



(11)

EP 3 348 079 B1

(12)

FASCICULE DE BREVET EUROPEEN

(45) Date de publication et mention
de la délivrance du brevet:
13.05.2020 Bulletin 2020/20

(51) Int Cl.:
H04S 7/00 (2006.01) G06T 17/00 (2006.01)
G06F 17/18 (2006.01)

(21) Numéro de dépôt: **16736088.2**

(86) Numéro de dépôt international:
PCT/EP2016/065839

(22) Date de dépôt: **05.07.2016**

(87) Numéro de publication internationale:
WO 2017/041922 (16.03.2017 Gazette 2017/11)

(54) **PROCÉDÉ ET SYSTÈME D'ÉLABORATION D'UNE FONCTION DE TRANSFERT RELATIVE À LA TÊTE ADAPTÉE À UN INDIVIDU**

VERFAHREN UND SYSTEM ZUR ENTWICKLUNG EINER AN EIN INDIVIDUUM ANGEPASSTEN KOPFBEZOGENEN ÜBERTRAGUNGSFUNKTION

METHOD AND SYSTEM FOR DEVELOPING A HEAD-RELATED TRANSFER FUNCTION ADAPTED TO AN INDIVIDUAL

(84) Etats contractants désignés:
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB
GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO
PL PT RO RS SE SI SK SM TR**

(30) Priorité: **07.09.2015 FR 1558279**

(43) Date de publication de la demande:
18.07.2018 Bulletin 2018/29

(73) Titulaire: **Mimi Hearing Technologies GmbH
10245 Berlin (DE)**

(72) Inventeurs:
• **GHORBAL, Slim
35000 Rennes (FR)**
• **SEGUIER, Renaud
35690 Acigne (FR)**
• **BONJOUR, Xavier
78560 Le Port Marly (FR)**

(74) Mandataire: **Brunelli, Gérald
Marks & Clerk France
Immeuble Visium
22, avenue Aristide Briand
94117 Arcueil Cedex (FR)**

(56) Documents cités:
US-A1- 2006 067 548 US-B2- 7 209 564

- **RODRIGUEZ SERGIO G ET AL: "HRTF Individualization by Solving the Least Squares Problem", 118TH AES CONVENTION, 6438, 28 mai 2005 (2005-05-28), - 31 mai 2005 (2005-05-31), XP040372767,**
- **MESHRAM ALOK ET AL: "P-HRTF: Efficient personalized HRTF computation for high-fidelity spatial sound", 2014 IEEE INTERNATIONAL SYMPOSIUM ON MIXED AND AUGMENTED REALITY (ISMAR), IEEE, 10 septembre 2014 (2014-09-10), pages 53-61, XP032676177, DOI: 10.1109/ISMAR.2014.6948409**
- **TORRES-GALLEGOS EDGAR A ET AL: "Personalization of head-related transfer functions (HRTF) based on automatic photo-anthropometry and inference from a database", APPLIED ACOUSTICS, vol. 97, 7 avril 2015 (2015-04-07), pages 84-95, XP029221944, ISSN: 0003-682X, DOI: 10.1016/J.APACoust.2015.04.009**
- **GUILLON PIERRE ET AL: "Head-Related Transfer Function Customization by Frequency Scaling and Rotation Shift Based on a New Morphological Matching Method", AES CONVENTION 125; OCTOBER 2008, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, 1 octobre 2008 (2008-10-01), XP040508788,**

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

EP 3 348 079 B1

- ZOTKIN D N ET AL: "HRTF personalization using anthropometric measurements", APPLICATIONS OF SIGNAL PROCESSING TO AUDIO AND ACOUSTICS, 2003 IEEE WORKSHOP ON. NEW PALTZ, NY, USA OCT., 19-22, 2003, PISCATAWAY, NJ, USA, IEEE, 19 octobre 2003 (2003-10-19), pages 157-160, XP010697926, DOI: 10.1109/ASPAA.2003.1285855 ISBN: 978-0-7803-7850-6 cité dans la demande

Description

[0001] L'invention porte sur un procédé et un système d'élaboration d'une fonction de transfert relative à la tête adaptée à un individu.

[0002] La présente invention a trait à la personnalisation des procédés de spatialisation sonore, aussi qualifiés d'écoute binaurale. Plus particulièrement, il s'agit d'une méthode d'individualisation des fonctions de transfert relatives à la tête ou "Head-Related Transfer Functions" en langue anglaise, d'acronyme HRTF, éléments piliers de l'audition tridimensionnelle de tout individu.

[0003] L'écoute binaurale est un domaine de recherche visant à comprendre les mécanismes permettant à l'être humain de percevoir l'origine spatiale des sons. Partant du postulat que cette origine est déterminée grâce à la morphologie de chacun, l'écoute binaurale stipule notamment que la position et la forme des oreilles d'un individu en sont des éléments déterminants. Ces dernières agissent en effet comme des filtres fréquentiels et directionnels sur les sons nous parvenant.

[0004] Si les relations entre la morphologie et l'audition ont de longue date été étudiées, on note depuis près d'un quart de siècle un intérêt croissant dans la communauté scientifique pour le problème de l'individualisation, c'est-à-dire de la prise en compte des spécificités propres à chacun.

[0005] En particulier, l'attention s'est portée sur l'individualisation des fonctions de transfert relatives à la tête ou HRTFs, représentations mathématiques de la coloration fréquentielle des sons que nous percevons. On entend par coloration fréquentielle, les variations de densité spectrale de puissance des signaux sonores. Les spectres des bruits blanc, rose ou encore gris en sont des exemples. Il est maintenant connu de nombreuses méthodes que l'on peut classer en deux grandes familles: les méthodes de synthèse, qui visent à calculer ou recréer des jeux ou ensembles d'HRTFs, et les méthodes adaptatives, qui cherchent à découvrir, parmi un ensemble donné et au prix éventuel de transformations mineures, la fonction de transfert la plus adaptée à un individu.

[0006] Parmi les méthodes de synthèse, on peut tout d'abord distinguer les calculs exacts des approches statistiques et probabilistes.

[0007] Développée depuis plus de vingt ans, la famille des méthodes aux éléments finis vise à modéliser puis résoudre le problème aux dérivées partielles posé par la propagation du son de la source aux tympans du sujet. Cette famille comprend notamment les variantes connues sous les appellations anglaises: "Direct Boundary Element Method", d'acronyme DBEM, "Indirect Boundary Element Method" d'acronyme IBEM, "In_nite-Finite Element Method" d'acronyme IFEM, ou "Fast-Multipole Boundary Element Method" d'acronyme FM-BEM.

[0008] Réputées offrir des solutions exactes au problème traité, ces méthodes souffrent néanmoins de quelques handicaps notables. Tout d'abord, elles nécessitent la donnée d'un maillage 3D du sujet d'autant plus fin que

l'on souhaite calculer les HRTFs dans les hautes fréquences, et que le temps de calcul devient rapidement prohibitif à mesure que l'on affine le maillage (et donc que l'on souhaite disposer de résultats fiables dans les hautes fréquences). On entend par hautes fréquences des fréquences supérieures à 4 kHz. Enfin, la modélisation physique du problème nécessite d'introduire beaucoup d'a priori et d'approximations. Ainsi, chaque surface se voit attribuer une impédance propre (traduisant les phénomènes d'absorption/réflexion) dont la valeur est empirique. De même, la chevelure est classiquement modélisée par une surface d'impédance différente de la peau, ne tenant donc pas compte de sa nature volumique.

[0009] Une approche alternative au calcul direct des HRTFs consiste, à partir d'un ensemble représentatif d'HRTFs réelles, d'en faire émerger les principaux modes de variation.

[0010] C'est notamment ce que réalisent les travaux de Sylvain Busson ("Individualisation d'Indices Acoustiques pour la Synthèse Binaurale"; PhD thesis, Université de la Méditerranée-Aix-Marseille II, 2006.) sur les réseaux de neurones artificiels (RNA). L'idée développée est de réaliser une prédiction des HRTFs à partir de la mesure d'un nombre restreint d'entre-elles. Cela passe en particulier par l'utilisation conjointe d'une carte de Kohonen et d'une Classification Hiérarchique Ascendante, d'acronyme CHA, avant l'élection d'HRTFs représentatives. Par la suite, un réseau de neurone de type Multi Layer Perceptron en langue anglaise, d'acronyme MLP, à trois couches, est construit et les HRTFs représentatives de 44 sujets de la base CIPIC utilisés comme ensemble d'apprentissage. Bien que prometteuse, cette étude ne parvient pas à dégager de représentants universels, i.e. communs à tous les individus, ni ne présente de validation psycho-acoustique des résultats. De plus, il est également nécessaire de disposer d'un moyen d'accès auxdits représentants.

[0011] Les méthodes statistiques pour la synthèse d'HRTFs peuvent, en variante, se fonder sur l'analyse en composantes principales, d'acronyme ACP.

[0012] Kistler et Wightman ("A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction"; The Journal of the Acoustical Society of America, 91(3) :1637-1647, 1992) furent les premiers à proposer de décomposer les HRTFs selon cette méthode. L'ensemble des HRTFs est alors vu comme un sous-espace vectoriel de l'espace des mesures. La connaissance d'une base de ce sous-espace permet ensuite d'en atteindre n'importe quel représentant, i.e. n'importe quelle HRTF, par simple combinaison linéaire des vecteurs de base. C'est ce que permet l'ACP en fournissant une base orthonormée de l'espace engendré par les HRTFs d'apprentissage. La dernière étape de la résolution du problème d'individualisation consiste alors à faire le lien entre les paramètres morphologiques des individus et les coefficients de reconstruction par les vecteurs propres de la base. Pour

cela, des régressions linéaires multiples sont classiquement utilisées.

[0013] Partant des travaux de Kistler & Wightman, Xu et associés (Song Xu, Zhizhong Li, and Gavriel Salvendy: "improved method to individualize head-related transfer function using anthropométrie measurements"; *Acoustical Science and Technology*, 29(6) :388-390, 2008.) ont proposé de grouper les HRTFs des différents individus mesurés selon la direction (azimut, élévation) pointée avant d'effectuer l'ACP (une par groupe), espérant ainsi réduire l'erreur d'estimation.

[0014] Zhang et associés (R. A. Kennedy M. Zhang and T. D. Abhayapala; "Statistical method to identify key anthropométrie parameters in hrtf individualization"; In *Joint Workshop on Hands-free Speech Communication and Microphone Arrays*, 2011) ont quant à eux proposés une méthode statistique d'estimation des paramètres anthropomorphiques les plus pertinents pour réaliser l'étape de régression.

[0015] En 2007, Vast Audio Pty Ltd a déposé un brevet (G. Jin, P. Leong, J. Leung, S. Carlile, and A. Van Schaik; "Génération of customized three dimensional sound effects for individuals", April 24 2007, US 7209564) inspiré par ces idées. En pratique, ce dernier décrit tout d'abord la création d'une base d'HRTFs et d'une base de paramètres morphologiques. Est ensuite invoquée l'utilisation d'une méthode d'analyse statistique pour décomposer en composantes élémentaires les espaces de paramètres et d'HRTFs, à la manière de ce que permet l'ACP. Par la suite, à l'aide d'une autre méthode d'analyse statistique, les liens entre les coefficients de reconstruction des paramètres morphologiques et ceux des HRTFs sont déterminés.

[0016] Une autre approche statistique, cherchant à déterminer les liens entre les HRTFs et des images de la tête, du torse et des oreilles des individus correspondant, est proposée dans le document US 2006/0067548 A1.

[0017] Chaque variante proposée jusqu'à maintenant a généralement permis d'améliorer les résultats des méthodes antérieures sans toutefois offrir de rendu satisfaisant du point de vue psycho-acoustique, i.e. en conditions réelles. En particulier, le nombre et la localisation des paramètres morphologiques nécessaires sont très imprécis. De plus, dans le cas d'analyse simultanée de la morphologie et des HRTFs, la découverte des liens entre les coefficients des deux espaces est d'autant plus complexe que les données sont laissées brutes.

[0018] Un autre type de méthode de synthèse, notable par son caractère innovant, est la reconstruction d'HRTFs selon une approche Bayésienne. Présentée par Hofman & Van Opstal (Paul M Hofman and A John Van Opstal. Bayesian; "reconstruction of sound localization cues from responses to random spectra", *Biological cybernetics*, 86(4):305-316, 2002), qui veut recréer des HRTFs potentielles à partir d'une analyse probabiliste des réponses des sujets étudiés à des stimuli bien précis. Plus particulièrement, l'idée est de faire écouter aux sujets des sons convolués par des filtres mimant les types

de variations observables dans de véritables HRTFs et diffusés par un haut-parleur situé droit devant eux. La consigne donnée est de diriger le regard dans la direction dont semble leur provenir le son.

[0019] Bien qu'innovante, cette méthode présente toutefois de nombreuses contraintes jouant en sa défaveur comme le temps nécessaire à l'expérimentation ou l'impossibilité d'adresser les HRTFs hors du champ de vision, le sujet étant contraint à désigner du regard les directions d'où semblent leur provenir les sons.

[0020] Alors que les méthodes de synthèse citées précédemment visent à créer de tout nouveaux jeux d'HRTFs (sans parfois même en avoir jamais observé de réels, comme c'est le cas pour les méthodes aux éléments finis), les méthodes adaptatives visent, au contraire, à rester au plus près de l'existant. L'idée sous-jacente consiste en l'exécution des mesures sur de vrais sujets pour obtenir des jeux d'HRTFs au moins adaptés à une personne. Ils contiennent donc nécessairement suffisamment d'indices de localisation pour être utilisables, ce que les méthodes de synthèse ne peuvent promettre.

[0021] Les méthodes sélectives n'entraînent aucune altération des mesures; leur principe commun est l'élection d'un jeu d'HRTFs parmi plusieurs selon certains critères. Ces derniers sont le plus souvent psycho-acoustiques, sans pour autant y être limités.

[0022] Parmi les critères psycho-acoustiques, il convient en premier lieu de citer les travaux de Shimada et associés (Shoji Shimada, Nobuo Hayashi, et Shinji Hayashi; "A clustering method for sound localization transfer functions", *Journal of the Audio Engineering Society*, 42(7/8) :577-584, 1994). Partant d'une base conséquente d'HRTFs, ces derniers entendent réaliser des regroupements entre HRTFs similaires. Pour ce faire, ils opèrent une composition cepstrale de 16 coefficients. La distance euclidienne naturellement associée à cet espace à 16 dimensions permet alors le regroupement des HRTFs en classes (au nombre de 8). Des jeux d'HRTFs sont ensuite choisis aléatoirement au sein des classes et les sujets invités à élire le ou les classes qui leur offrent la meilleure impression d'externalisation et de directivité.

[0023] Plus récemment, on peut se référer aux travaux de Tame et associés (Robert P Tame, Daniele Barchiese, et Anssi Klapuri; "Headphone virtualization : Improved localization and externalization of nonindividualized hrtfs by cluster analysis", in *Audio Engineering Society Convention 133*; Audio Engineering Society, May 2012.) ou encore ceux de Xie et associés (Bosun Xie et Zhaojun Tian; "Improving binaural reproduction of 5.1 channel surround sound using individualized hrtf cluster in the wavelet domain", in *Audio Engineering Society Conference : 55th International Conference : Spatial Audio*, Audio Engineering Society, August 2014.) qui utilisent respectivement des gaussiennes et une décomposition en ondelettes pour réaliser le regroupement des HRTFs.

[0024] Une fois la classe (ou cluster en langue

anglaise) sélectionnée, une autre étape de sélection peut être ajoutée pour sélectionner un jeu bien précis. Là encore, de multiples méthodes ont été publiées. Ainsi, Y. Iwaya (Yukio Iwaya, "Individualization of head-related transfer functions with tournament-style listening test : Listening with other's ears", *Acoustical science and technology*, 27(6): 340-343, 2006.) décrit une procédure de sélection d'un jeu d'HRTFs parmi 32 disponibles en reprenant le principe des tournois d'échec. Une trajectoire sonore dans le plan horizontal est simulée par convolution d'un bruit rose avec les jeux d'HRTFs. Un bruit rose est un bruit dont la puissance sonore est constante pour une largeur de bande fréquentielle donnée dans un espace logarithmique (ex : même puissance émise sur la bande 40-60Hz que sur la bande 4000-6000Hz). 32 trajectoires sont donc obtenues et mises en compétitions. A chaque rencontre, le sujet déclare vainqueur l'une des deux trajectoires selon qu'elle ressemble le plus ou non à la trajectoire de consigne. Le jeu sortant vainqueur du tournoi est déclaré le plus adapté au sujet.

[0025] Une autre approche, de Seeber et associés (Bernhard U Seeber et Hugo Fastl; "Subjective selection of non-individual head-related transfer functions", July 2003.), présente une sélection en deux étapes d'un jeu parmi 12. L'objectif affiché est d'être rapide sans entraînement préalable tout en fournissant un résultat minimisant l'impression de son intra-cranien. La première étape consiste à désigner les 5 jeux présentant un meilleur rendu en termes de spatialisation dans la zone frontale. La seconde consiste à en éliminer 4 selon qu'ils pêchent à reproduire différents comportements tels que le déplacement d'une source sonore à vitesse constante, à élévation constante ou encore à distance constante. Une dizaine de minutes est nécessaire à la réalisation de la procédure.

[0026] Enfin, on cite également les travaux de Martens (William L Martens; "Rapid psychophysical calibration using bisection scaling for individualized control of source elevation in auditory display"; in *Proc. Int. Conf. on Auditory Display*, pages 199-206, July 2002) connus comme bisection scaling. L'idée est de créer, à l'aide d'un test psycho-acoustique, une table de correspondance entre les directions réelles associées à un jeu d'HRTFs et les directions perçues par le sujet. En pratique, pour un azimut donné il faut trouver l'HRTF correspondant le mieux à la sensation d'une élévation à 45°. Les élévations extrêmes (0° et 90°) étant supposées correctement perçues, une interpolation polynomiale du second ordre est ensuite opérée pour construire la table évoquée ci-avant.

[0027] D'autres protocoles encore ont été proposés par la communauté scientifique mais aucun ne permet d'éviter les inconvénients inhérents à ce type de méthodologie. En effet, même si l'objectif n'est pas de trouver les HRTFs exactes du sujet (il faudrait faire appel aux méthodes de synthèse) mais de sélectionner ou de s'adapter au mieux à l'existant, il n'en reste pas moins que la qualité de la meilleure solution possible est tou-

jours limitée par la variabilité des jeux d'HRTFs ouverts à la sélection. Ainsi, pour un protocole donné, les résultats sont d'autant meilleurs que la base de données d'entrée est importante. Or l'augmentation de cette dernière allonge de fait la durée de l'expérimentation, ce qui est d'autant plus gênant qu'elle repose sur la participation active du sujet.

[0028] Remettant au premier plan l'importance de la morphologie propre à chacun, Zotkin et associés (D.N. Zotkin, J. Hwang, R. Duraiswaini, et L.S. Davis; "Hrtf personalization using anthropométrie measurements", in *Applications of Signal Processing to Audio and Acoustics*, 2003 IEEE Workshop on, pages 157-160, Oct 2003.) décrivent l'oreille au travers de sept paramètres morphologiques mesurables sur une vue de profil de l'oreille. Ces paramètres permettent de définir une distance entre les individus qui est utilisée pour sélectionner le plus proche voisin dans la base CIPIC d'un sujet donné. On note que les HRTFs ainsi sélectionnées ont ensuite fait l'objet d'une modification pour les fréquences inférieures à 3 kHz. En effet, pour les basses fréquences ($f \leq 500\text{Hz}$), un modèle Tête et Torse, d'acronyme HAT pour "Head-And-Torso" en langue anglaise est utilisé pour synthétiser les HRTFs. Entre 500 Hz et 3 kHz, un recollement affine est opéré pour passer progressivement des HRTFs de synthèse aux HRTFs sélectionnées.

[0029] En 2001, la société Arkamys et le CNRS ont déposé un brevet (B.F. Katz and D. Schönstein, "Procédé de sélection de filtres hrtf perceptivement optimale dans une base de données à partir de paramètres morphologiques", WO2011128583) portant sur une méthode de sélection morphologique. L'idée est de constituer trois bases de données. La première contient les HRTFs d'un ensemble d'individus, la deuxième contient un jeu de paramètres morphologiques de ces individus et la troisième contient les préférences d'écoute de ces individus, i.e. pour chaque sujet, la classification qu'il fait des HRTFs de la première base. Une fois cela posé, une étude des corrélations entre les deuxième et troisième bases de données est réalisée pour classer les paramètres morphologiques par ordre d'importance. Du côté des HRTFs, une analyse dimensionnelle de l'espace est menée (par exemple une ACP) pour en obtenir une base dans laquelle elles deviennent représentables. Les liens entre K paramètres morphologiques les plus importants et les coordonnées des HRTFs dans l'espace précité sont alors calculés, établissant un lien entre morphologie et HRTFs. Etant donné un nouvel individu, la mesure des K paramètres morphologiques mis en lumière précédemment permet ensuite de se positionner dans l'espace des HRTFs. Le plus proche voisin présent en base est recherché et constitue le résultat de la personnalisation.

[0030] Le problème rencontré par les précédentes méthodes utilisant des paramètres morphologiques, à savoir, de définir leur nombre et leur localisation. En effet, la notion de hauteur d'une oreille, par exemple, n'a rien de naturel et sa mesure sera très dépendante de la subjectivité de l'expérimentateur qui devra avant toute chose

déterminer si l'oreille doit être tournée et où se situent ses points les plus "bas" et "haut". Par ailleurs, se pose la question des critères de définition de la distance utilisée car c'est de cette dernière que dépend le résultat de la sélection.

[0031] Enfin viennent les méthodes de sélection adaptée, dont le représentant le plus explicite est sans doute la mise à l'échelle en fréquences ou "Frequency Scaling" en langue anglaise, introduite par Middlebrook (John C Middlebrooks, "Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency", The Journal of the Acoustical Society of America, 106(3) :1493-1510, 1999); cette opération repose sur l'idée que l'interaction d'une onde sonore de fréquence donnée avec un solide dépend des dimensions de ce dernier. En particulier, toute homothétie opérée sur l'objet doit s'accompagner, si l'on souhaite toujours observer la même interaction, d'une homothétie de rapport inverse sur la fréquence. Appliquée à l'individualisation, cette idée revient à dire qu'en connaissant les HRTFs d'un individu de référence (ou même d'un mannequin) et le rapport d'échelle ("scaling factor" en langue anglaise) entre la morphologie de cette référence et celle d'un sujet à individualiser, il est possible d'améliorer la sensation de localisation apportée à celui-ci par les HRTFs de référence en leur appliquant une mise à l'échelle de rapport inverse.

[0032] En parallèle à la mise à l'échelle en fréquences ou "Frequency Scaling", Maki et Furukawa (Katsuhiro Maki et Shigeto Furukawa; "Reducing individual differences in the external-ear transfer functions of the mongolian gerbil; The Journal of the Acoustical Society of America, 118(4), 2005) ont montré que, partant de la donnée de l'angle entre un pavillon d'oreille de référence et un pavillon test, une rotation du système de coordonnées donnant la direction des HRTFs permet de réduire significativement les différences inter-individus. En d'autres termes, ce procédé utilise le fait, en le restreignant au pavillon d'oreille, qu'une rotation du sujet induit la même rotation au niveau des HRTFs mesurées.

[0033] Ces approches, si utiles soient-elles, ne sauraient néanmoins constituer à elles seules des procédés complets de personnalisation. Cela reviendrait à réduire la variabilité des HRTFs à seulement 1 ou 2 paramètres. Toutefois, elles peuvent être vues comme de bons compléments à d'autres méthodes.

[0034] En dépit de la multiplicité des approches connues visant à personnaliser l'écoute binaurale, aucune n'est encore parvenue à se détacher clairement des autres par son efficacité et sa simplicité. De plus, des problèmes peuvent en découler comme des temps de personnalisation prohibitifs ou un manque de fiabilité des solutions, si ce n'est les deux simultanément.

[0035] Un but de l'invention est d'élaborer une fonction de transfert relative à la tête (HRTF) adaptée à un individu avec une rapidité et une fiabilité améliorées.

[0036] Dans la suite de la description, l'expression "données d'oreilles", "espace des oreilles" ou "oreilles"

signifie des photos 2D d'oreilles ou des oreilles 3D représentées par un nuage de points 3D décrivant la surface de l'oreille.

[0037] Aussi, il est proposé, selon un aspect de l'invention, un procédé d'élaboration d'une fonction de transfert relative à la tête ou HRTF adaptée à un individu, à partir d'une base de données comprenant des données d'oreilles 3D ou 2D et des fonctions de transfert correspondantes relatives à la tête, le procédé comprenant les étapes consistant à:

- effectuer une analyse statistique de l'espace des oreilles 3D ou 2D, de la base de données;
- effectuer une analyse statistique de l'espace des fonctions de transfert relatives à la tête, de la base de données;
- effectuer une analyse des liens entre lesdits paramètres statistiques de l'espace des oreilles 3D ou 2D et lesdits paramètres statistiques de l'espace des fonctions de transfert relatives à la tête; et
- déterminer, à partir de ladite analyse des liens et de ladite analyse statistique de l'espace des oreilles 3D ou 2D, une fonction de calcul d'une fonction de transfert relative à une tête à partir de données représentatives d'au moins une oreille.

[0038] Ainsi, les relations entre HRTFs et données d'oreilles étant déterminées en amont, il est possible de les utiliser dans des applications temps réel. Par ailleurs, le caractère statistique des analyses permet de s'affranchir des simplifications introduites par les modèles physiques et des approximations qui en découlent.

[0039] Bien entendu, une HRTF est liée à une direction de l'espace, et pour recréer un environnement virtuel auditif complet, il faut donc disposer d'HRTFs pour un nombre conséquent de directions, ce que permet de faire la présente invention pour un nombre quelconque de directions souhaitées.

[0040] Selon un mode de mise en œuvre, le procédé comprend, en outre, une étape consistant à mettre en correspondance dense, ou "dense registration en langue anglaise, des points relatifs à des positions respectives des oreilles de la base de données.

[0041] Dans un mode de mise en œuvre, le procédé comprend, en outre, une étape de calcul d'une fonction de transfert relative à la tête, adaptée à l'individu, à partir de ladite fonction de calcul et d'au moins une photographie d'au moins une oreille de l'individu.

[0042] Ainsi, l'utilisation de la fonction de calcul permet la détermination de la fonction de transfert en un temps compatible avec une application temps réel.

[0043] Selon un mode de mise en œuvre, ladite étape de calcul d'une fonction de transfert relative à la tête est itérative.

[0044] Dans un mode de mise en œuvre, ladite étape itérative de calcul d'une fonction de transfert relative à la tête comprend :

- une première sous-étape itérative d'estimation d'au moins un paramètre de pose de l'individu lors de la ou lesdites photographies; et
- une deuxième sous-étape itérative d'estimation de paramètres statistiques optimisés représentant au moins une oreille de l'individu dans l'espace des oreilles.

[0045] Ainsi, il est possible de reconstituer une oreille en 3D à partir d'une photographie qui ne nécessite pas que l'utilisateur prenne de précaution particulière lors de la prise du cliché.

[0046] Selon un mode de mise en œuvre, lesdites données représentant des oreilles 3D sont des nuages de points.

[0047] Ainsi, la visualisation et l'étude des propriétés, notamment géométriques, des données sont facilitées.

[0048] Dans un mode de mise en œuvre, on utilise lesdites étapes divulguées pour élaborer une fonction de transfert, pour de hautes fréquences supérieures à un seuil, relative à la tête adaptée à l'individu, ledit procédé comprenant, en outre, une étape d'élaboration d'une fonction de transfert, pour de basses fréquences inférieures audit seuil, relative à la tête adaptée à l'individu.

[0049] Ainsi, chaque partie du spectre fréquentielle se voit adaptée en fonction des structures physiques qui l'impactent le plus.

[0050] Selon un mode de mise en œuvre, ladite étape d'élaboration d'une fonction de transfert, pour de basses fréquences inférieures audit seuil, relative à la tête adaptée à l'individu comprend les sous-étapes suivantes, consistant à:

- échantillonner de plages de valeurs possibles de paramètres morphologiques humains d'une base de données relatives à la morphologie humaine,
- déterminer d'un maillage de modèle paramétrique desdits paramètres morphologiques,
- calculer des fonctions de transfert gabarit de basses fréquences, associées audit maillage,
- estimer la valeur des paramètres morphologiques de l'individu à partir d'au moins une photo de l'individu de face ou de profil, et
- calculer une fonction de transfert, pour de basses fréquences, relative à la tête, adaptée à l'individu à partir de la valeur estimée des paramètres morphologiques et desdites fonctions calculées de transfert gabarit de basses fréquences.

[0051] Ainsi, la plupart des calculs est menée en amont, permettant l'utilisation du procédé au sein d'applications en temps réel.

[0052] Dans un mode de mise en œuvre, une fonction de transfert relative à la tête de l'individu est élaborée à partir desdites fonctions de transfert respectivement pour de hautes et basses fréquences et de ladite ou lesdites photos de l'individu de face ou de profil, comprenant les étapes consistant à:

- estimer, à partir de ladite ou lesdites photos de l'individu de face ou de profil, la taille d'oreilles relativement au reste du corps de l'individu;
- mettre à l'échelle en fréquences les fonctions de transfert relatives à la tête, pour les hautes fréquences; et
- fusionner les fonctions de transfert, respectivement pour de hautes et basses fréquences, pour obtenir la fonction de transfert relative à la tête de l'individu.

[0053] Pour un individu, la photo d'une seule oreille, peut suffire, en supposant une symétrie des oreilles d'un individu, mais en variante, une meilleure précision est obtenue avec des photos des deux oreilles d'un individu.

[0054] Il est également proposé, selon un autre aspect de l'invention, un système d'élaboration d'une fonction de transfert relative à la tête ou HRTF adaptée à un individu, à partir d'une base de données comprenant des données d'oreilles et des fonctions de transfert correspondantes relatives à la tête, comprenant un calculateur configuré pour mettre en œuvre le procédé selon l'une des revendications précédentes.

[0055] L'invention sera mieux comprise à l'étude de quelques modes de réalisation décrits à titre d'exemples nullement limitatifs et illustrés par les dessins annexés sur lesquels les figures 1 à 4 illustrent schématiquement le procédé selon l'invention.

[0056] Sur la figure 1, une base de données OH_1 comprend des données O_1 d'oreilles et des fonctions de transfert H_1 correspondantes relatives à la tête. On entend par "correspondantes", le fait que pour cette base de données, on enregistre lors de sa conception, pour les individus servant à concevoir la base de données, les données représentatives des oreilles des personnes de la base, ainsi que leurs fonctions de transfert relatives à la tête, en gardant le lien entre les données d'oreilles et la fonction de transfert correspondant de la base de données.

[0057] Les données O_1 d'oreilles peuvent être des nuages de points.

[0058] Une étape S1, optionnelle, permet de mettre en correspondance dense des points relatifs à des positions respectives des oreilles O_1 de la base de données OH_1 .

[0059] On entend par mise en correspondance dense, la spécification des correspondances entre les points constitutifs d'un nuage ou les pixels d'une image 2D d'oreille et ceux constitutifs d'un autre nuage ou d'une autre image 2D d'oreille. À titre d'exemple, si l'extrémité du lobe est représentée par le point 2048 sur une oreille et par le point 157 sur une autre, la spécification de cette équivalence de rôle constitue une mise en correspondance. On pourra parler de classe d'équivalence, tous les points d'une même classe jouant un rôle similaire au sein de leur oreille d'appartenance.

[0060] Il est possible de n'utiliser qu'une oreille, en supposant une symétrie des oreilles d'un utilisateur.

[0061] Une étape S2 permet ensuite d'effectuer une analyse statistique de l'espace des oreilles O_1 , de la base

de données OH_1 . Cette analyse statistique peut se faire aux moyen de techniques utilisant une base d'exemples d'oreilles et réalisant une réduction de dimension (analyse en composantes principales, analyse en composantes indépendantes, codage de type *sparse* ou parcimonieux, réseaux de neurones de type auto-encodeurs). Ces techniques permettent de convertir la représentation d'une oreille 2D ou 3D (sous la forme d'un nuage de points ou de pixels dans une image) en un vecteur de paramètres statistiques de nombre restreint.

[0062] Une étape S3 permet d'effectuer une analyse statistique de l'espace des fonctions de transfert relatives à la tête H_1 , de la base de données OH_1 . Cette analyse statistique est du même type que celle décrite dans le paragraphe précédent. Elle permet donc de représenter les HRTF par un vecteur de paramètres statistiques de nombre restreint.

[0063] Une étape S4 permet d'effectuer une analyse des liens entre lesdits paramètres statistiques de l'espace des oreilles de l'étape S2 et lesdits paramètres statistiques de l'espace des fonctions de transfert relatives à la tête de l'étape S3.

[0064] Enfin, une étape S5 permet de déterminer, à partir de ladite analyse des liens de l'étape S4, et de ladite analyse statistique de l'espace des oreilles de l'étape S2, une fonction de calcul OH'_1 d'une fonction de transfert S_1 relative à une tête à partir de données représentatives d'au moins une oreille.

[0065] Les analyses statistiques S2 et S3 doivent aboutir à la création de représentations paramétriques des oreilles et des fonctions de transfert relatives à la tête. En particulier, les données d'apprentissage de la base de données OH_1 doivent pouvoir être reconstruites à partir des sorties de l'analyse.

[0066] Il est notamment possible d'utiliser, dans les étapes d'analyse S2 et S3, des analyses en composantes principales d'acronyme ACP.

[0067] A titre d'exemple, lorsque l'ACP est choisie pour réaliser la réduction de dimension, elle consiste à calculer, à partir d'une base d'exemples des données à analyser, les vecteurs propres qui représentent le mieux ces données au sens des moindres carrés. Les paramètres statistiques qui représentent la donnée à analyser (oreille 3D ou 2D ou fonction de transfert relative à la tête) ne sont ni plus ni moins que les coefficients de projection cette donnée projetée sur les vecteurs propres.

[0068] Alternativement, tout type d'analyse dimensionnelle linéaire ou non, convient, pour autant qu'elle réponde à l'exigence de reconstruction précitée, comme les méthodes d'analyse en composantes indépendantes, d'acronyme ACI, ou de codage clairsemé ou "sparse-coding" en langue anglaise.

[0069] L'analyse des liens de l'étape S4 entre les jeux de paramètres statistiques de l'espace des oreilles et les paramètres statistiques de l'espace des fonctions de transfert relatives à la tête, dans une configuration nominale, peut se faire par régression linéaire multivariée sur

les valeurs des paramètres utilisés pour la reconstruction des données d'apprentissage de la base de données OH_1 .

[0070] Alternativement, toute méthode permettant de trouver les valeurs du jeu de paramètres des fonctions de transfert relatives à la tête à partir des valeurs du jeu de paramètres statistiques et assurant une bonne reconstruction des fonctions de transfert relatives à la tête de la base de données OH_1 , comme des méthodes à base de réseaux de neurones, à base d'analyse en composantes multiples, d'acronyme ACM, ou de partitionnement en k-moyennes.

[0071] Comme illustré sur la figure 2, le procédé peut comprendre, en outre, une étape de calcul S6 d'une fonction de transfert S_1 relative à la tête, adaptée à l'individu,

à partir de ladite fonction de calcul OH'_1 , OH_1 et d'au moins une photographie U_1 d'une oreille de l'individu.

[0072] L'étape de calcul S6 d'une fonction de transfert S_1 relative à la tête peut être itérative, et comprendre une première sous-étape itérative S7 d'estimation d'au moins un paramètre de pose de l'individu lors de la ou lesdites photographies, et une deuxième sous-étape itérative S8 d'estimation de paramètres statistiques optimisés représentant au moins une oreille de l'individu dans l'espace des oreilles.

[0073] Bien entendu l'étape itérative de calcul S6 d'une fonction de transfert S_1 relative à la tête comprend alors également une sous-étape S6a d'initialisation ou mise à jour des paramètres statistiques de forme et des paramètres de pose, ainsi qu'une sous-étape S6b de test de convergence de l'étape de calcul S6 ou d'atteinte d'un nombre limite d'itérations.

[0074] Les première et deuxième sous-étapes itératives S7 et S8 comprennent bien sûr chacune un test de convergence de l'estimation respective ou d'atteinte d'un nombre limite d'itérations.

[0075] Les paramètres de pose dont il est question font référence aux angles sous lesquels sont photographiées les oreilles des utilisateurs.

[0076] Les première et deuxième sous-étapes itératives S7 et S8 d'estimation font intervenir des modèles actifs d'apparence ou "active appearance models" en langue anglaise, d'acronyme AAM. Dans une configuration nominale, ils sont basés sur l'utilisation de matrices de régression.

[0077] En variante, il est possible d'utiliser toute méthode permettant de faire converger la projection en 2D du modèle vers les images 2D des utilisateurs comme des AAM basés sur des descentes de gradient, des algorithmes génétiques ou des simplex.

[0078] Comme illustré sur la figure 3, on utilise lesdites étapes divulguées pour élaborer une fonction de transfert S_H , pour de hautes fréquences supérieures à un seuil, relative à la tête adaptée à l'individu, ledit procédé comprenant, en outre, une étape d'élaboration d'une fonction de transfert S_B , pour de basses fréquences inférieures

audit seuil, relative à la tête adaptée à l'individu.

[0079] L'étape d'élaboration d'une fonction de transfert S_B , pour de basses fréquences inférieures audit seuil, relative à la tête, adaptée à l'individu comprend les sous-étapes suivantes, consistant à:

- échantillonner S9 des plages de valeurs possibles de paramètres morphologiques humains d'une base de données M_1 relatives à la morphologie humaine,
- déterminer S10 un maillage de modèle paramétrique desdits paramètres morphologiques,
- calculer S11 des fonctions de transfert gabarit de basses fréquences (M'_1), associées audit maillage,
- estimer S12 la valeur des paramètres morphologiques de l'individu à partir d'au moins une photo U_2 de l'individu de face ou de profil, et
- calculer S13 une fonction de transfert S_B , pour de basses fréquences, relative à la tête, adaptée à l'individu à partir de la valeur estimée des paramètres morphologiques et desdites fonctions calculées de transfert gabarit de basses fréquences.

[0080] Les fonctions de transfert gabarit de basses fréquences M'_1 sont calculées hors ligne et servent de base de référence de fonctions de transfert relatives à la tête en basses fréquences (fréquences inférieures à un seuil, par exemple 2 kHz).

[0081] Par exemple, il est possible d'utiliser un modèle boules de neige ou "snowball" en langue anglaise. En variante, tout modèle paramétrique à peu d'entrées et permettant d'obtenir un maillage de la tête et du torse convient, comme une modélisation de la tête et du torse par des ellipsoïdes de révolution.

[0082] Par exemple, les paramètres macroscopiques peuvent être la largeur des épaules et le diamètre de la tête. Le choix des paramètres est dicté par le choix du modèle utilisé pour le calcul des gabarits.

[0083] Comme illustré sur la figure 4, une fonction de transfert relative à la tête S_1 de l'individu est élaborée à partir desdites fonctions de transfert S_H , S_B , respectivement pour de hautes et basses fréquences et de ladite ou lesdites photos U_2 de l'individu de face ou de profil, comprenant les étapes consistant à:

- estimer S14, à partir de ladite ou lesdites photos U_2 de l'individu de face ou de profil, la taille d'oreille de l'individu;
- utiliser ladite taille d'oreille estimée de l'individu pour ajuster S15 les fonctions de transfert relatives à la tête S_H à la bande de fréquences la plus adaptée selon la méthode de mise à l'échelle en fréquences ou "frequency scaling" en langue anglaise, pour les hautes fréquences;
- fusionner S16 les fonctions de transfert S_H , S_B , respectivement pour de hautes et basses fréquences,

pour obtenir la fonction de transfert relative à la tête S_1 de l'individu.

[0084] Les dimensions de l'oreille peuvent être normalisées, auquel cas il faut prévoir une remise à l'échelle du spectre fréquentiel généré pour l'oreille.

[0085] En effet, deux oreilles identiques à un facteur d'échelle près ont des HRTFs identiques à l'inverse de ce même facteur d'échelle près. Ceci est très important lorsque l'on travaille avec un modèle d'oreille normalisé et sans information, tout du moins en début d'algorithme, sur les dimensions réelles de l'oreille du sujet. Par conséquent, si le modèle reconstruit une oreille de 5 cm de haut là où l'oreille du sujet en faisait 10 cm, il faudra comprimer les HRTFs par un facteur 0.5.

[0086] En variante, si les oreilles ne font pas l'objet de normalisation en taille, l'étape 15 de mise à l'échelle devient sans objet.

[0087] La fusion des deux parties du spectre par leur sommation après application d'un filtre passe-haut et d'un filtre passe-bas respectivement au spectre de hautes fréquences et au spectre de basses fréquences.

[0088] Les étapes du procédé décrit ci-dessus peuvent être effectuées par un ou plusieurs processeurs programmables exécutant un programme informatique pour exécuter les fonctions de l'invention en opérant sur des données d'entrée et générant des données de sortie.

[0089] Un programme informatique peut être écrit dans n'importe quelle forme de langage de programmation, y compris les langages compilés ou interprétés, et le programme d'ordinateur peut être déployé dans n'importe quelle forme, y compris en tant que programme autonome ou comme un sous-programme, élément ou autre unité appropriée pour une utilisation dans un environnement informatique. Un programme d'ordinateur peut être déployée pour être exécuté sur un ordinateur ou sur plusieurs ordinateurs à un seul site ou répartis sur plusieurs sites et reliées entre elles par un réseau de communication.

[0090] Le mode de réalisation préféré de la présente invention a été décrit. Diverses modifications peuvent être apportées sans s'écarter de l'esprit et de la portée de l'invention. Par conséquent, d'autres mises en œuvre sont dans la portée des revendications suivantes.

Revendications

1. Procédé mis en œuvre par ordinateur d'élaboration d'une fonction de transfert relative à la tête (S_1) adaptée à un individu, à partir d'une base de données (OH_1) comprenant des données (O_1) d'oreilles 3D ou 2D et des fonctions de transfert (H_1) correspondantes relatives à la tête, le procédé comprenant les étapes consistant à:

- effectuer une analyse statistique menant à une réduction de dimension (S_2) de l'espace des

oreilles 3D ou 2D, de la base de données (OH_1) et représenter chaque oreille 3D ou 2D par un vecteur de paramètres statistiques dont les valeurs des composantes sont les valeurs des projections de chaque oreille dans l'espace des oreilles de dimension réduite;

- effectuer une analyse statistique menant à une réduction de dimension (S_3) de l'espace des fonctions de transfert relatives à la tête, de la base de données (OH_1) et représenter chaque fonction de transfert par un vecteur de paramètres statistiques dont les valeurs des composantes sont les valeurs des projections de chaque fonction de transfert dans l'espace des fonctions de transfert de dimension réduite;

- effectuer une analyse des liens (S_4) entre lesdits paramètres statistiques de l'espace des oreilles 3D ou 2D et lesdits paramètres statistiques de l'espace des fonctions de transfert relatives à la tête; et

- déterminer (S_5), à partir de ladite analyse des liens et de ladite analyse statistique de l'espace des oreilles 3D ou 2D, une fonction de calcul (OH'_1) d'une fonction de transfert (S_1) relative à une tête à partir de données représentatives d'au moins une oreille

ledit procédé utilisant lesdites étapes divulguées pour élaborer une fonction de transfert (S_H), pour de hautes fréquences supérieures à un seuil, relative à la tête adaptée à l'individu, ledit procédé comprenant, en outre, une étape d'élaboration d'une fonction de transfert (S_B), pour de basses fréquences inférieures audit seuil, relative à la tête adaptée à l'individu,

ladite étape d'élaboration d'une fonction de transfert (S_B), pour de basses fréquences inférieures audit seuil, relative à la tête, adaptée à l'individu comprenant les sous-étapes suivantes, consistant à:

- échantillonner (S_9) des plages de valeurs possibles de paramètres morphologiques humains d'une base de données (M_1) relatives à la morphologie humaine,
- déterminer (S_{10}) un maillage de modèle paramétrique desdits paramètres morphologiques,
- calculer (S_{11}) des fonctions de transfert gabarit de basses fréquences, associées audit maillage,
- estimer (S_{12}) la valeur des paramètres morphologiques de l'individu à partir d'au moins une photo (U_2) de l'individu de face ou de profil, et
- calculer (S_{13}) une fonction de transfert (S_B), pour de basses fréquences, relative à la tête, adaptée à l'individu à partir de la valeur estimée des paramètres morphologiques et desdites fonctions calculées de transfert gabarit de bas-

ses fréquences.

2. Procédé selon la revendication 1, comprenant en outre une étape consistant à mettre en correspondance dense (S_1) des points relatifs à des positions respectives des oreilles de la base de données (OH_1);
3. Procédé selon la revendication 1 ou 2, comprenant, en outre, une étape de calcul (S_6) d'une fonction de transfert (S_1) relative à la tête, adaptée à l'individu, à partir de ladite fonction de calcul (OH'_1) et d'au moins une photographie (U_1) d'au moins une oreille de l'individu.
4. Procédé selon la revendication 3, dans lequel ladite étape de calcul (S_6) d'une fonction de transfert (S_1) relative à la tête est itérative.
5. Procédé selon la revendication 4, dans lequel ladite étape itérative de calcul d'une fonction de transfert relative à la tête comprend :
 - une première sous-étape itérative (S_7) d'estimation d'au moins un paramètre de pose de l'individu lors de la ou lesdites photographies; et
 - une deuxième sous-étape itérative (S_8) d'estimation de paramètres statistiques optimisés représentant au moins une oreille de l'individu dans l'espace des oreilles.
6. Procédé selon l'une des revendications précédentes, dans lequel lesdites données (O_1) représentant des oreilles sont des nuages de points.
7. Procédé selon la revendication 6, dans lequel une fonction de transfert relative à la tête (S_1) de l'individu est élaborée à partir desdites fonctions de transfert (S_H , S_B), respectivement pour de hautes et basses fréquences et de ladite ou lesdites photos (U_2) de l'individu de face ou de profil, comprenant les étapes consistant à:
 - estimer (S_{14}), à partir de ladite ou lesdites photos (U_2) de l'individu de face ou de profil, la taille d'oreilles relativement au reste du corps de l'individu;
 - mettre à l'échelle en fréquences (S_{15}) les fonctions de transfert relatives à la tête (S_H), pour les hautes fréquences; et
 - fusionner (S_{16}) les fonctions de transfert (S_H , S_B), respectivement pour de hautes et basses fréquences, pour obtenir la fonction de transfert relative à la tête (S_1) de l'individu.
8. Système d'élaboration d'une fonction de transfert relative à la tête adaptée à un individu, à partir d'une

base de données comprenant des données d'oreilles et des fonctions de transfert correspondantes relatives à la tête, comprenant un ordinateur configuré pour mettre en œuvre le procédé selon l'une des revendications précédentes.

Patentansprüche

1. Verfahren zur computergestützten Erstellung einer an ein Individuum angepassten kopfbezogenen Übertragungsfunktion (S_1) anhand einer Datenbank (OH_1), welche Daten (O_1) von 3D- oder 2D-Ohren und entsprechende kopfbezogene Übertragungsfunktionen (H_1) beinhaltet, wobei das Verfahren folgende Schritte beinhaltet:

- Durchführen einer statistischen Analyse der Datenbank (OH_1), welche zu einer dimensionalen Reduzierung (S_2) des Raumes der 3D- oder 2D-Ohren führt, und Darstellen eines jeden 3D- oder 2D-Ohres durch einen Statistikparameter-Vektor, deren Komponentenwerte die Werte von Projektionen eines jeden Ohres in den Raum der Ohren reduzierter Dimension sind;
- Durchführen einer statistischen Analyse der Datenbank (OH_1), welche zu einer Reduzierung der Dimension (S_3) des Raumes der kopfbezogenen Übertragungsfunktionen führt, und Darstellen einer jeden Übertragungsfunktion durch einen Statistikparameter-Vektor, deren Komponentenwerte die Werte von Projektionen einer jeden Übertragungsfunktion in den Raum der Übertragungsfunktionen mit reduzierter Dimension sind;
- Durchführen einer Analyse der Verbindungen (S_4) zwischen den statistischen Parametern des Raumes der 3D- oder 2D-Ohren und den statistischen Parametern des Raumes der kopfbezogenen Übertragungsfunktionen; und
- Bestimmen (S_5), anhand der Analyse der Verbindungen und der statistischen Analyse des Raumes der 3D- oder 2D-Ohren, einer Rechenfunktion (OH_1) einer kopfbezogenen Übertragungsfunktion (S_1) anhand von Daten, welche mindestens ein Ohr darstellen

wobei das Verfahren die offengelegten Schritte verwendet, um eine auf ein Individuum angepasste kopfbezogene Übertragungsfunktion (S_H), für hohe Frequenzen oberhalb eines Grenzwertes zu erstellen, wobei das Verfahren weiter einen Schritt des Entwickelns einer auf das Individuum angepassten kopfbezogenen Übertragungsfunktion (S_B) für niedrige Frequenzen unterhalb des Grenzwertes beinhaltet, wobei der Schritt des Erstellens einer auf das Individuum angepassten kopfbezogenen Übertra-

gungsfunktion (S_B) für niedrige Frequenzen unterhalb des Grenzwertes folgende Teilschritte beinhaltet:

- Abtasten (S_9) von möglichen Wertebereichen menschlicher morphologischer Parameter einer Datenbank (M_1), welche sich auf die menschliche Morphologie beziehen,
 - Bestimmen (S_{10}) eines Rasters eines parametrischen Modells der morphologischen Parameter,
 - Berechnen (S_{11}) der Schablonen-Übertragungsfunktionen für niedrige Frequenzen, welche dem Raster zugeordnet sind,
 - Schätzen (S_{12}) des Wertes der morphologischen Parameter des Individuums anhand mindestens einer Fotoaufnahme (U_2) des Individuums von vorn oder im Profil, und
 - Berechnen (S_{13}) einer auf das Individuum angepassten kopfbezogenen Übertragungsfunktion (S_B) für niedrige Frequenzen anhand des geschätzten Wertes der morphologischen Parameter und der berechneten Schablonen-Übertragungsfunktionen für niedrige Frequenzen.
2. Verfahren nach Anspruch 1, zudem beinhaltend einen Schritt des dichten Abstimmens (S_1) der Punkte der Datenbank (OH_1), welche sich auf jeweilige Positionen der Ohren beziehen.
 3. Verfahren nach Anspruch 1 oder 2, zudem beinhaltend einen Schritt des Berechnens (S_6) einer auf das Individuum angepassten kopfbezogenen Übertragungsfunktion (S_1) anhand der Berechnungsfunktion (OH_1) und mindestens einer Fotografie (U_1) von mindestens einem Ohr des Individuums.
 4. Verfahren nach Anspruch 3, bei welchem der Schritt des Berechnens (S_6) einer kopfbezogenen Übertragungsfunktion (S_1) ein iterativer Schritt ist.
 5. Verfahren nach Anspruch 4, bei welchem der iterative Schritt des Berechnens einer kopfbezogenen Übertragungsfunktion Folgendes beinhaltet:
 - einen ersten iterativen Teilschritt (S_7) des Schätzens mindestens eines Posen-Parameters des Individuums bei der oder den Fotoaufnahmen; und
 - einen zweiten iterativen Teilschritt (S_8) des Schätzens optimierter statistischer Parameter, welche mindestens ein Ohr des Individuums im Raum der Ohren darstellen.
 6. Verfahren nach einem der vorhergehenden Ansprüche, bei welchem die Daten (O_1), welche Ohren darstellen, Punktwolken sind.

7. Verfahren nach Anspruch 6, bei welchem eine kopfbezogene Übertragungsfunktion (S_1) des Individuums anhand der Übertragungsfunktionen (S_H , S_B) jeweils für hohe und niedrige Frequenzen und der Fotoaufnahme(n) (U_2) des Individuums von vorn oder im Profil erstellt wird, folgende Schritte beinhaltend:

- Schätzen (S14), anhand der Fotoaufnahme(n) (U_2) des Individuums von vorn oder im Profil, der Größe von Ohren in Bezug auf den Rest des Körpers des Individuums;
- Frequenz-Skalieren (S15) der kopfbezogenen Übertragungsfunktionen (S_H) für die hohen Frequenzen; und
- Verschmelzen (S16) der Übertragungsfunktionen (S_H , S_B), jeweils für hohe und niedrige Frequenzen, um die kopfbezogene Übertragungsfunktion (S_1) des Individuums zu erzeugen.

8. System zum Erstellen einer auf ein Individuum angepassten kopfbezogenen Übertragungsfunktion anhand einer Datenbank, welche Daten von Ohren beinhaltet, und der entsprechenden kopfbezogenen Übertragungsfunktionen, beinhaltend einen Rechner, welcher konfiguriert ist, um das Verfahren nach einem der vorhergehenden Ansprüche umzusetzen.

Claims

1. A computer-implemented method for generating an individual-specific head-related transfer function (S_1) from a database (OH_1) containing 3D or 2D ear data (O_1) and corresponding head-related transfer functions (H_1), the method comprising the steps of:

- performing a statistical analysis of the database (OH_1) leading to a reduction in the dimensionality (S_2) of the 3D or 2D ear space, and representing each 3D or 2D ear by a vector of statistical parameters the values of the components of which are the values of the projections of each ear into the ear space of reduced dimensionality;
- performing a statistical analysis of the database (OH_1) leading to a reduction in the dimensionality (S_3) of the head-related-transfer-function space, and representing each transfer function by a vector of statistical parameters the values of the components of which are the values of the projections of each transfer function into the transfer-function space of reduced dimensionality;
- performing an analysis (S4) of the relationships between said statistical parameters of the 3D or 2D ear space and said statistical parameters of the head-related-transfer-function space; and

- determining (S5), from said relationship analysis and said statistical analysis of the 3D or 2D ear space, a function (OH'_1) for calculating a head-related transfer function (S_1) from data representative of at least one ear;

said method using said disclosed steps to generate an individual-specific head-related transfer function (S_H) for high frequencies above a threshold, said method furthermore comprising a step of generating an individual-specific head-related transfer function (S_B) for low frequencies below said threshold; wherein said step of generating an individual-specific head-related transfer function (S_B) for low frequencies below said threshold comprises the following substeps of:

- sampling (S9) ranges of possible values of human morphological parameters from a database (M_1) of data relating to human morphology;
- defining (S10) a mesh of a parametric model of said morphological parameters;
- calculating (S11) low-frequency template transfer functions associated with said mesh;
- estimating (S12) the value of morphological parameters of the individual from at least one face-on or profile photograph (U_2) of the individual; and
- calculating (S13) an individual-specific head-related transfer function (S_B) for low frequencies from the estimated value of the morphological parameters and said calculated low-frequency template transfer functions.

2. The method according to claim 1, furthermore comprising a step consisting in densely matching (S1) points of the database (OH_1) relating to respective positions of the ears.

3. The method according to claim 1 or 2, furthermore comprising a step of calculating (S6) an individual-specific head-related transfer function (S_1) using said calculating function (OH'_1) and at least one photograph (U_1) of at least one ear of the individual.

4. The method according to claim 3, wherein said step of calculating (S6) a head-related transfer function (S_1) is iterative.

5. The method according to claim 4, wherein said iterative step of calculating a head-related transfer function comprises:

- a first iterative substep (S7) of estimating at least one postural parameter of the individual in said one or more photographs; and
- a second iterative substep (S8) of estimating optimized statistical parameters representing at

least one ear of the individual in the ear space.

6. The method according to one of the preceding claims, wherein said ear-representing data (O_1) are point clouds. 5
7. The method according to claim 6, wherein a head-related transfer function (S_1) of the individual is generated on the basis of said transfer functions (S_H , S_B) for high and low frequencies, respectively, and of said one or more face-on or profile photographs (U_2) of the individual, comprising the steps of: 10
 - estimating (S_{14}), from said one or more face-on or profile photographs (U_2) of the individual, ear size relative to the rest of the body of the individual; 15
 - frequency scaling (S_{15}) the head-related transfer functions (S_H), for the high frequencies; 20
 - and
 - fusing (S_{16}) the transfer functions (S_H , S_B) for high and low frequencies, respectively, in order to obtain the head-related transfer function (S_1) of the individual. 25
8. A system for generating an individual-specific head-related transfer function from a database containing ear data and corresponding head-related transfer functions, comprising a processor configured to implement the method according to one of the preceding claims. 30

35

40

45

50

55

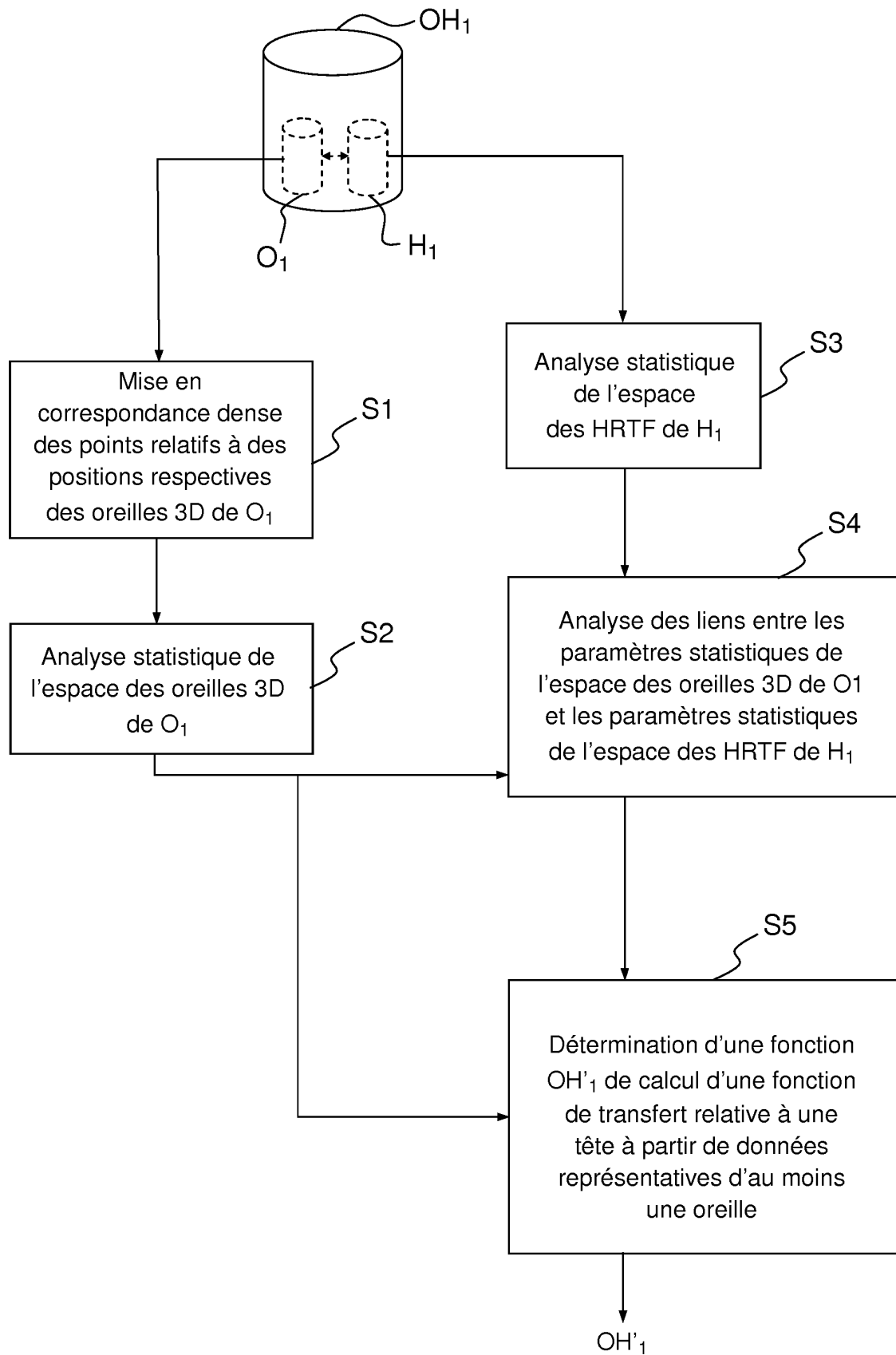


FIG.1

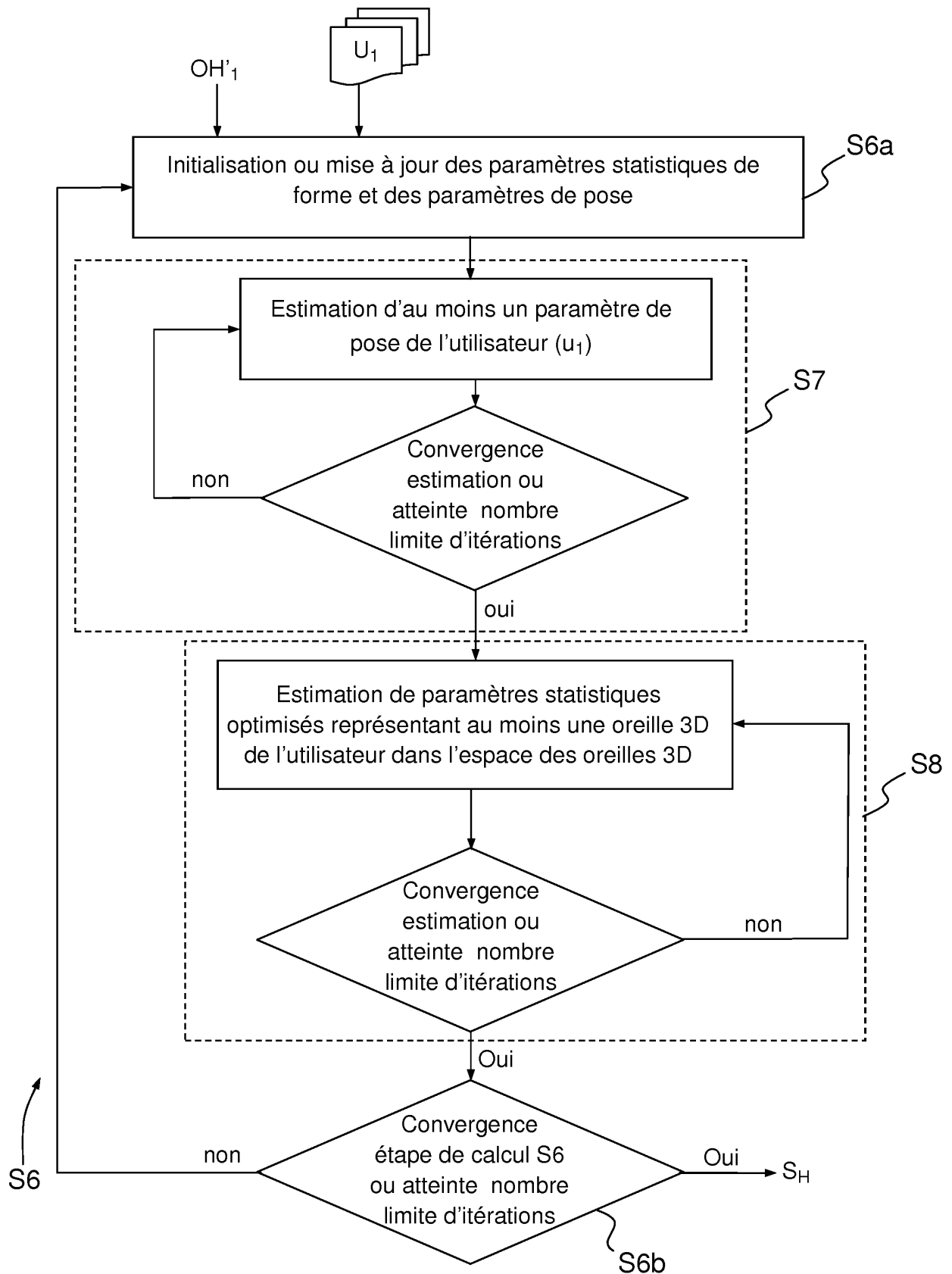


FIG.2

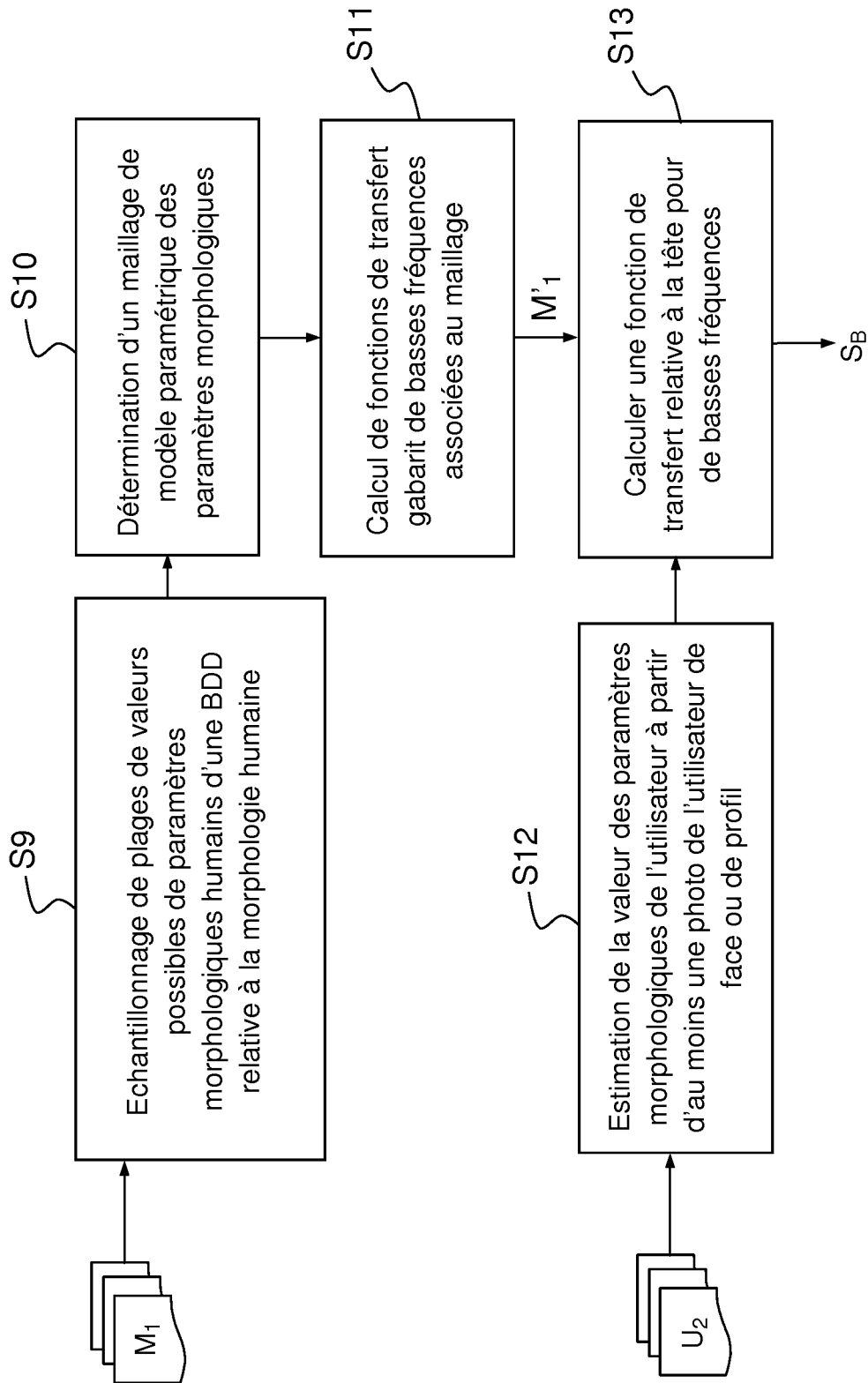


FIG.3

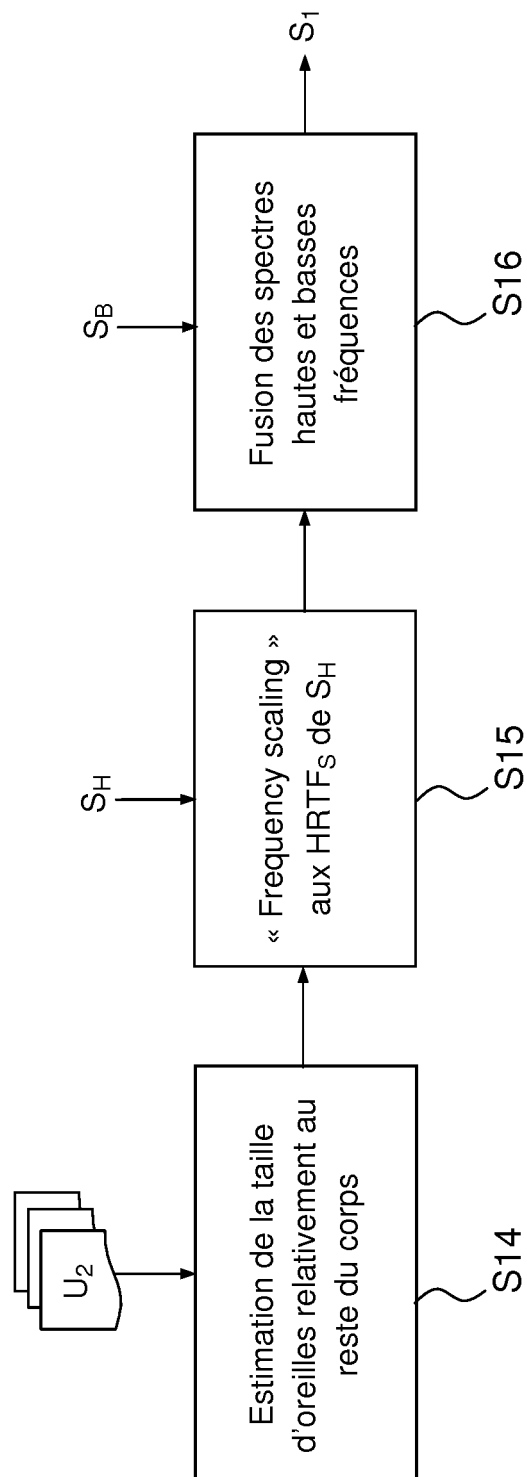


FIG.4

RÉFÉRENCES CITÉES DANS LA DESCRIPTION

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

Documents brevets cités dans la description

- US 7209564 B [0015]
- US 20060067548 A1 [0016]
- WO 2011128583 A [0029]

Littérature non-brevet citée dans la description

- **SYLVAIN BUSSON**. Individualisation d'Indices Acoustiques pour la Synthèse Binaurale. *PhD thesis, Université de la Méditerranée-Aix-Marseille II*, 2006 [0010]
- **KISTLER ; WIGHTMAN**. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *The Journal of the Acoustical Society of America*, 1992, vol. 91 (3), 1637-1647 [0012]
- **SONG XU ; ZHIZHONG LI ; GAVRIEL SALVENDY**. improved method to individualize head-related transfer function using anthropométrie measurements. *Acoustical Science and Technology*, 2008, vol. 29 (6), 388-390 [0013]
- **R. A. KENNEDY ; M. ZHANG ; T. D. ABHAYAPALA**. Statistical method to identify key anthropométrie parameters in hrtf individualization. *Joint Workshop on Hands-free Speech Communication and Microphone Arrays*, 2011 [0014]
- **PAUL M HOFMAN ; A JOHN VAN OPSTAL. BAYESIAN**. reconstruction of sound localization cues from responses to random spectra. *Biological cybernetics*, 2002, vol. 86 (4), 305-316 [0018]
- **SHOJI SHIMADA ; NOBUO HAYASHI ; SHINJI HAYASHI**. A clustering method for sound localization transfer functions. *Journal of the Audio Engineering Society*, 1994, vol. 42 (7/8), 577-584 [0022]
- Headphone virtualization : Improved localization and externalization of nonindividualized hrtfs by cluster analysis. **ROBERT P TAME ; DANIELE BARCHIESE ; ANSSI KLAPURI**. Audio Engineering Society Convention. Audio Engineering Society, Mai 2012, vol. 133 [0023]
- Improving binaural reproduction of 5.1 channel surround sound using individualized hrtf cluster in the wavelet domain. **BOSUN XIE ; ZHAOJUN TIAN**. Audio Engineering Society Conference : 55th International Conference : Spatial Audio. Audio Engineering Society, Août 2014 [0023]
- **YUKIO IWAYA**. Individualization of head-related transfer functions with tournament-style listening test : Listening with other's ears. *Acoustical science and technology*, 2006, vol. 27 (6), 340-343 [0024]
- **BERNHARD U SEEBER ; HUGO FASTL**. *Subjective selection of non-individual head-related transfer functions*, Juillet 2003 [0025]
- **WILLIAM L MARTENS**. Rapid psychophysical calibration using bisection scaling for individualized control of source elevation in auditory display. *Proc. Int. Conf. on Auditory Display*, Juillet 2002, 199-206 [0026]
- **D.N. ZOTKIN ; J. HWANG ; R. DURAIWAINI ; L.S. DAVIS**. Hrtf personalization using anthropométrie measurements. *Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop on*, Octobre 2003, 157-160 [0028]
- **JOHN C MIDDLEBROOKS**. Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency. *The Journal of the Acoustical Society of America*, 1999, vol. 106 (3), 1493-1510 [0031]
- **KATUHIRO MAKI ; SHIGETO FURUKAWA**. Reducing individual differences in the external-ear transfer functions of the mongolian gerbil. *The Journal of the Acoustical Society of America*, 2005, vol. 118 (4 [0032]