



(11) **EP 1 836 876 B1**

(12) **FASCICULE DE BREVET EUROPEEN**

(45) Date de publication et mention
de la délivrance du brevet:
18.07.2018 Bulletin 2018/29

(21) Numéro de dépôt: **06709051.4**

(22) Date de dépôt: **09.01.2006**

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

(86) Numéro de dépôt international:
PCT/FR2006/000037

(87) Numéro de publication internationale:
WO 2006/075077 (20.07.2006 Gazette 2006/29)

(54) **PROCÉDÉ ET DISPOSITIF D'INDIVIDUALISATION DE HRTFS PAR MODÉLISATION**

VERFAHREN UND VORRICHTUNG ZUR INDIVIDUALISIERUNG VON HRTFS DURCH
MODELLIERUNG

METHOD AND DEVICE FOR INDIVIDUALIZING HRTFS BY MODELING

(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 NL PL PT RO SE SI
SK TR**

(30) Priorité: **10.01.2005 FR 0500218**

(43) Date de publication de la demande:
26.09.2007 Bulletin 2007/39

(73) Titulaire: **Orange**
75015 Paris (FR)

(72) Inventeurs:
• **NICOL, Rozenn**
F-22450 La Roche Derrien (FR)
• **BUSSON, Sylvain**
F-22300 Ploumilliau (FR)
• **LEMAIRE, Vincent**
F-22560 Trebeurden (FR)

(74) Mandataire: **Cabinet Plasseraud**
66, rue de la Chaussée d'Antin
75440 Paris Cedex 09 (FR)

(56) Documents cités:
WO-A-97/25834 WO-A-2004/028205
DE-A1- 19 910 372 US-A1- 2003 138 107

- **KISTLER D J ET AL: "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction" JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA USA, vol. 91, no. 3, mars 1992 (1992-03), pages 1637-1647, XP002099514 ISSN: 0001-4966**
- **Rick L Jenison ET AL: "A Spherical Basis Function Neural Network for Modeling Auditory Space", NEURAL COMPUTATION., vol. 8, no. 1, 1 January 1996 (1996-01-01), pages 115-128, XP055356520, US ISSN: 0899-7667, DOI: 10.1162/neco.1996.8.1.115**

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] La présente invention concerne la modélisation de fonctions de transfert individuelles dites HRTFs (pour "*Head Related Transfer Functions*"), relatives à l'audition d'un individu dans l'espace tridimensionnel.

[0002] L'invention s'inscrit en particulier dans le cadre de services de télécommunication proposant une diffusion du son spatialisée (par exemple une audioconférence entre plusieurs locuteurs, une diffusion de bande annonce de cinéma). Sur des terminaux de télécommunication, notamment mobiles, il est envisagé un rendu sonore avec un casque d'écoute stéréophonique. La technique la plus efficace de positionnement des sources sonores dans l'espace est alors la synthèse binaurale.

[0003] La synthèse binaurale repose sur l'utilisation de filtres, dits "*binauraux*", qui reproduisent les fonctions de transfert acoustiques entre la source sonore et les oreilles de l'auditeur. Ces filtres servent à simuler les indices de localisation auditive, indices qui permettent à un auditeur de localiser les sources sonores en situation d'écoute réelle. Ces filtres prennent en compte l'ensemble des phénomènes acoustiques (notamment la diffraction par la tête, les réflexions sur le pavillon de l'oreille et le haut du torse) qui modifient l'onde acoustique dans son trajet entre la source et les oreilles de l'auditeur. Ces phénomènes varient fortement avec la position de la source sonore (principalement avec sa direction) et ces variations permettent à l'auditeur de localiser la source dans l'espace. En effet, ces variations déterminent une sorte de codage acoustique de la position de la source. Le système auditif d'un individu sait, par apprentissage, interpréter ce codage pour localiser les sources sonores. Néanmoins, les phénomènes acoustiques de diffraction/réverbération dépendent tout aussi fortement de la morphologie de l'individu. Une synthèse binaurale de qualité repose donc sur des filtres binauraux qui reproduisent au mieux le codage acoustique que produit naturellement le corps de l'auditeur, en prenant en compte les spécificités individuelles de sa morphologie. Lorsque ces conditions ne sont pas respectées, on observe une dégradation des performances du rendu binaural, qui se traduit notamment par une perception intracrânienne des sources et des confusions avant/arrière. Les sources situées à l'avant sont perçues à l'arrière et vice versa.

[0004] Parmi les technologies du son 3D ou de spatialisation sonore, en traitement du signal audio appliqué notamment à la simulation de phénomènes acoustiques et psycho-acoustiques, certaines visent la génération de signaux à diffuser sur haut-parleurs ou sur écouteurs, afin de donner à l'auditeur l'illusion auditive de sources sonores placées à des positions respectives particulières autour de lui. On parle alors de création de sources et d'images sonores virtuelles.

[0005] Les techniques binaurales décrites ci-avant sont appliquées au traitement d'un son 3D destiné à une diffusion sur casque à deux oreillettes gauche et droite. Ces techniques visent la reconstruction du champ sonore

au niveau des oreilles d'un auditeur, de telle sorte que ses tympans perçoivent un champ sonore pratiquement identique à celui qu'auraient induit les sources réelles dans l'espace 3D. Les techniques binaurales sont donc basées sur une paire de signaux binauraux qui alimentent respectivement les deux écouteurs du casque. Ces signaux binauraux peuvent être obtenus de deux façons :

- par une prise de son directe, au moyen de deux microphones insérés à l'entrée du canal auditif d'un individu ou d'un mannequin à morphologie standard ("*tête artificielle*"), ou
- par traitement du signal, en filtrant un signal monophonique par deux filtres binauraux, ces filtres reproduisant les propriétés de la propagation acoustique entre la source placée à une position donnée et les deux oreilles d'un auditeur.

[0006] Les techniques binaurales mettant en oeuvre des filtres binauraux définissent le domaine de la synthèse binaurale dans un contexte avantageux de la présente invention. La synthèse binaurale repose sur les filtres binauraux qui modélisent la propagation de l'onde acoustique entre la source et les deux oreilles de l'auditeur. Ces filtres représentent des fonctions de transfert acoustiques appelées HRTFs qui modélisent les transformations engendrées par le torse, la tête et le pavillon de l'auditeur sur le signal provenant d'une source sonore. A chaque position de source sonore est associée une paire de HRTFs (une HRTF pour l'oreille droite, une HRTF pour l'oreille gauche). De plus, les HRTFs portent l'empreinte acoustique de la morphologie de l'individu sur lequel elles ont été mesurées.

[0007] Les HRTFs dépendent donc non seulement de la direction du son, mais aussi de l'individu. Elles sont ainsi fonction de la fréquence f , de la position (θ, φ) de la source sonore (où l'angle θ représente l'azimut et l'angle φ l'élévation), de l'oreille (gauche ou droite) et de l'individu.

[0008] De manière classique, les HRTFs sont obtenues par la mesure. On fixe initialement une sélection de directions qui couvrent plus ou moins finement l'ensemble de l'espace entourant l'auditeur. Pour chaque direction, les HRTFs gauche et droite sont mesurées au moyen de microphones insérés à l'entrée du conduit auditif d'un sujet. La mesure doit être réalisée dans une chambre anéchoïque (ou "*chambre sourde*"). Au final, si M directions sont mesurées, on obtient, pour un sujet donné, une base de données de $2M$ fonctions de transfert acoustiques représentant chaque position de l'espace pour chaque oreille.

[0009] Dans le contexte avantageux de la synthèse binaurale, l'effet de spatialisation repose sur l'utilisation de HRTFs qui, pour des performances optimales, doivent prendre en compte les phénomènes de propagation acoustique entre la source et les oreilles, mais aussi les spécificités individuelles de la morphologie de l'auditeur. La mesure expérimentale des HRTFs directement sur

un individu est, à l'heure actuelle, la solution la plus fiable pour obtenir des filtres binauraux de qualité et réellement individualisés (tenant compte des spécificités individuelles de la morphologie de l'individu). On rappelle qu'il s'agit de mesurer la fonction de transfert entre une source située à une position donnée (θ_1 , ϕ_1) et les deux oreilles du sujet au moyen de microphones placés à l'entrée des conduits auditifs de cette personne.

[0010] Toutefois, la mesure de ces fonctions de transfert HRTFs présente quelques difficultés. Elle requiert un équipement spécifique et coûteux (typiquement une chambre sourde, un microphone, un dispositif mécanique de positionnement de sources). Cette opération est longue car il faut notamment mesurer les fonctions de transfert pour un grand nombre de directions afin de couvrir de façon homogène l'ensemble d'une sphère 3D entourant l'auditeur.

[0011] Cette mesure des HRTFs devient très difficile, voire impossible, dans le cadre d'applications de la synthèse binaurale destinée au grand public. La mesure des HRTFs pose en fait au moins trois problèmes principaux :

- La mesure des HRTF en elle-même est difficile à mettre en oeuvre, car elle nécessite un équipement spécifique. La mesure doit être réalisée dans une chambre anéchoïque. Elle requiert aussi un dispositif mécanique pour déplacer et piloter le haut-parleur de mesure afin d'effectuer des mesures pour un grand nombre de directions uniformément réparties en azimut et en élévation autour de l'auditeur. En outre, la procédure de mesure dans l'ensemble est pénible pour le sujet, à cause des contraintes imposées au sujet par le système de mesure et à cause de la durée de la mesure.
- Un second problème réside dans la nécessité de mesurer les HRTFs dans un grand nombre de directions pour offrir un échantillonnage spatial suffisant et homogène de la sphère 3D entourant l'auditeur. Plus le nombre de directions mesurées est élevé, plus la durée du test est longue, ce qui accroît l'inconfort du sujet.
- Un troisième problème concerne la mesure d'un individu en particulier. Offrir une synthèse binaurale performante à un individu quelconque suppose d'utiliser ses HRTFs propres, qu'il aura fallu mesurer au préalable, ce qui est en général impossible.

On a donc recherché des solutions nécessitant un minimum de mesures de HRTFs et implémentant davantage des techniques de modélisation. En particulier, on a recherché des modèles mathématiques de HRTFs consistant en une fonction F permettant d'exprimer une HRTF (Y) à partir d'un jeu de paramètres (X) donnés a priori, tels que $Y = F(X)$. Souvent, deux éléments essentiels interviennent :

- la mise au point du modèle mathématique (fonction F), et

- la spécification du jeu de paramètres à appliquer en entrée du modèle.

On présente ci-après l'état de l'art à la connaissance des inventeurs sur les modélisations des HRTFs mises en oeuvre à ce jour, en portant attention au choix des paramètres d'entrée des modèles.

[0012] L'article de Rick L Jenison ET AL: "A Spherical Basis Function Neural Network for Modeling Auditory Space", NEURAL COMPUTATION, vol. 8, no. 1, janvier 1996, pages 115-128, présente une technique de modélisation appliquée à l'interpolation de HRTFs.

[0013] Dans le document US-2003/138107, on a présenté un modèle statistique de HRTFs à partir de données morphologiques. Cette approche part d'une analyse statistique appliquée à une base de données incluant des HRTFs et des données morphologiques. Une analyse en composantes principales est d'abord appliquée d'une part aux HRTFs et d'autre part aux données morphologiques, ce qui permet de décrire l'ensemble des données avec un nombre réduit de composantes. Ensuite, une régression linéaire est effectuée entre les composantes issues de l'analyse en composantes principales des HRTFs et les composantes issues de celle des données morphologiques. On établit ainsi un modèle statistique reliant les données morphologiques aux HRTFs. Il suffit ensuite de mesurer les paramètres morphologiques d'un individu quelconque pour prédire ses HRTFs à partir du modèle statistique obtenu.

[0014] Un mode de réalisation de ce document prévoit en particulier d'enrichir les données morphologiques d'un individu, en entrée du modèle, par quelques HRTFs mesurées sur cet individu et dans des directions respectives spécifiques. Ainsi, seul un nombre restreint de directions de mesure est utile pour obtenir les HRTFs de l'individu dans toutes les directions de l'espace.

[0015] Néanmoins, même si le nombre de mesures est restreint dans ce document, il faut encore respecter le protocole de mesure de HRTFs, en particulier prévoir une chambre anéchoïque pour les mesures et positionner strictement les sources à des distances bien précises des microphones qui sont accolés aux oreilles de l'individu.

[0016] La mise en oeuvre de la présente invention s'affranchit de telles contraintes.

[0017] La présente invention vise à cet effet un procédé de modélisation de fonctions de transfert HRTFs propres à un individu, dans lequel :

- a) on constitue une base de données incluant une pluralité de HRTFs suivant une multiplicité de directions de l'espace et pour une pluralité d'individus,
- b) par apprentissage sur ladite base de données, on construit un modèle propre à donner des HRTFs pour ladite multiplicité de directions, à partir d'un jeu de mesures représentatives de HRTFs dans des directions respectives sélectionnées parmi ladite multiplicité de directions, et

c) pour un individu quelconque :

- c1) on mesure un jeu de fonctions représentatives des HRTFs de l'individu dans lesdites directions sélectionnées seulement,
- c2) on applique le modèle auxdites mesures dans les directions sélectionnées, et
- c3) on obtient les HRTFs de l'individu dans toute ladite multiplicité de directions.

[0018] En outre, dans le procédé au sens de l'invention :

- les conditions et les directions de mesure pour obtenir ledit jeu de mesures sont fixées arbitrairement pendant l'étape d'apprentissage b), et
- on applique, à l'étape c1), des conditions de mesure sensiblement reproductibles avec les conditions de mesure de l'étape b).

[0019] Ainsi, selon un aspect de l'invention, on peut fixer arbitrairement, dès l'étape d'apprentissage, les conditions et les directions dans lesquelles les fonctions représentatives des HRTFs vont être mesurées. On entend par le terme "*arbitrairement*" le fait que ces mesures ne sont pas forcément des directions privilégiées pour que le modèle donne de meilleurs résultats. On comprendra donc que ces conditions et/ou ces directions de mesure peuvent être choisies pour des raisons indépendantes du bon fonctionnement du modèle. Par ailleurs, les conditions de mesure ne sont pas nécessairement optimales. C'est la raison pour laquelle on parle ici de "*mesures représentatives de HRTFs*" au lieu de "*mesures de HRTFs*".

[0020] Toutefois, les conditions de mesure de l'étape c1), sur un individu quelconque, doivent préférentiellement être reproductibles avec celles qui ont permis de constituer le modèle à l'étape b). Ainsi, ces conditions de mesure peuvent être choisies selon des critères complètement indépendants du fonctionnement du modèle, l'essentiel étant qu'elles soient reproductibles entre le moment où l'on constitue le modèle, à l'étape b), et le moment où l'on mène les mesures sur un individu quelconque, à l'étape c).

[0021] Ainsi, selon l'un des avantages que procure la présente invention, l'obtention des HRTFs complètes d'un individu quelconque peut s'effectuer en mesurant grossièrement ses HRTFs dans quelques directions seulement, avec une procédure de mesure allégée (c'est-à-dire n'impliquant qu'un nombre réduit de directions de mesure et/ou un dispositif de mesure simplifié).

[0022] Dans une réalisation préférée, le modèle est construit en mettant en oeuvre un réseau de neurones artificiels. Cette catégorie de modèles mathématiques performants est capable d'identifier et de reproduire des dépendances de haut niveau entre les variables d'entrée et de sortie, sans se limiter à des solutions triviales. Il est alors possible d'appliquer en entrée du modèle des pa-

ramètres dont la relation avec les HRTFs n'est pas forcément évidente, mais à partir desquels le modèle sera néanmoins capable d'extraire des informations permettant de calculer les HRTFs complètes d'un individu quelconque.

La présente invention vise aussi une installation pour la mise en oeuvre du procédé ci-avant et, plus particulièrement, pour l'estimation de fonctions de transfert HRTFs propres à un individu. Cette installation comporte :

- une cabine de mesure de fonctions de transfert représentatives de HRTFs dans un jeu de directions choisies, et
- une unité de traitement pour récupérer un jeu de mesures sur un individu dans lesdites directions choisies et évaluer les HRTFs de l'individu dans une multiplicité de directions de l'espace incluant lesdites directions choisies, à partir d'un modèle capable de donner des HRTFs pour une multiplicité de directions, à partir d'un jeu de mesures représentatives de HRTFs dans quelques directions seulement fixées arbitrairement parmi ladite multiplicité de directions,

dans laquelle les directions de mesure dans ladite cabine correspondent auxdites directions fixées arbitrairement, et dans laquelle, la cabine comportant des repères sur des axes respectifs dont l'intersection définit une position des oreilles de l'individu dans la cabine, des sources sonores, prévues dans ladite cabine pour effectuer lesdites mesures, sont placées à des distances différentes de ladite intersection. Au sens de l'invention, les directions de mesure dans la cabine précitée correspondent alors auxdites directions fixées arbitrairement, pour respecter les conditions de mesure entre l'apprentissage du modèle et son utilisation ultérieure.

La présente invention vise aussi un produit programme d'ordinateur pour constituer le modèle. Ce programme peut être stocké dans une mémoire d'une unité de traitement ou sur un support amovible propre à coopérer avec un lecteur de cette unité de traitement, ou encore être transmis d'un serveur vers l'unité de traitement, notamment via un réseau étendu. Le programme comporte alors des instructions sous forme de code informatique pour construire un modèle capable de donner des fonctions de transfert HRTFs d'un individu pour une multiplicité de directions, à partir d'un jeu de mesures, effectuées sur cet individu, représentatives de HRTFs dans quelques directions seulement, et fixées arbitrairement parmi ladite multiplicité de directions, le programme mettant en oeuvre, à partir d'une base de données incluant une pluralité de HRTFs suivant une multiplicité de directions de l'espace et pour une pluralité d'individus, au moins une phase d'apprentissage.

La présente invention vise aussi un second produit programme d'ordinateur, destiné à être stocké dans une mémoire d'une unité de traitement ou sur un support amovible propre à coopérer avec un lecteur de ladite unité

de traitement, ou destiné à être transmis d'un serveur vers ladite unité de traitement. Ce second programme comporte, quant à lui, des instructions sous forme de code informatique pour mettre en oeuvre un modèle basé sur un réseau de neurones artificiels et capable de donner des fonctions de transfert HRTFs d'un individu pour une multiplicité de directions, à partir d'un jeu de mesures effectuées sur cet individu, représentatives de HRTFs dans quelques directions seulement, et fixées arbitrairement parmi ladite multiplicité de directions, les mesures effectuées sur l'individu étant réalisées dans une cabine dans laquelle:

- les directions de mesure correspondent auxdites directions fixées arbitrairement, et
- la cabine comportant des repères sur des axes respectifs dont l'intersection définit une position des oreilles de l'individu dans la cabine, des sources sonores, prévues dans ladite cabine pour effectuer lesdites mesures, sont placées à des distances différentes de ladite intersection.

Ainsi, le premier programme décrit ci-avant permet de construire le modèle, tandis que le second programme consiste en des instructions informatiques représentant le modèle lui-même.

D'autres caractéristiques et avantages de l'invention apparaîtront à l'examen de la description détaillée ci-après, et des dessins annexés sur lesquels :

- la figure 1 illustre schématiquement les étapes de fonctionnement d'un modèle mettant en oeuvre un réseau de neurones artificiels, pouvant correspondre alors à un organigramme représentant schématiquement le déroulement du second programme informatique décrit ci-avant,
- la figure 2 illustre schématiquement les étapes de construction du modèle, pouvant correspondre alors à un organigramme représentant schématiquement le déroulement du premier programme informatique décrit ci-avant,
- la figure 3 représente la variation d'une erreur de validation dans l'étape de construction du modèle en fonction du nombre total de mesures à effectuer pour utiliser le modèle,
- la figure 4a illustre schématiquement les étapes a) et b) du procédé au sens de l'invention,
- la figure 4b illustre schématiquement l'étape c) du procédé au sens de l'invention,
- la figure 4c illustre schématiquement une réalisation avantageuse pour la construction du modèle aux étapes a) et b) du procédé au sens de l'invention, et
- la figure 5 représente schématiquement une installation pour la mise en oeuvre de l'invention.

[0023] On rappelle que la présente invention propose de calculer les fonctions de transfert au moyen d'un mo-

dèle mathématique basé sur une fonction F qui permet d'exprimer une fonction de transfert à partir de plusieurs paramètres d'entrée. Plus précisément, si la fonction de transfert cherchée est représentée sous la forme d'un

vecteur Y ($Y \in \mathbb{R}^n$, $n \in \mathbb{N}$) et si les paramètres d'entrée sont décrits sous la forme d'un vecteur

X ($X \in \mathbb{R}^m$, $m \in \mathbb{N}$), la fonction F définit la relation

suivante : $Y = F(X)$. En d'autres termes, la fonction F permet de déduire une fonction de transfert d'un ensemble donné de paramètres connus a priori. L'intérêt du modèle mathématique réside dans l'utilisation de paramètres d'entrée qu'il est facile d'acquérir pour un individu quelconque, en gardant à l'esprit toutefois que leur relation avec la fonction de transfert n'est pas forcément directe ou évidente. Le modèle mathématique doit en particulier être capable d'extraire l'information plus ou moins cachée dans les paramètres d'entrée afin d'en déduire la fonction de transfert cherchée. Le procédé de l'invention repose essentiellement sur deux points :

- la définition de la fonction F ,
- la détermination des paramètres d'entrée X .

Le modèle mathématique des HRTFs repose sur une fonction F permettant d'exprimer une HRTF à partir d'un nombre donné de paramètres d'entrée. Les paramètres d'entrée sont regroupés dans un vecteur

X ($X \in \mathbb{R}^m$, $m \in \mathbb{N}$) qui constitue donc le vecteur d'entrée de la fonction F . Le vecteur de sortie de la fonction est une HRTF qui est représentée par un vecteur

Y ($Y \in \mathbb{R}^n$, $n \in \mathbb{N}$). Par exemple, ce vecteur Y peut être constitué de coefficients fréquentiels décrivant le module du spectre de la fonction de transfert définie par la HRTF.

De manière équivalente, Y peut être constitué :

- de coefficients temporels décrivant la réponse impulsionnelle associée à la fonction de transfert définie par la HRTF,
- ou de coefficients fréquentiels décrivant le spectre complexe de la fonction de transfert définie par la HRTF.

La fonction F est donc une fonction de $\mathbb{R}^m$ dans $\mathbb{R}^n$. Le problème de la modélisation consiste à déterminer la fonction F , en association avec un jeu pertinent de paramètres (X), telle que toute HRTF (Y) soit solution de : $Y = F(X)$.

[0024] Spécifiquement pour l'estimation des HRTFs d'un individu, le vecteur d'entrée X du modèle contient principalement des informations relatives à :

- la direction dans laquelle on désire calculer une HRTF, préférentiellement sous la forme d'un angle d'azimut (θ) et d'un angle d'élévation (φ),
- et des paramètres "*individuels*" (tels que des HRTFs mesurées dans quelques directions de l'espace seulement, comme on le verra plus loin), ces paramètres individuels étant destinés à apporter au modèle des informations relatives aux spécificités de l'individu dont on désire calculer les HRTFs.

[0025] Le vecteur de sortie Y du modèle est constitué de coefficients associés à une représentation donnée d'une HRTF. Comme indiqué ci-avant, le vecteur Y peut correspondre aux coefficients fréquentiels décrivant le module du spectre d'une HRTF, mais d'autres représentations peuvent être considérées (analyse en composantes principales, filtre IIR, ou autres).

[0026] Ici, le modèle est appliqué à des fins d'interpolation. Un nombre réduit de HRTFs est mesuré sur un individu. Le modèle est alors utilisé pour calculer les HRTFs de cet individu dans toutes les directions couvrant la sphère 3D. Les HRTFs précédemment mesurées sont donc utilisées comme paramètres d'entrée du modèle. La modélisation consiste essentiellement à :

- déterminer la fonction F qui approche au mieux la relation entre X et Y,
- déterminer le jeu X de paramètres d'entrée le mieux adapté, en relation avec la fonction F, notamment en termes de qualité et de quantité des informations apportées par les paramètres et qui pourront être exploitées par le modèle utilisé.

La détermination de F et du vecteur X ne sont bien évidemment pas indépendants.

[0027] Il existe une grande variété de procédés mathématiques pour déterminer ces deux entités F et X. Le procédé de l'invention est préférentiellement basé sur des algorithmes d'apprentissage statistique et, dans un mode de réalisation préféré, sur des algorithmes du type à réseaux de neurones artificiels. Ces algorithmes sont brièvement présentés ci-après.

[0028] Les algorithmes d'apprentissage statistique sont des outils de prédiction des processus statistiques. Ils ont été utilisés avec succès pour la prédiction de processus pour lesquels on peut identifier plusieurs variables explicatives. Les réseaux de neurones artificiels définissent une catégorie particulière de ces algorithmes. L'intérêt des réseaux de neurones réside dans leur capacité à capter les dépendances de haut niveau, c'est-à-dire des dépendances qui impliquent plusieurs variables à la fois. La prédiction du processus tire profit de la connaissance et de l'exploitation de dépendances de haut niveau. Il existe une grande variété de domaines applicatifs des réseaux de neurones, notamment dans les techniques financières pour prédire les fluctuations de marché, en pharmacie, dans le domaine bancaire pour la détection des fraudes sur les cartes de cré-

dit, dans le marketing pour prévoir le comportement des consommateurs, ou autres. Les réseaux de neurones sont souvent considérés comme des prédicteurs universels, au sens où ils sont capables de prédire des données quelconques à partir de variables explicatives quelconques, dès lors que le nombre d'unités cachées est suffisant. En d'autres termes, ils permettent de modéliser une

fonction mathématique quelconque de $\mathfrak{R}^m$ dans $\mathfrak{R}^n$, si le nombre d'unités cachées m est suffisant.

[0029] En référence à la figure 1; un réseau de neurones se compose de trois couches : une couche d'entrée 10, une couche cachée 11 et une couche de sortie 12. La couche d'entrée 11 correspond aux variables explicatives, c'est-à-dire les variables d'entrée (le vecteur X précité), à partir desquelles s'effectue la prédiction, et qui sera décrit en détail plus loin. La couche de sortie 12 définit les valeurs prédites (le vecteur Y précité).

[0030] Dans la couche cachée, une première étape 111 consiste à calculer des combinaisons linéaires des variables explicatives de façon à combiner les informations provenant potentiellement de plusieurs variables. Une seconde étape 112 consiste à appliquer une transformation non linéaire (par exemple une fonction du type "*tangente hyperbolique*") à chacune des combinaisons linéaires afin d'obtenir les valeurs des unités cachées ou neurones qui constituent la couche cachée. Cette transformation non linéaire définit la fonction d'activation des neurones. Enfin, les unités cachées sont recombinaisons linéairement, à l'étape 113, afin de calculer la valeur prédite par le réseau de neurones.

[0031] Initialement, la mise au point d'un réseau de neurones passe par trois opérations :

- l'apprentissage, consistant à optimiser les paramètres de la couche cachée à partir d'une série d'exemples d'entraînement (formant ensemble d'apprentissage), à partir desquels le réseau de neurones cherche à minimiser son erreur de prédiction ;
- la procédure de validation, menée en parallèle de l'apprentissage et destinée à optimiser le nombre de couches cachées du réseau, afin que le réseau de neurones ne sur-apprenne pas l'ensemble d'apprentissage. Le réseau ne modélise que les relations fondamentales de dépendance et ne cherche pas à reproduire des relations qui ne sont dues qu'à des fluctuations statistiques de l'ensemble d'apprentissage. En complément de l'erreur d'apprentissage, une erreur de prédiction est ainsi évaluée sur des exemples issus d'un ensemble de validation, qui est distinct de l'ensemble d'apprentissage. Cette erreur définit l'erreur de validation. Elle commence par décroître lorsqu'on augmente le nombre de couches cachées, atteint un minimum, puis croît lorsque le nombre de couches cachées devient trop important. Le minimum définit donc un nombre optimal de couches cachées du réseau ;
- le calcul de l'erreur de prédiction finale, sur un tri-

sième ensemble de test, distinct des deux ensembles précédents.

[0032] Il existe différentes catégories de réseau de neurones se distinguant par leur architecture (type d'interconnexion entre neurones, choix des fonctions d'activation, ou autres) et le mode d'apprentissage utilisé.

[0033] Les réseaux de neurones ne sont pas utilisés qu'à des fins de prédiction. Ils sont également utilisés pour la classification et/ou le regroupement de données (en anglais "*clustering*") dans une perspective de réduction d'information. En effet, un réseau de neurones est capable, dans un ensemble de données, d'identifier des caractéristiques communes entre les éléments de cet ensemble, pour les regrouper ensuite en fonction de leur ressemblance. A chaque groupe ainsi constitué est associé ensuite un élément représentatif de l'information contenue dans le groupe, dit "*représentant*". On peut alors substituer ce représentant à l'ensemble du groupe. L'ensemble des données peut ainsi être décrit au moyen d'un nombre réduit d'éléments, ce qui constitue une réduction de données. Les cartes de Kohonen ou cartes auto-organisatrices (en anglais SOM pour "*Self Organizing Map*") peuvent être des réseaux de neurones dédiés à cette tâche de regroupement.

[0034] Une question s'était posée par rapport au choix des directions des HRTFs à mesurer pour mener l'étape c) décrite ci-avant.

[0035] La méthode qui semblait la plus immédiate consistait en une sélection uniforme dans laquelle on choisissait un sous-ensemble de directions en cherchant à couvrir de la manière la plus homogène et régulière possible l'ensemble de la sphère 3D. Cette méthode reposait sur un échantillonnage régulier de la sphère 3D. Or, il s'est avéré que les HRTFs ne variaient pas de façon uniforme en fonction de la direction. De ce point de vue, une sélection uniforme des HRTFs n'était pas réellement efficace.

[0036] Une méthode plus prometteuse consistait à appliquer la technique de regroupement précitée ("*clustering*") afin d'identifier les directions des HRTFs les plus "*pertinentes*", c'est-à-dire les mieux représentatives des caractéristiques des HRTFs observées sur l'ensemble de la sphère 3D. Lorsqu'elle est appliquée à la détermination des HRTFs d'un individu, cette technique de regroupement peut consister :

- dans une première étape, à identifier les redondances entre les HRTFs de directions voisines,
- dans une seconde étape, à regrouper les HRTFs selon un critère de ressemblance,
- dans une troisième étape, l'ensemble de la sphère 3D entourant l'auditeur est ainsi subdivisé en un nombre réduit de zones qui correspondent aux différents groupes de HRTFs précédemment identifiés, et
- dans une quatrième étape, à chaque groupe est associée une HRTF qui est considérée comme le re-

présentant du groupe.

Cette HRTF "*représentante*" est l'une des HRTFs du regroupement et elle est sélectionnée comme la HRTF minimisant un critère de distance avec toutes les autres HRTFs du groupe. La HRTF représentante contient l'essentiel de l'information des HRTFs du groupe. Au final, l'ensemble des HRTFs représentantes ainsi obtenues constitue une description compacte des propriétés des HRTFs pour toute la sphère 3D.

[0037] Cette technique avait donné de bons résultats quant au modèle. Le premier résultat est une réduction de données. La procédure de clustering apporte en outre une information supplémentaire quant aux directions associées aux HRTFs représentantes, cette information permettant de définir une sélection de HRTFs destinées à alimenter l'entrée du modèle de calcul de HRTFs. Cette sélection est a priori non uniforme, mais plus efficace, et garantit une meilleure "*représentativité*" de l'ensemble de la sphère 3D.

[0038] Néanmoins, il est apparu aux inventeurs que cette étape de regroupement n'était pas nécessaire et qu'en fait, quelques directions de mesure de HRTFs pouvaient être choisies initialement, de façon arbitraire, sans que le modèle ne soit faussé ou que ses performances soient moindres pour autant. Un avantage considérable alors est que ces directions peuvent être choisies librement selon des conditions de mesures préférées qui seront décrites en détail plus loin.

Ainsi, la présente invention propose l'utilisation, comme paramètres d'entrée du modèle, d'une sélection de HRTFs correspondant à des directions quelconques en ce sens que ces directions ne sont pas nécessairement "*représentatives*" (au sens de la technique de clustering exposée ci-avant). Toutefois, ces directions restent exploitables en ce sens que le modèle est capable d'extraire des informations spécifiques relatives à chaque individu.

[0039] Préférentiellement, l'invention utilise des algorithmes d'apprentissage statistique de type "*réseau de neurones artificiels*", en tant qu'outil de modélisation pour le calcul des HRTFs (par exemple avec un réseau de neurone de type "*Multi Layer Perceptron*" ou MLP). Les paramètres d'entrée du réseau de neurones sont au minimum l'angle d'azimut (θ_1) et d'élévation (φ_1) spécifiant la direction d'une HRTF à calculer. Ces paramètres sont éventuellement complétés par des paramètres "*individuels*" associés à l'individu dont on veut calculer les HRTFs. Ces paramètres individuels comportent une sélection de HRTFs de l'individu qui ont été mesurées au préalable. Néanmoins, il n'est pas exclu d'ajouter en entrée du modèle des paramètres morphologiques de l'individu pour enrichir l'information à fournir au modèle.

[0040] Les paramètres de sortie du modèle sont alors les coefficients du vecteur décrivant la HRTF pour la direction (θ_1 , φ_1) et pour l'individu spécifié en entrée.

[0041] En se référant à nouveau à la figure 1, le principe du calcul des HRTFs par la mise en oeuvre d'un réseau de neurones artificiels (par exemple de type MLP)

se compose :

- de la couche d'entrée 10 constituée des paramètres d'entrée incluant alors :
 - les HRTFs déjà mesurées pour quelques directions de l'espace seulement et notées $HRTF(\varphi_i^{mes}, \theta_i^{mes})$, avec i compris entre 1 et n ,
 - les directions pour lesquelles on veut calculer les HRTFs, spécifiées préférentiellement sous la forme d'un angle d'élévation (φ_j^{cal}) et un angle d'azimut (θ_j^{cal}), avec j compris entre 1 et N , N étant beaucoup plus grand que n ,
- de la couche de sortie 12 donnant les HRTFs de l'individu dans les directions ($\varphi_j^{cal}, \theta_j^{cal}$) spécifiées en entrée, et
- d'une ou plusieurs couches cachées 11 qui vont chercher, en ajustant les poids et les fonctions d'activation des neurones, à modéliser au mieux les relations entre la couche d'entrée et la couche de sortie.

[0042] En se référant maintenant à la figure 2, la mise en oeuvre d'un réseau de neurones passe par trois étapes :

- la phase d'apprentissage 21,
- la phase de validation 22, et
- la phase de test 23.

Pour mener à bien ces trois phases, on dispose initialement d'une base de données 20 de HRTFs collectées sur un ou plusieurs individus. Ainsi, on comprendra qu'une étape préalable de collecte de mesures de HRTFs de plusieurs individus dans toutes les directions de l'espace est mise en oeuvre. C'est ainsi que l'on constitue la base de données 20.

Cette base de données 20 est décomposée en trois ensembles distincts :

- un ensemble d'apprentissage (APPR),
- un ensemble de validation (VALID),
- un ensemble de test (TEST).

Pour la phase d'apprentissage 21, on dispose de couples combinant :

- un vecteur d'entrée X (décrivant la direction de la HRTF à calculer et les paramètres individuels tels que la mesure des HRTFs dans quelques directions),
- et un vecteur de sortie Y (correspondant à la HRTF que doit estimer au mieux le réseau de neurones).

[0043] L'apprentissage consiste, pour chaque couple ainsi formé issu de l'ensemble d'apprentissage :

- à optimiser le réseau de neurones (en termes des poids et des fonctions d'activation des neurones),
- et à comparer le résultat obtenu par le réseau de neurones et le résultat attendu (HRTF mesurée sur l'individu), de façon à minimiser un critère d'erreur donné.

[0044] Un risque de la phase d'apprentissage est le sur-apprentissage qui se traduit comme suit : le réseau de neurones apprend "*par coeur*" l'ensemble d'apprentissage et cherche à reproduire des variations propres à l'ensemble d'apprentissage, alors qu'elles n'existent pas au niveau global. Pour éviter le sur-apprentissage, la phase de validation 22 est menée conjointement à la phase d'apprentissage 21. En se référant à la figure 3, elle consiste à évaluer l'erreur de prédiction du réseau de neurones sur un ensemble de validation (distinct de l'ensemble d'apprentissage), ce qui définit l'erreur de validation Err_{valid} commence par décroître, puis se remet à croître lorsque le sur-apprentissage se manifeste. Le minimum MIN de l'erreur de validation détermine donc la fin de l'apprentissage.

[0045] En fait, cette observation retentit directement sur le nombre de HRTFs mesurées à fournir en entrée du modèle, après la phase d'apprentissage, c'est-à-dire à l'étape c) décrite ci-avant. En effet, plus ce nombre de mesures est faible et moins le modèle dispose d'informations pour calculer les HRTFs : l'erreur de validation est élevée. Toutefois, plus ce nombre de mesures est élevé et plus le risque de sur-apprentissage est élevé. On retiendra donc qu'une caractéristique optionnelle avantageuse du procédé au sens de l'invention prévoit, à l'étape d'apprentissage b), de déterminer un nombre optimum N_{opt} (figure 3) de HRTFs mesurées (Nb_HRTF^{mes}) à fournir en entrée du modèle pour la mise en oeuvre de l'étape c).

[0046] La phase de test est menée une fois l'apprentissage terminé et consiste à évaluer l'erreur de prédiction sur l'ensemble de test. Cette erreur, dite "*erreur de test*", décrit finalement les performances finales du réseau de neurones.

[0047] A l'issue de ces trois phases, on dispose d'un réseau de neurones opérationnel, auquel il suffit de soumettre des paramètres d'entrée pour obtenir les HRTFs d'un individu dans une direction.

[0048] Ainsi, en référence à la figure 4a, le procédé au sens général de l'invention comporte donc une étape a) au cours de laquelle on constitue une base de données 20 en mesurant une pluralité de HRTFs dans une multiplicité de directions de l'espace et pour une pluralité d'individus. Cette étape de mesure référencée 40 sur la figure 4a consiste à collecter les mesures de HRTFs dans N directions de l'espace, pour plusieurs individus préférentiellement de morphologie (ou de "*morphotype*") différente, pour obtenir une base de données exhaustives selon les spécificités des individus. Plus généralement, plus le nombre d'individus pris en compte lors de l'ap-

prentissage est élevé et meilleures sont les performances du réseau de neurones, notamment en termes "d'universalité".

[0049] L'étape b) suivante consiste en l'apprentissage du modèle en utilisant la base de données 20. On choisit arbitrairement, à l'étape 41, des directions i de mesures représentatives de HRTFs en nombre restreint n (avec $n < N$). Cette étape 41 sera décrite en détail plus loin, en référence à la figure 4c. On mène ensuite les trois phases d'apprentissage 21, de validation 22 et de test 23, pour construire le modèle à l'étape 44. On notera qu'il est possible d'ajuster le nombre restreint de mesures n pour éviter le phénomène de sur-apprentissage décrit ci-avant. Ainsi, on peut déterminer un nombre optimum N_{opt} de mesures nécessaires au bon fonctionnement du modèle (étape 42) et adopter ce nombre optimum (étape 43) pour la définition du modèle. On obtient finalement le réseau de neurones 44 pour le calcul des HRTFs. Le réseau de neurones 44 est alors capable de calculer les HRTFs de n'importe quel individu, dans n'importe quelle direction, dès lors qu'on dispose de quelques HRTFs de l'individu dans les directions prédéterminées $\varphi_i^{mes}, \theta_i^{mes}$.

[0050] Une fois le modèle construit (étape 44), on peut, au cours d'une étape c) ultérieure, déterminer les HRTFs d'un individu quelconque dans toutes les directions de l'espace. Ainsi, en référence à la figure 4b :

- c1) on mesure les HRTFs de l'individu dans les directions i de mesures ($HRTF(\varphi_i^{mes}, \theta_i^{mes})$) et l'on indique au modèle les directions dans lesquelles on souhaite un calcul de HRTFs ($\varphi_j^{cal}, \theta_j^{cal}$), dans une étape 45,
- c2) on applique ensuite le modèle 44 à ces mesures de HRTFs, et
- c3) on obtient les HRTFs de l'individu, calculées dans les directions souhaitées $\varphi_j^{cal}, \theta_j^{cal}$ (étape 46).

[0051] Toutefois, on rappelle que les conditions de mesure de l'étape c1) doivent être sensiblement reproductibles avec les conditions de mesure pour des HRTFs dans les directions i (étape 41 de la figure 4a).

[0052] En référence à la figure 4c, on précise maintenant un aspect optionnel de l'invention pour une réalisation préférée de l'apprentissage du modèle. En fait, la base de données 20 doit être constituée dans des conditions les plus classiques et les plus standard pour offrir, en sortie du modèle, des HRTFs de qualité qui pourront être appliquées à des dispositifs de restitution en offrant un confort d'écoute satisfaisant. En revanche, il est préférentiellement effectué un second type de mesures, parallèlement à la constitution de la base de données 20, dans des conditions qui peuvent être différentes, voire "dégradées", et dans un nombre restreint de directions. Ce second type de mesures est effectué sur les mêmes individus que ceux sur lesquels les mesures constituant la base de données 20 ont été menées. Ces mesures "dégradées" sont notées $HRTF(\varphi_i^{mes}, \theta_i^{mes})$ et effectuées à une étape 48 de la figure 4c.

[0053] Ensuite, au cours d'une étape 49, on précise, en entrée du modèle, dans quelles directions ($\varphi_j^{cal}, \theta_j^{cal}$) les HRTFs devront être calculées par le modèle. Préférentiellement, il s'agira bien entendu du plus grand nombre possible de directions de l'espace 3D. Une version du modèle 44b, en état d'apprentissage, calcule les HRTFs dans ces directions ($\varphi_j^{cal}, \theta_j^{cal}$) à partir des jeux de mesures "dégradées" $HRTF(\varphi_i^{mes}, \theta_i^{mes})$, dans une étape suivante 46b. Le modèle compare ces HRTFs calculées aux HRTFs de la base de données 20 dans les mêmes directions ($\varphi_j^{cal}, \theta_j^{cal}$). Si l'écart est jugé trop grand (flèche n), le modèle en apprentissage 44b est perfectionné jusqu'à ce que cet écart soit réduit à une erreur acceptable (flèche o) : le modèle devient alors définitif (étape de fin 44).

[0054] On retiendra donc qu'à l'étape a), parallèlement à la constitution de la base de données 20 pour une pluralité d'individus, on mesure en outre, sur cette même pluralité d'individus, des jeux respectifs de fonctions représentatives des HRTFs (notées $HRTF(\varphi_i^{mes}, \theta_i^{mes})$) dans les conditions et directions de mesure fixées arbitrairement. Pour la construction du modèle à l'étape b) :

- on applique alors en entrée du modèle ces jeux respectifs de mesure $HRTF(\varphi_i^{mes}, \theta_i^{mes})$, et
- on applique en sortie du modèle la base de données 20 pour une comparaison des HRTFs calculées avec celles de la base de données.

[0055] Bien entendu, cette réalisation optionnelle de la figure 4c est avantageuse en particulier si les mesures $HRTF(\varphi_i^{mes}, \theta_i^{mes})$ sont réellement dégradées par rapport à celles qui ont permis de constituer la base de données 20. On rappelle encore que ces conditions de mesures de $HRTF(\varphi_i^{mes}, \theta_i^{mes})$ doivent être sensiblement les mêmes que celles de l'étape c1) menées sur un individu quelconque.

[0056] En référence à la figure 5, on décrit maintenant un exemple de réalisation de ces conditions de mesure. L'individu IND est placé dans une cabine CAB qui n'est pas nécessairement anéchoïque. Il dispose d'un casque CAS portant au moins un microphone MIC accolé contre l'une de ses oreilles. Préférentiellement, le casque CAS est porté par une tige rigide télescopique en hauteur (le long de l'axe y). Cette tige est par ailleurs fixée à un repère REP1 de la cabine CAB. Cette réalisation permet de maintenir immobile l'individu IND (par rapport aux autres axes x et z) et de le positionner correctement par rapport au repère REP1 et, par conséquent, par rapport aux sources sonores $S1, S2, \dots, S_n$ de la cabine CAB. Par ailleurs, un autre repère REP2, tel qu'un repère visuel sur un miroir, permet à l'individu de se positionner en hauteur (le long de l'axe y). Typiquement, l'individu peut être assis sur un siège ajustable en hauteur et régler cette hauteur jusqu'à ce que ses oreilles coïncident avec le repère REP2, sur le miroir.

[0057] On comprend déjà que l'un des avantages de la mise en oeuvre de l'invention est d'éviter la technique

de regroupement (*clustering*) et de laisser un libre choix au niveau de l'emplacement des sources sonores S1-Sn. Par exemple, on peut disposer ces sources ailleurs qu'au niveau du miroir portant le repère REP2, ou encore ailleurs qu'au niveau de la base de la tige REP1. Typiquement, dans l'exemple de la figure 5, la source S2 est légèrement décalée par rapport au repère REP1.

[0058] Le nombre de sources S1-Sn à prévoir dépend, en principe, du nombre de HRTFs que l'on souhaite calculer à partir du modèle. Typiquement, pour calculer des HRTFs dans tout l'espace 3D, entre 25 et 30 directions de mesures préalables dans la cabine CAB sont recommandées. Néanmoins, pour un confort d'écoute satisfaisant, une quinzaine de mesures est suffisante.

[0059] Enfin, dans l'absolu, une unique mesure devrait suffire pour obtenir une unique HRTF estimée. On choisira alors la direction de mesure la plus proche de la direction de HRTF à calculer.

[0060] De façon plus générale, on retiendra que le nombre optimum de directions de mesure, et donc le nombre de mesures Nopt (figure 3), sont de l'ordre d'une vingtaine.

[0061] On indique en outre qu'entre 700 et 1000 directions de mesure (par oreille) sont habituellement nécessaires pour obtenir une bonne base de données des HRTFs d'un individu, selon la technique de l'art antérieur. On appréciera alors la réduction du nombre de mesures utiles, au sens de l'invention.

[0062] On constatera en outre sur la figure 5 que les sources S1 à Sn ne sont pas nécessairement disposées sur une même surface de portion de sphère. En effet, le but du protocole de mesure de la figure 5 n'est pas d'obtenir des HRTFs au sens strict du terme, mais plus exactement des fonctions de transfert d'un individu, ces fonctions de transfert étant partiellement représentatives de ses HRTFs. Ces fonctions de transfert sont destinées à être utilisées comme paramètres d'entrée du modèle 44. Les inventeurs ont en effet constaté que le modèle était capable d'extraire et d'exploiter l'information individuelle contenue dans ces fonctions de transfert, même si cette information est partielle ou brouillée. Ce qui importe n'est pas la qualité des HRTFs mesurées selon ce protocole, mais leur reproductibilité. C'est essentiellement sur cette reproductibilité que se base le modèle de HRTFs. Un avantage qu'offre ce protocole de mesure est de relâcher les contraintes de la procédure de mesure, sans pour autant affecter le bon fonctionnement du modèle.

[0063] On retiendra donc que, dans l'installation telle que représentée sur la figure 5, les sources sonores S1-Sn prévues dans la cabine CAB peuvent être dans des positions respectives appartenant à des surfaces de sphère distinctes.

[0064] On comprendra en outre que les mesures appliquées en entrée du modèle ne sont pas forcément des HRTFs réelles, mais des fonctions de transfert représentatives de HRTFs. Par ailleurs, ces fonctions de transfert présentées à l'entrée du modèle peuvent avoir des formes variées (correspondant à différentes représenta-

tions de HRTFs), notamment :

- un spectre complexe de la fonction de transfert,
- un module du spectre de la fonction de transfert,
- une phase du spectre de la fonction de transfert,
- une réponse impulsionnelle associée à la fonction de transfert,
- ou une combinaison de ces différents éléments.

[0065] On indique en outre qu'au moins un paramètre complémentaire qui peut être fourni en entrée du modèle peut être de nature morphologique et propre à l'individu IND, tel que la distance entre ses deux oreilles. Dans ce cas, les phases d'apprentissage, de validation et de test du réseau de neurones sont réalisées à partir d'une base de données comportant, en complément des HRTFs, des paramètres morphologiques des individus, tels que :

- la distance entre les oreilles, comme indiqué ci-avant,
- et/ou une position et/ou une forme des pavillons des oreilles de l'individu,
- et/ou des dimensions d'ellipsoïdes représentant sa tête et/ou son torse,
- et/ou les dimensions d'un cylindre représentant son cou.

[0066] En se référant à nouveau à la figure 5, les signaux mesurés par le microphone MIC sont recueillis par une interface 51 d'une unité centrale UC (par exemple une carte d'acquisition audio), qui les convertit en données numériques. Ces données, le cas échéant enrichies par une mesure du ou des paramètre(s) morphologique(s) de l'individu, sont ensuite traitées par le modèle 44 au sens de l'invention. Le modèle 44 peut être stocké sous la forme d'un produit programme d'ordinateur dans une mémoire de l'unité centrale UC. Les HRTFs calculées pour toutes les directions de l'espace que donne le modèle peuvent ensuite être stockées en mémoire 52 ou enregistrées sur un support amovible (sur disquette ou gravées sur CD-ROM) ou encore communiquées via un réseau tel qu'Internet ou équivalent.

[0067] Ainsi, dans cette mise en oeuvre avantageuse, la couche d'entrée du réseau de neurones comporte une sélection de HRTFs de l'individu correspondant à des directions quelconques, mais fixées a priori, et obtenues dans des conditions non idéales. Ces HRTFs "approximatives" sont certes obtenues par la mesure directe sur l'individu IND, mais dans des conditions non idéales, notamment dans un environnement qui n'est pas nécessairement anéchoïque. Toutefois, le protocole de mesure doit être défini au préalable (typiquement à l'étape d'apprentissage b)) et doit être suivi rigoureusement à l'étape c) d'application du modèle à un individu quelconque. Le réseau de neurones ainsi obtenu est capable de calculer les HRTFs de n'importe quel individu, dans n'importe quelle direction, dès lors qu'on dispose des mesures dans les directions φ_i^{mes} et θ_i^{mes} choisies et obtenues

sous ces conditions prédéfinies.

[0068] Bien entendu, la présente invention ne se limite pas à la forme de réalisation décrite ci-avant à titre d'exemple ; elle s'étend à d'autres variantes.

[0069] Par exemple, au lieu de prévoir une pluralité de sources sonores S1-Sn dans la cabine décrite en référence à la figure 5, on peut prévoir en variante une source unique qui se déplace entre des positions S1 à Sn.

Revendications

1. Procédé de modélisation de fonctions de transfert HRTFs propres à un individu, dans lequel :

a) on constitue une base de données incluant une pluralité de HRTFs suivant une multiplicité de directions de l'espace et pour une pluralité d'individus,

b) par apprentissage sur ladite base de données, on construit un modèle propre à donner des HRTFs pour ladite multiplicité de directions, à partir d'un jeu de mesures représentatives de HRTFs dans des directions respectives sélectionnées parmi ladite multiplicité de directions, et

c) pour un individu quelconque :

c1) on mesure un jeu de fonctions représentatives des HRTFs de l'individu dans lesdites directions sélectionnées seulement,

c2) on applique le modèle auxdites mesures dans les directions sélectionnées, et

c3) on obtient les HRTFs de l'individu dans toute ladite multiplicité de directions,

et dans lequel :

- les conditions et les directions de mesure pour obtenir ledit jeu de mesures sont fixées arbitrairement pendant l'étape d'apprentissage b), et
- on applique, à l'étape c1), des conditions de mesure sensiblement reproductibles avec les conditions de mesure de l'étape b).

2. Procédé selon la revendication 1, dans lequel, à l'étape a), parallèlement à la constitution de ladite base de données pour ladite pluralité d'individus, on mesure en outre, sur ladite pluralité d'individus, des jeux respectifs de fonctions représentatives des HRTFs dans lesdites conditions et directions de mesure fixées arbitrairement, et, pour la construction du modèle à l'étape b) :

- on applique en entrée du modèle lesdits jeux respectifs, et
- on applique en sortie du modèle ladite base de

données.

3. Procédé selon l'une des revendications 1 et 2, dans lequel le modèle est construit en mettant en oeuvre un réseau de neurones artificiels.

4. Procédé selon la revendication 3, dans lequel l'étape b) comporte :

- une phase d'apprentissage,
- une phase de validation menée parallèlement avec la phase d'apprentissage, et
- une phase de test,

et dans lequel, pendant la phase de validation, on détermine un nombre optimum (N_{opt}) de mesures à fournir en entrée du modèle pour la mise en oeuvre de l'étape c), en vue de limiter un effet de sur-apprentissage du modèle.

5. Procédé selon la revendication 4, dans lequel le nombre optimum (N_{opt}) est de l'ordre de vingt.

6. Procédé selon l'une des revendications précédentes, dans lequel le modèle utilise en outre au moins un paramètre morphologique caractérisant un individu, et dans lequel, à l'étape c2), on fournit en outre au modèle une mesure dudit paramètre morphologique.

7. Procédé selon l'une des revendications précédentes, dans lequel, à l'étape c2), on fournit en entrée du modèle :

- le jeu de mesures dans lesdites directions sélectionnées, et
- au moins une direction (ϕ_j^{cal} , θ_j^{cal}) parmi ladite multiplicité de directions dans laquelle on souhaite une estimation de HRTFs.

8. Installation pour l'estimation de fonctions de transfert HRTFs propres à un individu, comportant :

- une cabine de mesure de fonctions de transfert représentatives de HRTFs dans un jeu de directions choisies, et
- une unité de traitement (UC) pour récupérer un jeu de mesures sur un individu dans lesdites directions choisies et évaluer les HRTFs de l'individu dans une multiplicité de directions de l'espace incluant lesdites directions choisies, à partir d'un modèle capable de donner des HRTFs pour une multiplicité de directions, à partir d'un jeu de mesures représentatives de HRTFs dans quelques directions seulement fixées arbitrairement parmi ladite multiplicité de directions,

dans laquelle les directions de mesure dans ladite

cabine correspondent auxdites directions fixées arbitrairement, **caractérisée en ce que**, la cabine comportant des repères (REP1, REP2) sur des axes respectifs (x, y) dont l'intersection définit une position des oreilles de l'individu (IND) dans la cabine (CAB), des sources sonores (S1-Sn), prévues dans ladite cabine (CAB) pour effectuer lesdites mesures, sont placées à des distances différentes de ladite intersection.

9. Produit programme d'ordinateur, destiné à être stocké dans une mémoire d'une unité de traitement ou sur un support amovible propre à coopérer avec un lecteur de ladite unité de traitement, ou destiné à être transmis d'un serveur vers ladite unité de traitement, comportant des instructions sous forme de code informatique pour construire un modèle basé sur un réseau de neurones artificiels et capable de donner des fonctions de transfert HRTFs d'un individu pour une multiplicité de directions à partir d'un jeu de mesures, effectuées sur cet individu, représentatives de HRTFs dans quelques directions seulement, et fixées arbitrairement parmi ladite multiplicité de directions, le programme mettant en oeuvre, à partir d'une base de données incluant une pluralité de HRTFs suivant une multiplicité de directions de l'espace et pour une pluralité d'individus, au moins une phase d'apprentissage.

10. Produit programme d'ordinateur, destiné à être stocké dans une mémoire d'une unité de traitement ou sur un support amovible propre à coopérer avec un lecteur de ladite unité de traitement, ou destiné à être transmis d'un serveur vers ladite unité de traitement, comportant des instructions sous forme de code informatique pour mettre en oeuvre un modèle basé sur un réseau de neurones artificiels et capable de donner des fonctions de transfert HRTFs d'un individu pour une multiplicité de directions, à partir d'un jeu de mesures effectuées sur cet individu, représentatives de HRTFs dans quelques directions seulement, et fixées arbitrairement parmi ladite multiplicité de directions, les mesures effectuées sur l'individu étant réalisées dans une cabine dans laquelle :

- les directions de mesure correspondent auxdites directions fixées arbitrairement, et
- la cabine comportant des repères (REP1, REP2) sur des axes respectifs (x, y) dont l'intersection définit une position des oreilles de l'individu (IND) dans la cabine (CAB) pour effectuer lesdites mesures, des sources sonores (S1-Sn), prévues dans ladite cabine (CAB), sont placées à des distances différentes de ladite intersection.

Patentansprüche

1. Verfahren zur Modellierung von einer Person eigenen HRTFs-Übertragungsfunktionen, wobei:

- a) eine Datenbank erstellt wird, die eine Vielzahl von HRTFs gemäß einer Vielfalt von Richtungen des Raums und für eine Vielzahl von Personen umfasst,
- b) durch Lernen in der Datenbank ein Modell konstruiert wird, das HRTFs für die Vielfalt von Richtungen angeben kann, ausgehend von einem Satz von Messwerten, die für HRTFs in aus der Vielfalt von Richtungen ausgewählten Richtungen repräsentativ sind, und
- c) für eine beliebige Person:

- c1) ein Satz von Funktionen gemessen wird, die für die HRTFs der Person nur in den ausgewählten Richtungen repräsentativ sind,
- c2) das Modell an die Messwerte in den ausgewählten Richtungen angewendet wird, und
- c3) die HRTFs der Person in der ganzen Vielfalt von Richtungen erhalten werden,

und wobei:

- die Messbedingungen und -richtungen zum Erhalt des Satzes von Messwerten willkürlich während des Lernschritts b) festgelegt werden, und
- im Schritt c1) Messbedingungen angewendet werden, die im Wesentlichen mit den Messbedingungen des Schritts b) reproduzierbar sind.

2. Verfahren nach Anspruch 1, wobei im Schritt a) parallel zur Erstellung der Datenbank für die Vielzahl von Personen außerdem an der Vielzahl von Personen jeweilige Sätze von Funktionen gemessen werden, die für die HRTFs unter den willkürlich festgelegten Messbedingungen und -richtungen repräsentativ sind, und für die Konstruktion des Modells im Schritt b):

- am Eingang des Modells die jeweiligen Sätze angewendet werden, und
- am Ausgang des Modells die Datenbank angewendet wird.

3. Verfahren nach einem der Ansprüche 1 und 2, wobei das Modell unter Verwendung eines Netzwerks aus künstlichen Neuronen aufgebaut wird.

4. Verfahren nach Anspruch 3, wobei der Schritt b) aufweist:

- eine Lernphase,

- eine Validierungsphase, die parallel zur Lernphase ausgeführt wird, und
- eine Testphase,

und wobei während der Validierungsphase eine optimale Anzahl (Nopt) von Messungen bestimmt wird, die am Eingang des Modells zur Durchführung des Schritts c) zu liefern sind, um eine Wirkung des Überlernens des Modells zu begrenzen.

5. Verfahren nach Anspruch 4, wobei die optimale Anzahl (Nopt) in der Größenordnung von zwanzig liegt.
6. Verfahren nach einem der vorhergehenden Ansprüche, wobei das Modell außerdem mindestens einen eine Person kennzeichnenden morphologischen Parameter verwendet, und wobei im Schritt c2) außerdem ein Messwert des morphologischen Parameters an das Modell geliefert wird.
7. Verfahren nach einem der vorhergehenden Ansprüche, wobei im Schritt c2) am Eingang des Modells geliefert wird:

- der Satz von Messwerten in den ausgewählten Richtungen, und
- mindestens eine Richtung (φ_j^{cal} , θ_j^{cal}) aus der Vielfalt von Richtungen, in der eine Schätzung von HRTFs gewünscht wird.

8. Anlage zur Schätzung von einer Person eigenen HRTFs-Übertragungsfunktionen, die aufweist:

- eine Messkabine für repräsentative Übertragungsfunktionen von HRTFs in einem Satz von ausgewählten Richtungen, und
- eine Verarbeitungseinheit (UC), um einen Satz von Messwerten an einer Person in den ausgewählten Richtungen wiederzugewinnen und die HRTFs der Person in einer Vielfalt von Richtungen des Raums zu bewerten, die die ausgewählten Richtungen umfassen, ausgehend von einem Modell, das fähig ist, HRTFs für eine Vielfalt von Richtungen ausgehend von einem Satz von Messwerten anzugeben, die für HRTFs nur in einigen Richtungen repräsentativ sind, die willkürlich unter der Vielfalt von Richtungen festgelegt werden,

wobei die Messrichtungen in der Kabine den willkürlich festgelegten Richtungen entsprechen, **dadurch gekennzeichnet, dass**, da die Kabine Bezugspunkte (REP1, REP2) auf Achsen (x, y) aufweist, deren Schnittstelle eine Stellung der Ohren der Person (IND) in der Kabine (CAB) definiert, Schallquellen (S1-Sn), die in der Kabine (CAB) zur Durchführung der Messungen vorgesehen sind, in unterschiedlichen Abständen von der Schnittstelle platziert sind.

9. Computerprogrammprodukt, das dazu bestimmt ist, in einem Speicher einer Verarbeitungseinheit oder auf einem entfernbaren Träger gespeichert zu werden, der mit einem Lesegerät der Verarbeitungseinheit zusammenwirken kann, oder dazu bestimmt ist, von einem Server zur Verarbeitungseinheit übertragen zu werden, das Anweisungen in Form eines Computercodes aufweist, um ein Modell zu konstruieren, das auf einem Netzwerk aus künstlichen Neuronen basiert und fähig ist, HRTFs-Übertragungsfunktionen einer Person für eine Vielfalt von Richtungen ausgehend von einem Satz von an dieser Person ausgeführten Messungen anzugeben, die für HRTFs nur in einigen Richtungen repräsentativ sind und willkürlich unter der Vielfalt von Richtungen festgelegt werden, wobei das Programm ausgehend von einer Datenbank, die eine Vielzahl HRTFs gemäß einer Vielfalt von Richtungen des Raums und für eine Vielzahl von Personen umfasst, mindestens eine Lernphase durchführt.

10. Computerprogrammprodukt, das dazu bestimmt ist, in einem Speicher einer Verarbeitungseinheit oder auf einem entfernbaren Träger gespeichert zu werden, der mit einem Lesegerät der Verarbeitungseinheit zusammenwirken kann, oder das dazu bestimmt ist, von einem Server zur Verarbeitungseinheit übertragen zu werden, das Anweisungen in Form eines Computercodes aufweist, um ein Modell anzuwenden, das auf einem Netzwerk aus künstlichen Neuronen basiert und fähig ist, HRTFs-Übertragungsfunktionen einer Person für eine Vielfalt von Richtungen ausgehend von einem Satz von an dieser Person ausgeführten Messungen anzugeben, die für HRTFs nur in einigen Richtungen repräsentativ sind und willkürlich unter der Vielfalt von Richtungen festgelegt werden, wobei die an der Person ausgeführten Messungen in einer Kabine durchgeführt werden, in der:

- die Messrichtungen den willkürlich festgelegten Richtungen entsprechen, und
- die Kabine Bezugspunkte (REP1, REP2) auf jeweiligen Achsen (x, y) aufweist, deren Schnittstelle eine Stellung der Ohren der Person (IND) in der Kabine (CAB) definiert, um die Messungen durchzuführen, wobei in der Kabine (CAB) vorgesehene Schallquellen (S1-Sn) in unterschiedlichen Abständen zur Schnittstelle platziert sind.

Claims

1. Method of modeling head-related transfer functions HRTFs specific to an individual, in which:

- a) a database is constructed including a plurality

of HRTFs in a multiplicity of directions in space and for a plurality of individuals,

b) by learning from said database, a specific model is constructed to give HRTFs for said multiplicity of directions, based on a series of measurements representative of HRTFs in respective directions selected from said multiplicity of directions, and

c) for any individual:

c1) a series of functions representative of the HRTFs of the individual only in said selected directions is measured,

c2) the model is applied to said measurements in the selected directions, and

c3) the HRTFs of the individual are obtained in all said multiplicity of directions,

and in which:

- the measurement conditions and directions to obtain said series of measurements are arbitrarily fixed during the learning step b), and
- measurement conditions roughly reproducible with the measurement conditions of the step b) are applied in the step c1).

2. Method according to Claim 1, in which, in the step a), parallel to the construction of said database for said plurality of individuals, respective sets of functions representative of the HRTFs are also measured, on said plurality of individuals, in said arbitrarily fixed measurement conditions and directions, and for the construction of the model in the step b):

- said respective sets are applied as input for the model, and
- said database is applied as output for the model.

3. Method according to one of Claims 1 and 2, in which the model is constructed by setting up an artificial neural network.

4. Method according to Claim 3, in which the step b) comprises:

- a learning phase,
- a validation phase conducted in parallel with the learning phase, and
- a test phase,

and in which, during the validation phase, an optimum number (Nopt) of measurements to be supplied as input for the model is determined for implementation of the step c), in order to limit an over-learning effect of the model.

5. Method according to Claim 4, in which the optimum number (Nopt) is around twenty.

6. Method according to one of the preceding claims, in which the model also uses at least one morphological parameter characterizing an individual, and in which, in the step c2), a measurement of said morphological parameter is also supplied to the model.

7. Method according to one of the preceding claims, in which, in the step c2), the model has supplied to it as input:

- the series of measurements in said selected directions, and
- at least one direction $(\varphi_j^{cal}, \theta_j^{cal})$ out of said multiplicity of directions in which an estimation of HRTFs is desired.

8. Installation for estimating head-related transfer functions HRTFs specific to an individual, comprising:

- a booth for measuring transfer functions representative of HRTFs in a set of chosen directions, and
- a processing unit (CPU) for recovering a series of measurements on an individual in said chosen directions and evaluating the HRTFs of the individual in a multiplicity of directions in space including said chosen directions, based on a model capable of giving HRTFs for the multiplicity of directions, based on a series of measurements representative of HRTFs only in a few arbitrarily fixed directions of said multiplicity of directions,

in which the measurement directions in said booth correspond to said arbitrarily fixed directions, **characterized in that** the booth comprising the reference points (REP1, REP2) on the respective axes (x, y), the intersection of which defines a position of the ears of the individual (IND) in the booth (CAB), the sound sources (S1-Sn), provided in said booth (CAB) for performing said measurements, are placed at different distances from said intersection.

9. Computer program product, designed to be stored in a memory of a processing unit or on a removable medium specifically for cooperating with a drive of said processing unit, or intended to be transmitted from a server to said processing unit, comprising instructions in computer code form to construct a model based on an artificial neural network and capable of giving transfer functions HRTFs of an individual for a multiplicity of directions, based on a series of measurements, performed on said individual, representative of HRTFs, only in a few arbitrarily fixed directions of said multiplicity of directions, the pro-

gram using a database including a plurality of HRTFs in a multiplicity of directions in space and for a plurality of individuals to implement at least one learning phase.

5

10. Computer program product, designed to be stored in a memory of a processing unit or on a removable medium specifically for cooperating with a drive of said processing unit, or intended to be transmitted from a server to said processing unit, comprising instructions in computer code form for implementing a model based on an artificial neural network and capable of giving transfer functions HRTFs of an individual for a multiplicity of directions, based on a series of measurements performed on said individual, representative of HRTFs, only in a few arbitrarily fixed directions of said multiplicity of directions, the measurements performed on the individual being carried out in a booth in which:

10

15

- the measurement directions correspond to said arbitrarily fixed directions, and
- the booth comprising the reference points (REP1, REP2) on the respective axes (x, y), the intersection of which defines a position of the ears of the individual (IND) in the booth (CAB) for performing said measurements, the sound sources (S1-Sn), provided in said booth (CAB), are placed at different distances from said intersection.

20

25

30

35

40

45

50

55

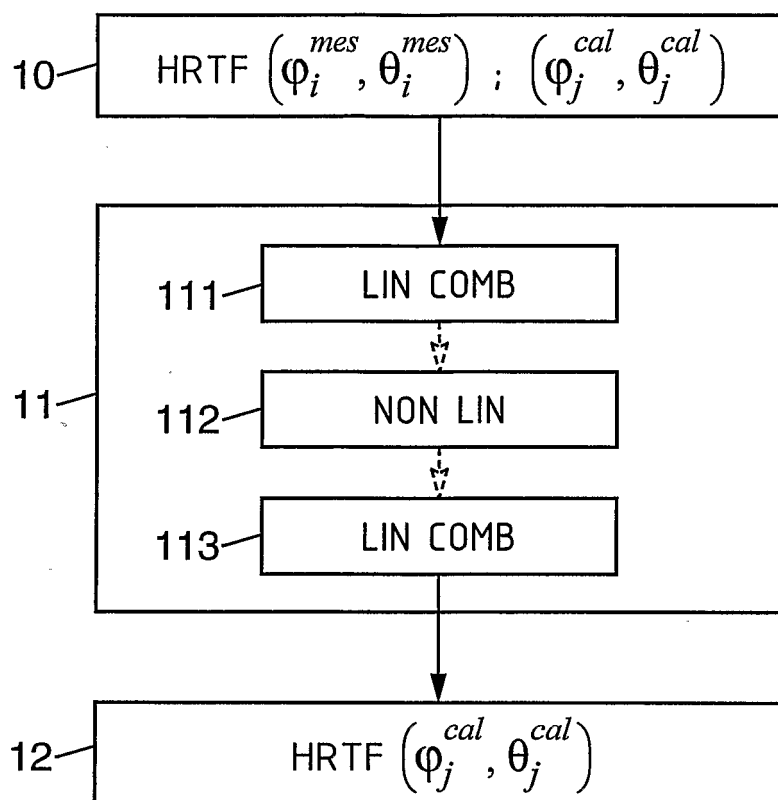


FIG. 1

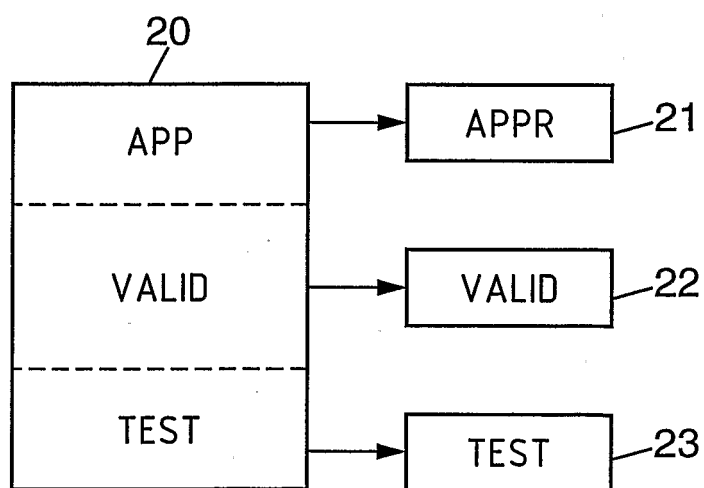


FIG. 2

FIG. 3

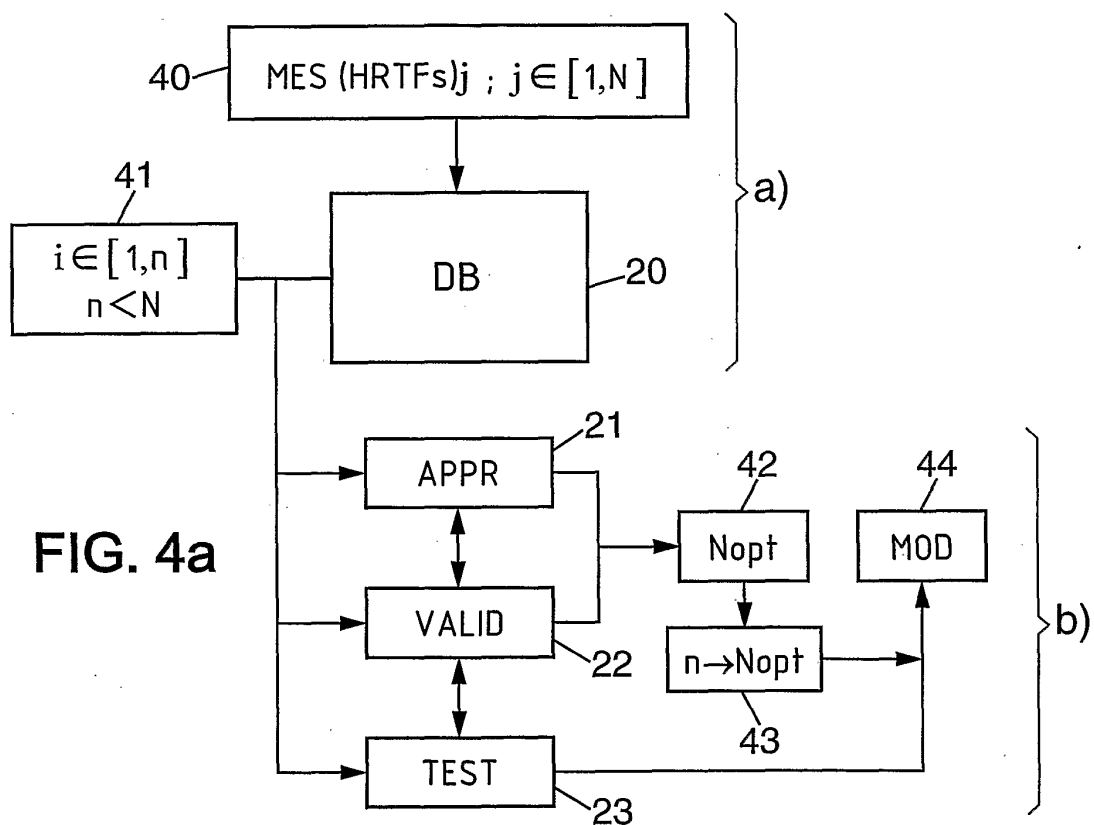
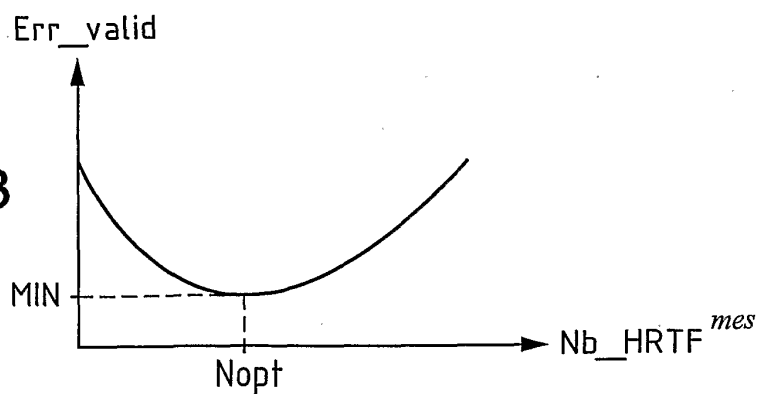
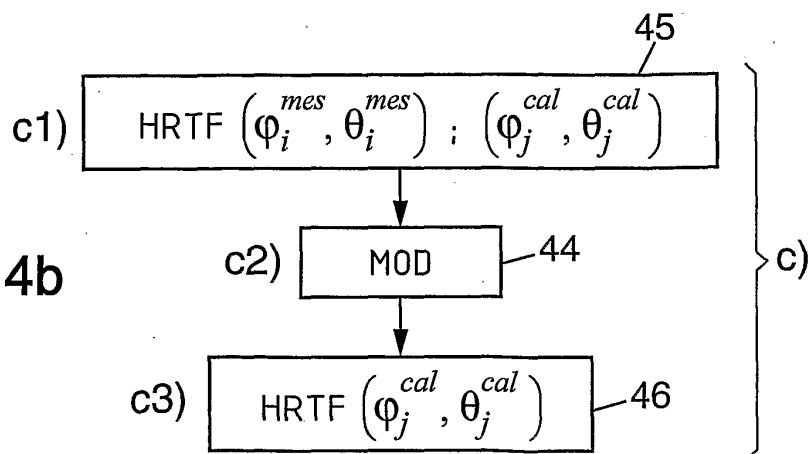


FIG. 4b



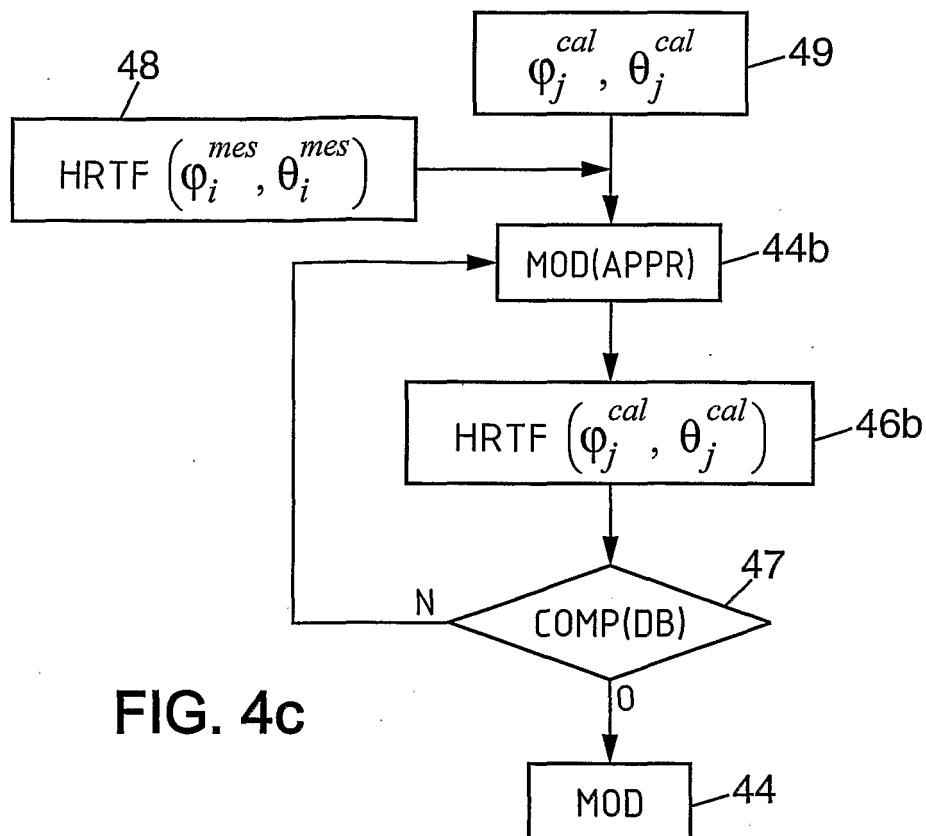


FIG. 4c

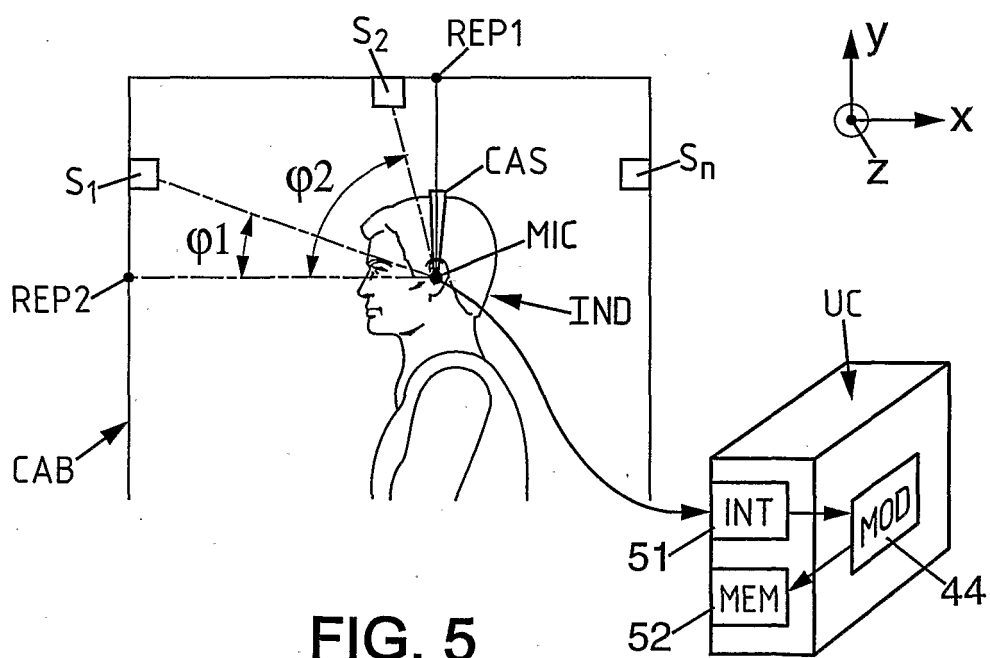


FIG. 5

RÉFÉRENCES CITÉES DANS LA DESCRIPTION

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

Documents brevets cités dans la description

- US 2003138107 A [0013]

Littérature non-brevet citée dans la description

- **RICK L JENISON et al.** A Spherical Basis Function Neural Network for Modeling Auditory Space. *NEURAL COMPUTATION*, Janvier 1996, vol. 8 (1), 115-128 [0012]