paramètres de spatialisation, une combinaison de filtres représentatifs chacun de fonctions de transfert (HRTF) entre cette oreille de l'auditeur et des haut-parleurs susceptibles d'être alimentés par des canaux respectifs du format multicanal initial, et
- appliquer aux données compressées la combinaison de filtres ($h_{L,L}$, $h_{L,R}$, $h_{L,C}$; $h_{R,R}$, $h_{R,L}$, $h_{R,C}$) associée à chaque voie de restitution (L-BIN ; R-BIN),

caractérisé en ce que le procédé comporte en outre les étapes :

- pour chaque voie de restitution associée à une oreille de l'auditeur, déterminer à partir desdits paramètres de spatialisation au moins une fonction de transfert d'un haut-parleur situé à l'arrière de l'oreille de l'auditeur et représentative d'une décorrélation entre les canaux du format multicanal respectivement associés au haut-parleur arrière et à au moins un haut-parleur situé devant l'oreille de l'auditeur, et
- pour chaque voie de restitution, intégrer ladite fonction de transfert représentative d'une décorrélation dans ladite combinaison de filtres associée à cette voie de restitution.

2. Procédé selon la revendication 1, **caractérisé en ce que** la combinaison de filtres associée à une voie de restitution (L-BIN) comporte au moins un premier regroupement, formant un premier filtre ($h_{L,L}$), à partir :

- de la fonction de transfert d'un haut-parleur avant (HRTF-A),
- de la fonction de transfert d'un haut-parleur arrière (HRTF-C), et
- d'une version (HRTF-C*) de la fonction de transfert du haut-parleur arrière, représentative d'une décorrélation entre canaux,

et en ce que les haut-parleurs avant et arrière sont situés sur un même premier côté par rapport à l'auditeur.

3. Procédé selon la revendication 2, **caractérisé en ce que** ledit regroupement comporte une pondération, selon un coefficient choisi (α_1 ; α_2), entre :

- la fonction de transfert du haut-parleur situé à l'arrière, et
- la version représentative d'une décorrélation de cette fonction de transfert du haut-parleur arrière.

4. Procédé selon la revendication 3, **caractérisé en ce que** l'encodage en compression met en oeuvre un codeur paramétrique (ENCOD) délivrant une information de décorrélation entre canaux du format multicanal, **et en ce que** le coefficient de pondération est représenté par une fonction variable dynamiquement en fonction de l'information de décorrélation (ICC_L ; ICC_R) que délivre le codeur paramétrique.

5. Procédé selon l'une des revendications 2 à 4, dans lequel les données sonores sont encodées en compression sur deux canaux (L,R),

caractérisé en ce que la combinaison de filtres associée à ladite voie de restitution (L-BIN) comporte, outre ledit premier regroupement formant filtre ($h_{L,L}$) de l'un des canaux compressés (L), un second regroupement formant filtre ($h_{L,R}$) de l'autre des canaux compressés (R) à partir :

- de la fonction de transfert d'un haut-parleur avant (HRTF-H) situé sur un second côté, opposé au premier côté par rapport à l'auditeur,
- de la fonction de transfert d'un haut-parleur arrière (HRTF-E) situé sur ledit second côté, et
- d'une version (HRTF-E*) de la fonction de transfert de ce haut-parleur arrière, représentative d'une décorrélation entre canaux.

6. Procédé selon l'une des revendications précédentes, **caractérisé en ce que**, les données sonores étant encodées en compression dans un domaine transformé, la combinaison de filtres est appliquée dans le domaine transformé en fonction d'énergies cibles associées aux canaux du format multicanal, ces énergies cibles étant déterminées à partir desdits paramètres de spatialisation.
7. Procédé selon l'une des revendications précédentes, **caractérisé en ce que** lesdites fonctions de transfert des haut-parleurs sont de type HRTF et représentent des perturbations acoustiques sur des chemins entre chaque haut-parleur et une oreille pour une voie de restitution associée à cette oreille.
8. Procédé selon les revendications 6 et 7, dans lequel le domaine transformé est le domaine des sous-bandes, **caractérisé en ce que** les versions décorrélées des fonctions HRTF des haut-parleurs arrière sont obtenues en appliquant aux fonctions HRTF initiales des haut-parleurs arrière un déphasage qui est fonction de chaque sous-bande de fréquences.
9. Module de décodage (DECOD BIN) pour une restitution spatialisée en trois dimensions sur deux voies de restitution, **caractérisé en ce qu'il** comporte des moyens de traitement de données sonores pour la mise en oeuvre du procédé selon l'une des revendications précédentes.
10. Programme informatique, destiné à être stocké dans une mémoire d'un module de décodage pour une restitution spatialisée en trois dimensions sur deux voies de restitution, **caractérisé en ce qu'il** comporte des instructions pour l'exécution du procédé selon l'une des revendications 1 à 8.

Claims

1. Method of processing sound data for three-dimensional spatialized reproduction on two reproduction tracks for the respective ears of a listener, the sound data being initially represented in a multichannel format and then compression encoded (ENCOD) on a reduced number of channels (L,R), said multichannel format consisting in providing for more than two channels able to feed respective loudspeakers, the method comprising the steps:
 - obtaining, with the data compressed on said reduced number of channels, spatialization parameters (SPAT),
 - forming, on the basis of said spatialization parameters, for each reproduction track associated with an ear of the listener, a combination of filters each representative of transfer functions (HRTF) between this ear of the listener and loudspeakers able to be fed by respective channels of the initial multichannel format, and
 - applying to the compressed data the filter combination ($h_{L,L}$, $h_{L,R}$, $h_{L,C}$; $h_{R,R}$, $h_{R,L}$, $h_{R,C}$) associated with each reproduction track (L-BIN; R-BIN),**characterized in that** the method furthermore comprises the steps:
 - determining on the basis of said spatialization parameters, for each reproduction track associated with an ear of the listener, at least one transfer function of a loudspeaker situated at the rear of the ear of the listener and representative of a decorrelation between the channels of the multichannel format which are respectively associated with the rear loudspeaker and with at least one loudspeaker situated in front of the ear of the listener, and
 - integrating, for each reproduction track, said transfer function representative of a decorrelation in said filter combination associated with this reproduction track.
2. Method according to Claim 1, **characterized in that** the filter combination associated with a reproduction track (L-BIN) comprises at least one first grouping, forming a first filter ($h_{L,L}$), on the basis:
 - of the transfer function of a front loudspeaker (HRTF-A),

- of the transfer function of a rear loudspeaker (HRTF-C), and
- of a version (HRTF-C*) of the transfer function of the rear loudspeaker, representative of a decorrelation between channels,

and **in that** the front and rear loudspeakers are situated on one and the same first side with respect to the listener.

3. Method according to Claim 2, **characterized in that** said grouping comprises a weighting, according to a chosen coefficient (α_1 ; α_2), between:

- the transfer function of the loudspeaker situated at the rear, and
- the version representative of a decorrelation of this transfer function of the rear loudspeaker.

4. Method according to Claim 3, **characterized in that** the compression encoding implements a parametric coder (ENCOD) delivering an information cue regarding the decorrelation between channels of the multichannel format, and **in that** the weighting coefficient is represented by a function that can vary dynamically as a function of the decorrelation information cue (ICC_L ; ICC_R) delivered by the parametric coder.

5. Method according to one of Claims 2 to 4, in which the sound data are compression encoded on two channels (L,R), **characterized in that** the filter combination associated with said reproduction track (L-BIN) comprises, in addition to said first filter-forming grouping ($h_{L,L}$) of one of the compressed channels (L), a second filter-forming grouping ($h_{L,R}$) of the other of the compressed channels (R) on the basis:

- of the transfer function of a front loudspeaker (HRTF-H) situated on a second side, opposite from the first side with respect to the listener,
- of the transfer function of a rear loudspeaker (HRTF-E) situated on said second side, and
- of a version (HRTF-E*) of the transfer function of this rear loudspeaker, representative of a decorrelation between channels.

6. Method according to one of the preceding claims, **characterized in that**, the sound data being compression encoded in a transformed domain, the combination of filters is applied in the transformed domain as a function of target energies associated with the channels of the multichannel format, these target energies being determined on the basis of said spatialization parameters.

7. Method according to one of the preceding claims, **characterized in that** said transfer functions of the loudspeakers are of HRTF type and represent acoustic disturbances on paths between each loudspeaker and an ear for a reproduction track associated with this ear.

8. Method according to Claims 6 and 7, in which the transformed domain is the sub-band domain, **characterized in that** the decorrelated versions of the HRTF functions of the rear loudspeakers are obtained by applying to the initial HRTF functions of the rear loudspeakers a phase shift which is dependent on each sub-band of frequencies.

9. Decoding module (DECOD BIN) for three-dimensional spatialized reproduction on two reproduction tracks, **characterized in that** it comprises means for processing sound data for the implementation of the method according to one of the preceding claims.

10. Computer program, intended to be stored in a memory of a decoding module for three-dimensional spatialized reproduction on two reproduction tracks, **characterized in that** it comprises instructions for the execution of the method according to one of Claims 1 to 8.

Patentansprüche

1. Verfahren zur Verarbeitung von Tondaten für eine dreidimensionale verräumlichte Wiedergabe auf zwei Wiedergabepfaden für die Ohren eines Hörers, wobei die Tondaten anfangs in einem Mehrkanalformat dargestellt und dann auf eine reduzierte Anzahl von Kanälen (L,R) kompressionscodiert (ENCOD) werden, wobei das Mehrkanalformat darin besteht, mehr als zwei Kanäle vorzusehen, die jeweilige Lautsprecher speisen können,

wobei das Verfahren die Schritte aufweist:

- mit den auf die reduzierte Anzahl von Kanälen komprimierten Daten Verräumlichungsparameter (SPAT) zu erhalten,
- für jeden einem Ohr des Hörers zugeordneten Wiedergabepfad ausgehend von den Verräumlichungsparametern eine Kombination von Filtern zu formen, die je für Transferfunktionen (HRTF) zwischen diesem Ohr des Hörers und Lautsprechern repräsentativ sind, die von Kanälen des Anfangs-Mehrkanalformats gespeist werden können, und
- an die komprimierten Daten die Kombination von Filtern ($h_{L,L}$, $h_{L,R}$, $h_{L,C}$; $h_{R,R}$, $h_{R,L}$, $h_{R,C}$) anzuwenden, die jedem Wiedergabepfad (L-BIN; R-BIN) zugeordnet ist,

dadurch gekennzeichnet, dass das Verfahren außerdem die Schritt aufweist:

- für jeden einem Ohr des Hörers zugeordneten Wiedergabepfad ausgehend von den Verräumlichungsparametern mindestens eine Transferfunktion eines hinter dem Ohr des Hörers befindlichen Lautsprechers zu bestimmen, die für eine Dekorrelation zwischen den Kanälen des Mehrkanalformats repräsentativ ist, welche dem hinteren Lautsprecher bzw. mindestens einem vor dem Ohr des Hörers befindlichen Lautsprecher zugeordnet sind, und
- für jeden Wiedergabepfad die für eine Dekorrelation repräsentative Transferfunktion in die Kombination von Filtern zu integrieren, die diesem Wiedergabepfad zugeordnet ist.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** die einem Wiedergabepfad (L-BIN) zugeordnete Kombination von Filtern mindestens eine erste Gruppierung aufweist, die ein erstes Filter ($h_{L,L}$) formt, ausgehend von:

- der Transferfunktion eines vorderen Lautsprechers (HRTF-A),
- der Transferfunktion eines hinteren Lautsprechers (HRTF-C), und
- einer Version (HRTF-C*) der Transferfunktion des hinteren Lautsprechers, repräsentativ für eine Dekorrelation zwischen Kanälen,

und dass der vordere und der hintere Lautsprecher sich bezüglich des Hörers auf der gleichen Seite befinden.

3. Verfahren nach Anspruch 2, **dadurch gekennzeichnet, dass** die Gruppierung eine Gewichtung, gemäß einem gewählten Koeffizienten (α_1 ; α_2), aufweist zwischen:

- der Transferfunktion des hinten befindlichen Lautsprechers, und
- der für eine Dekorrelation dieser Transferfunktion des hinteren Lautsprechers repräsentativen Version.

4. Verfahren nach Anspruch 3, **dadurch gekennzeichnet, dass** die Kompressionscodierung einen parametrischen Codierer (ENCOD) verwendet, der eine Dekorrelationsinformation zwischen Kanälen des Mehrkanalformats liefert, und dass der Gewichtungskoeffizient durch eine in Abhängigkeit von der Dekorrelationsfunktion (ICC_L ; ICC_R), die der parametrische Codierer liefert, dynamisch variable Funktion dargestellt wird.

5. Verfahren nach einem der Ansprüche 2 bis 4, bei dem die Tondaten auf zwei Kanälen (L,R) kompressionscodiert werden,

dadurch gekennzeichnet, dass die dem Wiedergabepfad (L-BIN) zugeordnete Kombination von Filtern außer der ersten ein Filter formenden Gruppierung ($h_{L,L}$) eines der komprimierten Kanäle (L) eine zweite ein Filter formende Gruppierung ($h_{L,R}$) des anderen der komprimierten Kanäle (R) aufweist, ausgehend von:

- der Transferfunktion eines vorderen Lautsprechers (HRTF-H), der sich auf einer zweiten Seite entgegengesetzt zur ersten Seite bezüglich des Hörers befindet,
- der Transferfunktion eines hinteren Lautsprechers (HRTF-E), der sich auf der zweiten Seite befindet, und
- einer Version (HRTF-E*) der Transferfunktion dieses hinteren Lautsprechers, die für eine Dekorrelation zwischen Kanälen repräsentativ ist.

6. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass**, wenn die Tondaten in einem transformierten Bereich kompressionscodiert werden, die Kombination von Filtern im transformierten Bereich in Abhängigkeit von Zielenergien angewendet wird, die den Kanälen des Mehrkanalformats zugeordnet sind, wobei diese Zielenergien ausgehend von den Verräumli-

chungsparametern bestimmt werden.

7. Verfahren nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet, dass die Transferfunktionen der Lautsprecher vom Typ HRTF sind und akustische Störungen auf Wegen zwischen jedem Lautsprecher und einem Ohr für einen diesem Ohr zugeordneten Wiedergabepfad darstellen.
8. Verfahren nach den Ansprüchen 6 und 7, bei dem der transformierte Bereich der Bereich der Unterbänder ist,
dadurch gekennzeichnet, dass die dekorrelierten Versionen der HRTF-Funktionen der hinteren Lautsprecher erhalten werden, indem an die HRTF-Anfangsfunktionen der hinteren Lautsprecher einer Phasenverzögerung angewendet wird, die von jedem Frequenz-Unterband abhängt.
9. Decodiermodul (DECOD BIN) für eine verräumlichte dreidimensionale Wiedergabe auf zwei Wiedergabepfaden,
dadurch gekennzeichnet, dass es Einrichtungen zur Verarbeitung von Tondaten zur Durchführung des Verfahrens nach einem der vorhergehenden Ansprüche aufweist.
10. EDV-Programm, das dazu bestimmt ist, in einem Speicher eines Decodiermoduls für eine verräumlichte dreidimensionale Wiedergabe auf zwei Wiedergabepfaden gespeichert zu werden, **dadurch gekennzeichnet, dass** es Anweisungen zur Ausführung des Verfahrens nach einem der Ansprüche 1 bis 8 aufweist.

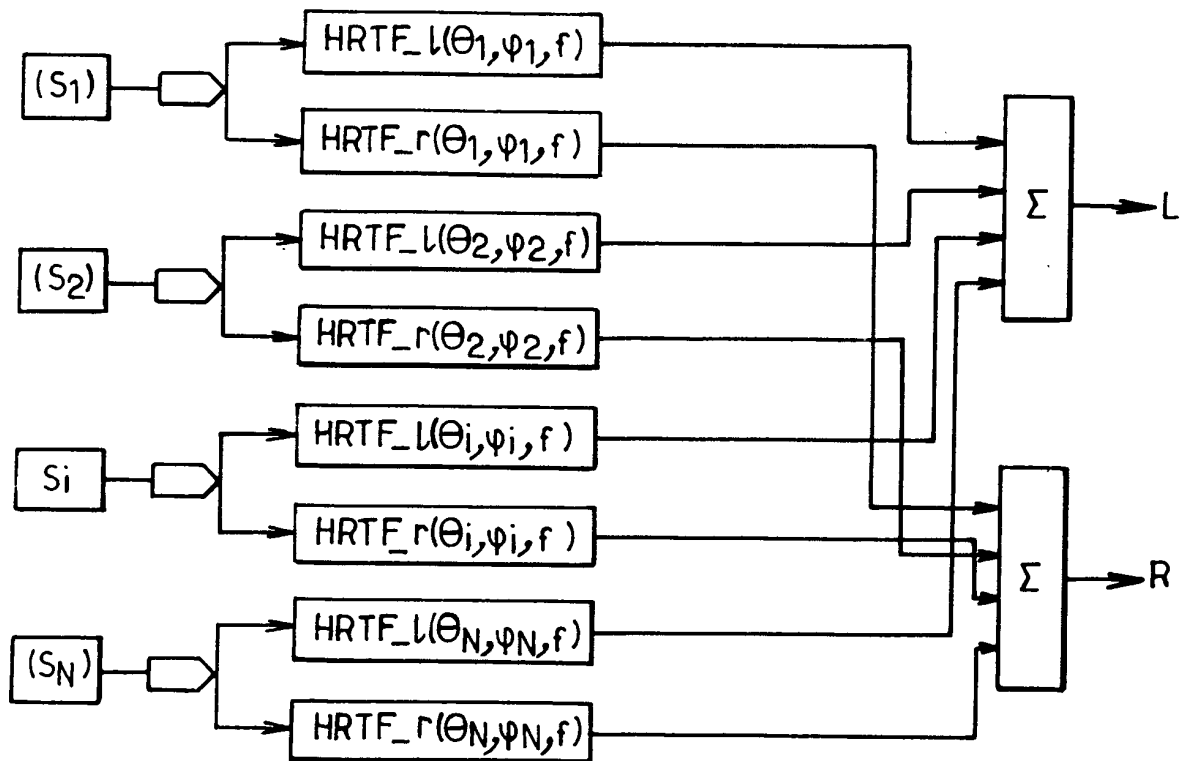


FIG.1.
(ART ANTÉRIEUR)

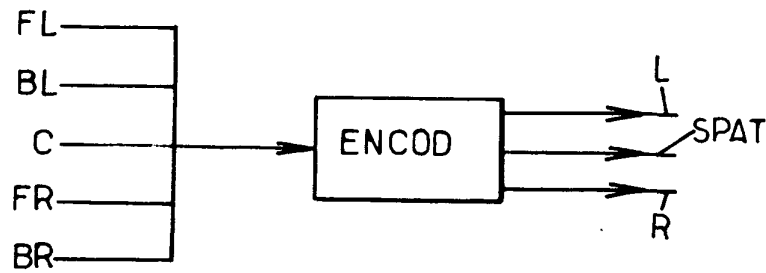


FIG.2A.

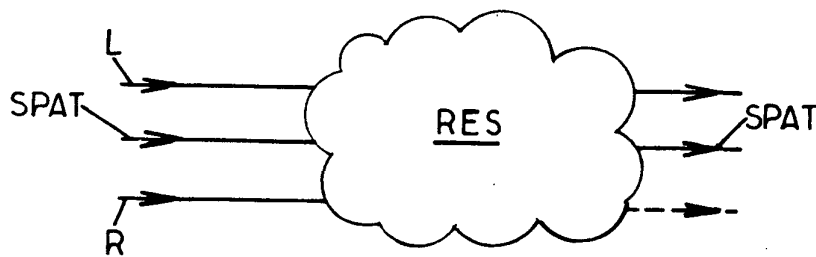


FIG.2B.

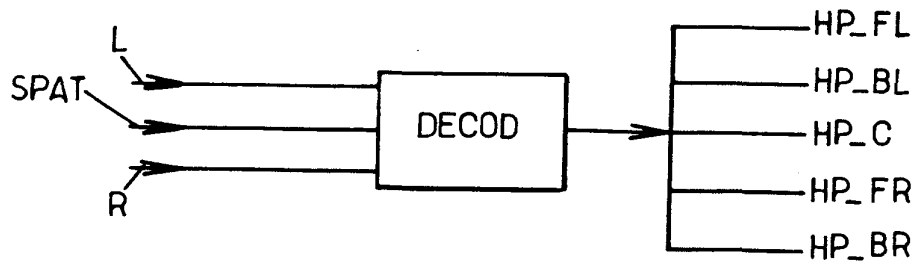


FIG.2C.

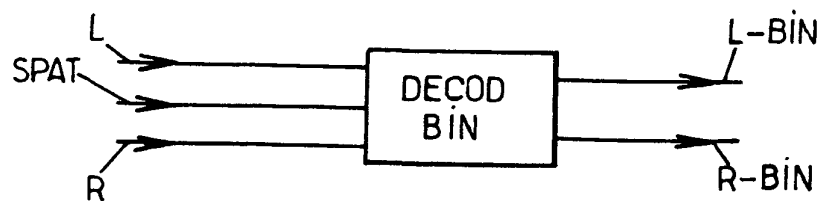
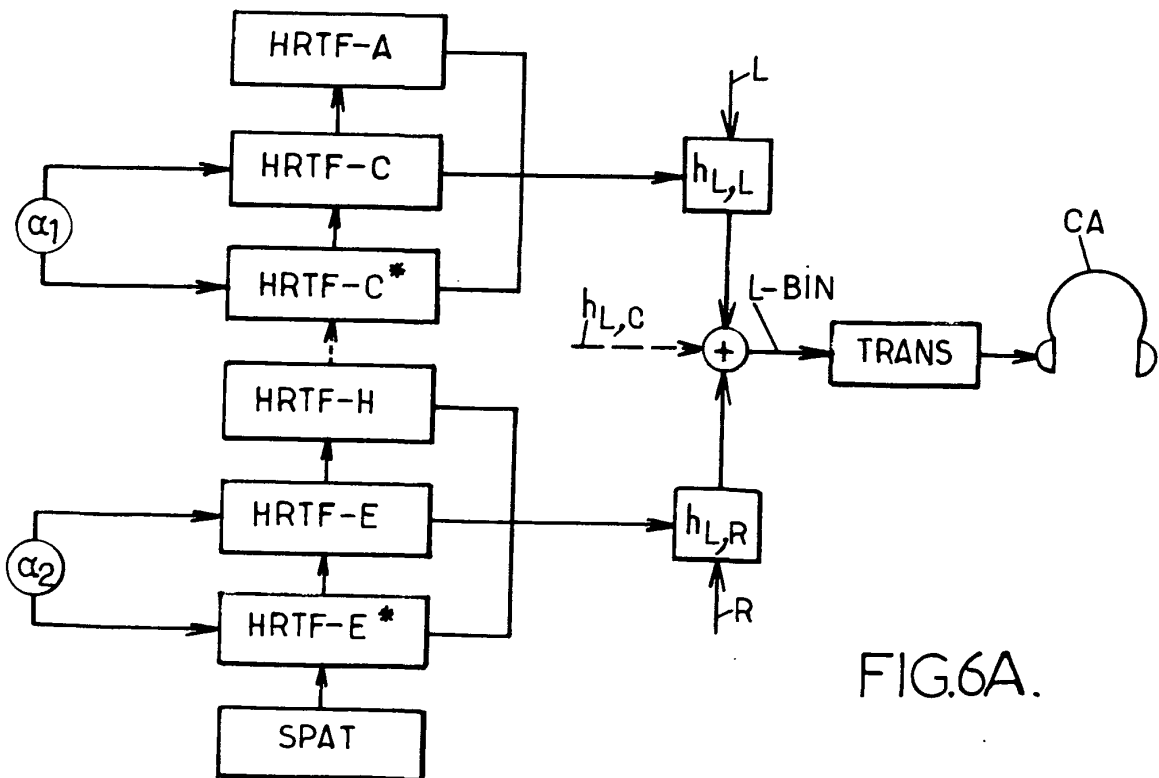
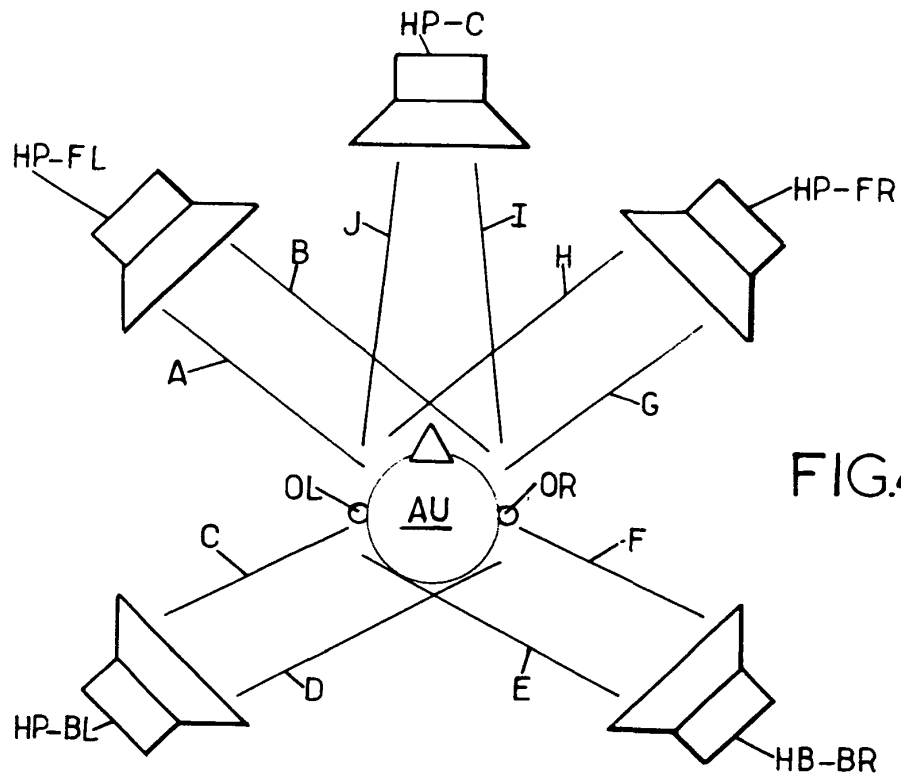


FIG.3.



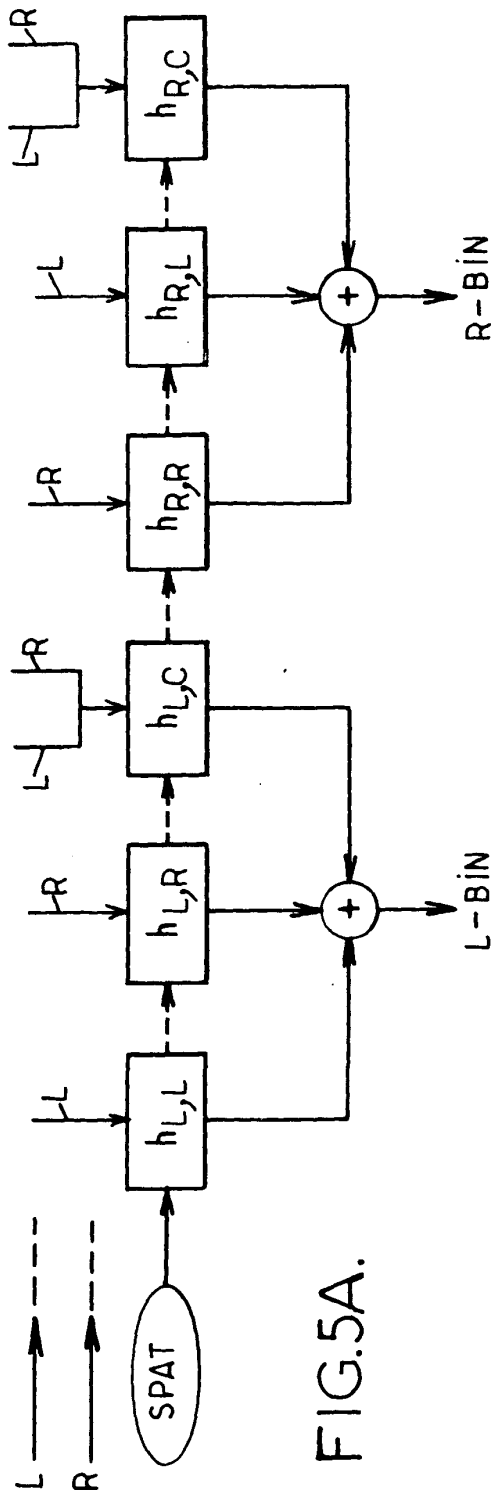


FIG.5A.

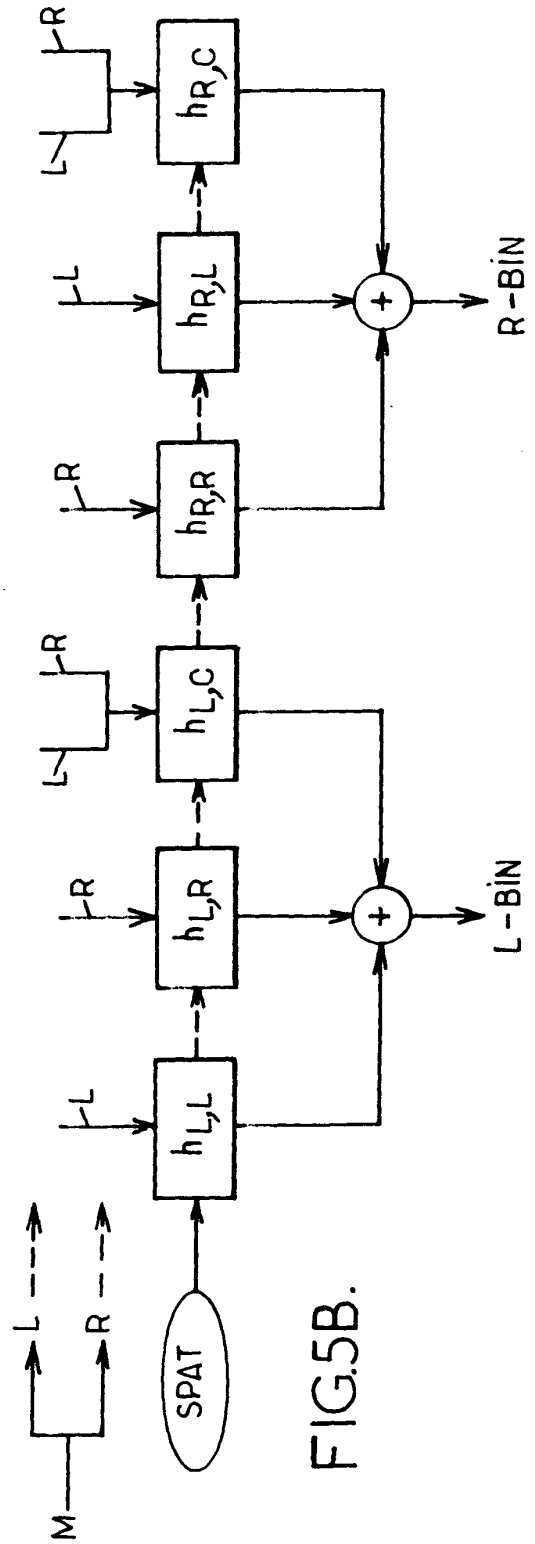


FIG.5B.

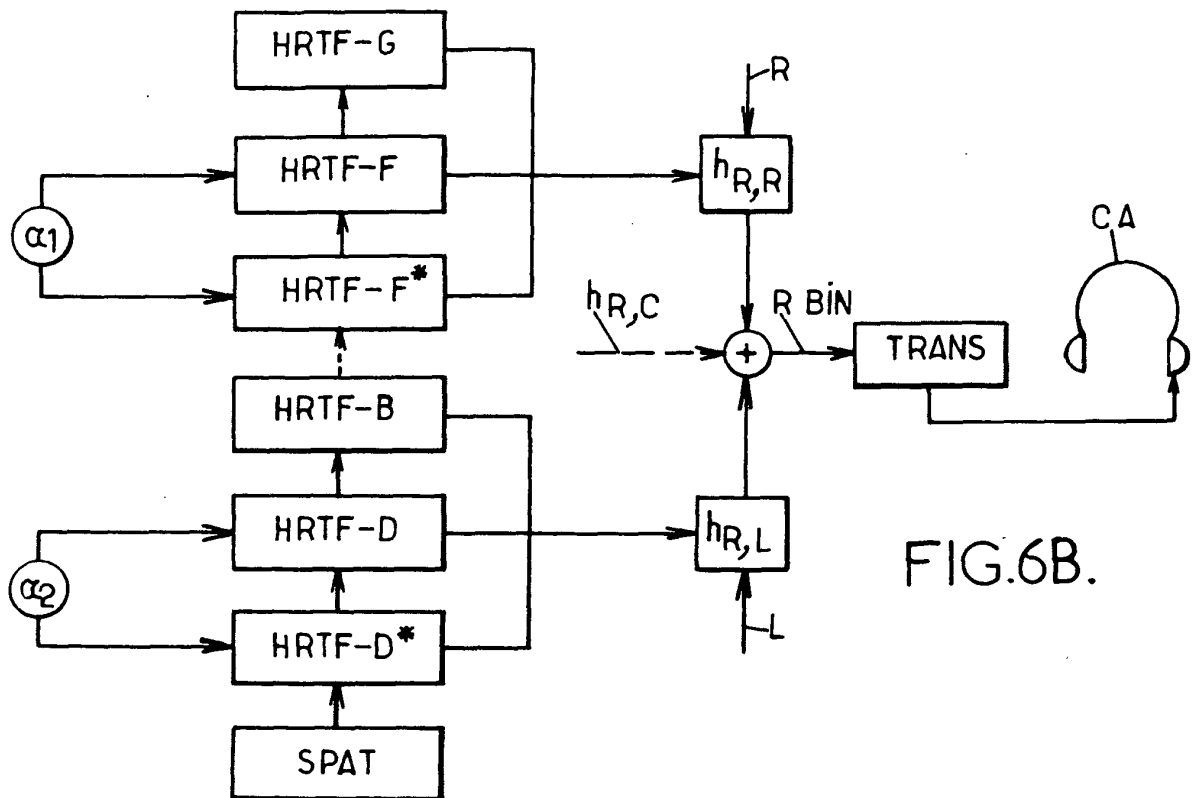


FIG. 6B.

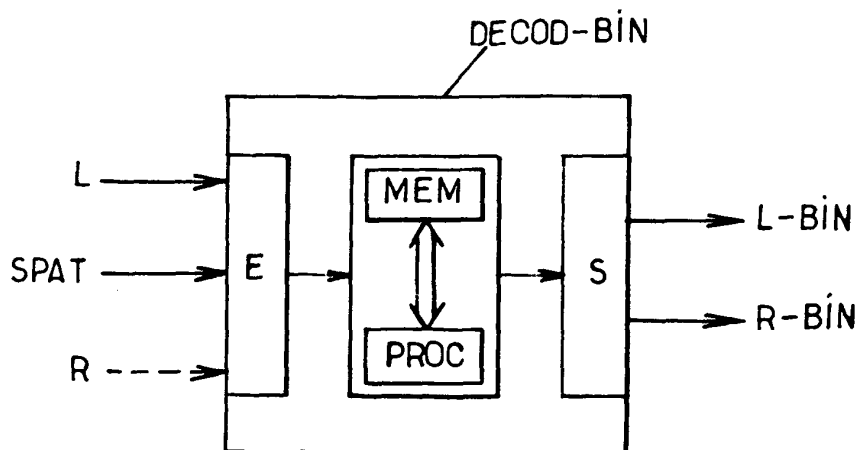


FIG. 7.

RÉFÉRENCES CITÉES DANS LA DESCRIPTION

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

Documents brevets cités dans la description

- US 2005047618 A [0006]

Littérature non-brevet citée dans la description

- **A.Benjelloun Touimi.** A Generic Framework for Filtering in Subband Domain. *Proceeding IEEE - 9th Workshop on Digital Signal Processing*, Octobre 2000 [0018]
- **J.Breebaart et al.** MPEG Spatial Audio Coding / MPEG Surround: Overview and Current Status. *119th Conv. Aud. Eng. Soc (AES)*, Octobre 2005 [0028]