

12 DEMANDE DE BREVET D'INVENTION

A1

22 Date de dépôt : 30.05.22.

30 Priorité :

43 Date de mise à la disposition du public de la demande : 01.12.23 Bulletin 23/48.

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

60 Références à d'autres documents nationaux apparentés :

○ Demande(s) d'extension :

71 Demandeur(s) : ELNO Société par actions simplifiée (SAS) — FR.

72 Inventeur(s) : LACROIX Arthur, Henri, Albert Clément, Jean-Baptiste, DEXHEIMER Mathieu, Clément, Nicolas et GAIFFE Thierry, Pierre, François.

73 Titulaire(s) : ELNO Société par actions simplifiée (SAS).

74 Mandataire(s) : Lavoix.

54 Dispositif électronique et procédé de traitement, appareil acoustique et programme d'ordinateur associés.

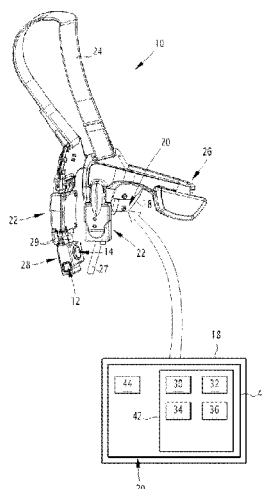
57 Dispositif électronique et procédé de traitement, appareil acoustique et programme d'ordinateur associés

Ce dispositif électronique de traitement (20) pour un appareil acoustique (10) comportant un premier microphone (12) aérien et un deuxième microphone (14) ostéophonique, est configuré pour être connecté aux premier et deuxième microphones (12,14), pour recevoir en entrée des premier, et respectivement deuxième, signaux analogiques issus des premier, et respectivement deuxième, microphones (12,14) et pour délivrer en sortie un signal corrigé.

Le dispositif de traitement (20) comprend :

- un module d'hybridation (30) configuré pour calculer un signal hybride à partir des premier et deuxième signaux analogiques ;
- un module d'estimation (32) configuré pour estimer un bruit dans le signal hybride ;
- un module de réduction de bruit (34) configuré pour calculer le signal corrigé en appliquant un algorithme de soustraction spectrale généralisée au signal hybride et en fonction du bruit estimé.

Figure pour l'abrégé : Figure 1



Description

Titre de l'invention : Dispositif électronique et procédé de traitement, appareil acoustique et programme d'ordinateur associés

- [0001] La présente invention concerne un dispositif électronique de traitement pour un appareil acoustique.
- [0002] L'invention concerne également un appareil acoustique comprenant un premier microphone comportant un transducteur électroacoustique apte à recevoir des ondes sonores acoustiques d'un signal sonore issu de cordes vocales d'un utilisateur et à transformer lesdites ondes acoustiques en un premier signal analogique ; un deuxième microphone comportant un transducteur à excitation mécanique osseuse apte à recevoir par conduction osseuse des oscillations vibratoires dudit signal sonore et à transformer lesdites oscillations vibratoires en un deuxième signal analogique ; et un tel dispositif électronique de traitement connecté aux premier et deuxième microphones, le dispositif de traitement étant configuré pour recevoir en entrée les premier et deuxième signaux analogiques, puis pour délivrer en sortie un signal corrigé.
- [0003] Le dispositif électronique de traitement comprend un module d'hybridation configuré pour calculer un signal hybride à partir des premier et deuxième signaux analogiques.
- [0004] L'invention concerne aussi un procédé de traitement mis en œuvre par un tel dispositif électronique de traitement ; ainsi qu'un programme d'ordinateur comportant des instructions logicielles qui, lorsqu'elles sont exécutées par un ordinateur, mettent en œuvre un tel procédé de traitement.
- [0005] On connaît du document FR 3 019 422 B1 un appareil acoustique du type précité. L'appareil acoustique comprend le premier microphone avec un tel transducteur électroacoustique, également appelé transducteur aérien ; le deuxième microphone avec un tel transducteur à excitation mécanique osseuse, également appelé transducteur solidien ; des moyens de calcul d'un signal électrique corrigé en fonction du premier signal électrique et du deuxième signal électrique, le signal électrique corrigé étant propre à être délivré en sortie de l'appareil acoustique ; et un dispositif de réduction du bruit connecté en sortie du transducteur électroacoustique pour réduire le bruit dans le premier signal électrique ; les moyens de calcul étant connectés, d'une part, en sortie du dispositif de réduction du bruit, et d'autre part, en sortie du transducteur à excitation mécanique osseuse.
- [0006] Toutefois, avec un tel appareil acoustique, la réduction de bruit n'est pas toujours optimale, et il subsiste parfois du bruit de fond relativement élevé dans le signal délivré en sortie de l'appareil acoustique.
- [0007] Le but de l'invention est alors de proposer un dispositif électronique de traitement, et

un procédé de traitement associé, permettant d'améliorer encore la réduction du bruit dans le signal délivré en sortie de l'appareil acoustique, c'est-à-dire de réduire la présence de bruit dans ledit signal.

- [0008] A cet effet, l'invention a pour objet un dispositif électronique de traitement pour un appareil acoustique,
- [0009] l'appareil acoustique comprenant un premier microphone comportant un transducteur électroacoustique apte à recevoir des ondes sonores acoustiques d'un signal sonore issu de cordes vocales d'un utilisateur et à transformer lesdites ondes acoustiques en un premier signal analogique ; et un deuxième microphone comportant un transducteur à excitation mécanique osseuse apte à recevoir par conduction osseuse des oscillations vibratoires dudit signal sonore et à transformer lesdites oscillations vibratoires en un deuxième signal analogique,
- [0010] le dispositif électronique de traitement étant configuré pour être connecté aux premier et deuxième microphones, pour recevoir en entrée les premier et deuxième signaux analogiques et pour délivrer en sortie un signal corrigé,
- [0011] le dispositif électronique de traitement comprenant :
- [0012] - un module d'hybridation configuré pour calculer un signal hybride à partir des premier et deuxième signaux analogiques ;
- [0013] - un module d'estimation connecté au module d'hybridation et configuré pour estimer un bruit dans le signal hybride ; et
- [0014] - un module de réduction de bruit connecté au module d'hybridation et au module d'estimation, le module de réduction de bruit étant configuré pour calculer le signal corrigé en appliquant un algorithme de soustraction spectrale généralisée au signal hybride et en fonction du bruit estimé.
- [0015] Avec le dispositif électronique de traitement selon l'invention, le fait d'estimer le bruit dans le signal hybride calculé à partir des premier et deuxième signaux analogiques, c'est-à-dire dans le signal hybride obtenu à partir des signaux issus d'une part du transducteur électroacoustique, ou aérien, et d'autre part du transducteur à excitation mécanique osseuse, également appelé transducteur ostéophonique, ou encore solidien, permet d'avoir une estimation plus précise du bruit, puis ensuite d'obtenir - via le module de réduction du bruit - un meilleur signal corrigé en appliquant l'algorithme de soustraction spectrale généralisée au signal de hybride et en fonction du bruit ainsi estimé.
- [0016] De préférence, le signal hybride comporte plusieurs tronçons successifs, chaque tronçon correspondant au signal hybride au cours d'une période temporelle, et le dispositif de traitement comporte en outre un module de détection d'activité vocale apte à déterminer si chaque tronçon du signal hybride comporte une présence de voix ou non, le module d'estimation étant alors configuré pour estimer le bruit dans le

signal hybride seulement à partir de chaque tronçon sans voix.

- [0017] La présence ou l'absence de voix est de préférence encore déterminée à partir du deuxième signal issu du transducteur ostéophonique, la présence ou l'absence de voix étant mieux détectable dans un signal provenant d'un microphone ostéophonique, plutôt que dans un signal provenant d'un microphone aérien.
- [0018] Suivant d'autres aspects avantageux de l'invention, le dispositif électronique de traitement comprend une ou plusieurs des caractéristiques suivantes, prises isolément ou suivant toutes les combinaisons techniquement possibles :
- [0019] - le signal hybride comporte plusieurs tronçons successifs, et le dispositif comprend en outre un module de détection d'activité vocale connecté au module d'hybridation et configuré pour déterminer une présence de voix ou une absence de voix dans chaque tronçon du signal hybride ; le module d'estimation étant alors configuré pour estimer le bruit dans le signal hybride en fonction de chaque tronçon avec une absence déterminée de voix ;
- [0020] - le module de détection d'activité vocale est configuré pour déterminer la présence de voix ou l'absence de voix à partir du deuxième signal issu du transducteur à excitation mécanique osseuse ;
- [0021] le module de détection d'activité vocale étant de préférence configuré pour déterminer la présence de voix ou l'absence de voix uniquement à partir du deuxième signal, sans prise en compte du premier signal ;
- [0022] - le deuxième signal comporte plusieurs tronçons successifs, et le module de détection d'activité vocale est configuré pour calculer une valeur RMS pour chaque tronçon du deuxième signal, puis pour déterminer la présence de voix ou l'absence de voix en fonction de valeur(s) RMS respective(s) ;
- [0023] - le module de détection d'activité vocale est configuré pour déterminer la présence de voix ou l'absence de voix en fonction d'une valeur moyenne de M dernière(s) valeur(s) RMS calculée(s) et/ou d'une variation de valeur RMS entre une valeur RMS courante et une valeur RMS précédente, M étant un nombre entier supérieur ou égal à 1 ;
- [0024] le module de détection d'activité vocale étant de préférence configuré pour déterminer la présence de voix si ladite valeur moyenne est supérieure ou égale à un seuil prédéfini de moyenne ou si ladite variation de valeur RMS est supérieure ou égale à un seuil prédéfini de variation ;
- [0025] - le module d'hybridation est configuré pour convertir le premier signal analogique en un premier signal numérique, au fur et à mesure de la réception du premier signal analogique, et pour générer des premiers tronçons successifs à partir du premier signal numérique, chaque nouveau premier tronçon généré comportant des échantillons d'un premier tronçon précédent et de nouveaux échantillons du premier signal numérique ;

et

- [0026] le module d'hybridation est configuré pour convertir le deuxième signal analogique en un deuxième signal numérique, au fur et à mesure de la réception du deuxième signal analogique, et pour générer des deuxièmes tronçons successifs à partir du deuxième signal numérique, chaque nouveau deuxième tronçon généré comportant des échantillons d'un deuxième tronçon précédent et de nouveaux échantillons du deuxième signal numérique ;
- [0027] des tronçons hybrides du signal hybride étant alors calculés au fur et à mesure à partir des premiers et deuxièmes tronçons générés ; le signal corrigé étant ensuite calculé à partir desdits tronçons hybrides ;
- [0028] - le module d'hybridation est configuré pour obtenir un premier signal filtré en appliquant au premier signal un premier filtre associé à une première plage de fréquences ; pour obtenir un deuxième signal filtré en appliquant au deuxième signal un deuxième filtre associé à une deuxième plage de fréquences ; puis pour calculer le signal hybride en sommant le premier signal filtré et le deuxième signal filtré, la deuxième plage de fréquences étant distincte de la première plage de fréquences ;
- [0029] la première plage de fréquences comportant de préférence des fréquences supérieures à celles de la deuxième plage de fréquences ;
- [0030] les première et deuxième plages de fréquences étant de préférence encore disjointes.
- [0031] L'invention concerne également un appareil acoustique comprenant :
- [0032] - un premier microphone comportant un transducteur électroacoustique apte à recevoir des ondes sonores acoustiques d'un signal sonore issu de cordes vocales d'un utilisateur et à transformer lesdites ondes acoustiques en un premier signal analogique ;
- [0033] - un deuxième microphone comportant un transducteur à excitation mécanique osseuse apte à recevoir par conduction osseuse des oscillations vibratoires dudit signal sonore et à transformer lesdites oscillations vibratoires en un deuxième signal analogique ;
- [0034] - un dispositif électronique de traitement connecté aux premier et deuxième microphones, le dispositif électronique de traitement étant configuré pour recevoir en entrée les premier et deuxième signaux analogiques, puis pour délivrer en sortie un signal corrigé ; le dispositif électronique de traitement étant tel que défini ci-dessus.
- [0035] Suivant un autre aspect avantageux de l'invention, l'appareil acoustique comprend en outre deux modules acoustiques latéraux en appui sur les flancs latéraux du crâne et propres à transmettre un signal sonore au nerf auditif.
- [0036] L'invention concerne aussi un équipement de tête pour opérateur comprenant un casque de protection, et un appareil acoustique tel que défini ci-dessus.
- [0037] L'invention a également pour objet un procédé de traitement, le procédé étant mis en œuvre par un dispositif électronique de traitement connecté à des premier et deuxième

microphones, le premier microphone comportant un transducteur électroacoustique apte à recevoir des ondes sonores acoustiques d'un signal sonore issu de cordes vocales d'un utilisateur et à transformer lesdites ondes acoustiques en un premier signal analogique ; et le deuxième microphone comportant un transducteur à excitation mécanique osseuse apte à recevoir par conduction osseuse des oscillations vibratoires dudit signal sonore et à transformer lesdites oscillations vibratoires en un deuxième signal analogique, le dispositif électronique de traitement étant configuré pour recevoir en entrée les premier et deuxième signaux analogiques et pour délivrer en sortie un signal corrigé,

[0038] le procédé de traitement comprenant :

[0039] - une étape d'hybridation comportant le calcul d'un signal hybride à partir des premier et deuxième signaux analogiques ;

[0040] - une étape d'estimation d'un bruit dans le signal hybride ; et

[0041] - une étape de réduction de bruit comportant le calcul du signal corrigé en appliquant un algorithme de soustraction spectrale généralisée au signal hybride et en fonction du bruit estimé.

[0042] L'invention concerne également un programme d'ordinateur comportant des instructions logicielles qui, lorsqu'elles sont exécutées par un ordinateur, mettent en œuvre un procédé de traitement tel que défini ci-dessus.

[0043] Ces caractéristiques et avantages de l'invention apparaîtront plus clairement à la lecture de la description qui va suivre, donnée uniquement à titre d'exemple non limitatif, et faite en référence aux dessins annexés, sur lesquels :

[0044] [Fig.1] la [Fig.1] est une vue d'ensemble en perspective d'un appareil acoustique selon l'invention, l'appareil acoustique comprenant un premier microphone aérien, un deuxième microphone ostéophonique, et un dispositif électronique de traitement à délivrer un signal électrique corrigé à partir des signaux électriques issus des premier et deuxième microphones ;

[0045] [Fig.2] la [Fig.2] est une représentation schématique sous forme d'un synoptique du dispositif de traitement de la [Fig.1], connecté au premier microphone aérien et au deuxième microphone ostéophonique ;

[0046] [Fig.3] la [Fig.3] est une représentation schématique d'une génération de tronçons chevauchés, effectuée par le dispositif de traitement de la [Fig.1] ;

[0047] [Fig.4] la [Fig.4] est un organigramme d'un procédé de traitement selon l'invention, le procédé étant mis en œuvre par le dispositif de traitement de la [Fig.1] ;

[0048] [Fig.5] la [Fig.5] est une vue représentant, en partie supérieure, un signal de voix bruité enregistré par un microphone aérien de l'état de la technique ; et en partie inférieure, un signal hybride obtenu avec les premier et deuxième microphones, et après réduction de bruit via le dispositif de traitement de la [Fig.1] ;

- [0049] [Fig.6] la [Fig.6] est une vue avec plusieurs courbes illustrant une détection d'activité vocale de l'état de la technique, via un microphone aérien et pour un seuil de détection bas ;
- [0050] [Fig.7] la [Fig.7] est une vue analogue à celle de la [Fig.6], pour un seuil de détection plus élevé ; et
- [0051] [Fig.8] la [Fig.8] est une vue analogue à celles des figures 6 et 7, illustrant une détection d'activité vocale selon l'invention, via un microphone ostéophonique.
- [0052] Dans la suite de la description, l'expression « sensiblement égal(e) à » définit une relation d'égalité à plus ou moins 20%, de préférence encore à plus ou moins 10%, de préférence encore à plus ou moins 5%.
- [0053] Sur la [Fig.1], un appareil acoustique 10 comprend un premier microphone 12, également appelé microphone aérien, apte à recevoir des ondes sonores acoustiques et à les transformer en un premier signal électrique, tel qu'un premier signal analogique, et un deuxième microphone 14, également appelé microphone ostéophonique ou encore microphone solidien, apte à recevoir par conduction osseuse des oscillations vibratoires et à les transformer en un deuxième signal électrique, tel qu'un deuxième signal analogique.
- [0054] L'appareil acoustique 10 comprend un boîtier de protection 18 et un dispositif de traitement 20 disposé à l'intérieur du boîtier de protection 18, le dispositif de traitement 20 étant connecté au premier microphone 12 et au deuxième microphone 14, et configuré pour recevoir en entrée les premier et deuxième signaux analogiques et délivrer en sortie un signal corrigé dans lequel le bruit a été réduit.
- [0055] En complément, l'appareil acoustique 10 comprend également deux modules acoustiques 22 latéraux, un arceau supérieur 24, un arceau arrière 26 de liaison des modules acoustiques et un câble de connexion 27, le câble de connexion 27 étant équipé à son extrémité d'un connecteur, non représenté. Les modules acoustiques latéraux 22, l'arceau supérieur 24, l'arceau arrière 26 et le câble de connexion 27 sont connus en soi, par exemple du document FR 3 019 422 B1.
- [0056] Le premier microphone 12 est connu, par exemple du document FR 3 019 422 B1, et comporte un transducteur électroacoustique, non représenté, apte à recevoir des ondes sonores acoustiques d'un signal sonore issu des cordes vocales et à transformer lesdites ondes acoustiques en le premier signal électrique. Le premier microphone 12 est connecté en entrée du dispositif de traitement 20.
- [0057] Le deuxième microphone 14 est également connu, par exemple du document FR 3 019 422 B1, et comporte un transducteur à excitation mécanique osseuse, non représenté, apte à recevoir par conduction osseuse, notamment à travers un os correspondant du crâne, les ondes vibratoires du signal sonore issu des cordes vocales de l'utilisateur et à le transformer en le deuxième signal électrique. Le transducteur à ex-

citation mécanique osseuse est également appelé transducteur ostéophonique, ou encore transducteur solidien. Le deuxième microphone 14 est aussi connecté en entrée du dispositif de traitement 20.

- [0058] Dans l'exemple de la [Fig.1], le premier microphone 12 et le deuxième microphone 14 ne sont pas disposés dans le boîtier de protection 18, mais sont disposés dans un boîtier additionnel 28, le boîtier additionnel 28 étant relié à l'un des deux modules acoustique 22 par deux bras de liaison 29. Le transducteur électroacoustique et le transducteur à excitation mécanique osseuse sont alors chacun disposés dans le boîtier additionnel 28. Ce boîtier additionnel 28 est de préférence destiné à être appliqué au contact du côté droit du crâne de l'utilisateur, et est alors de préférence relié au module acoustique 22 droit.
- [0059] En variante, comme illustré dans l'exemple de la figure 13 du document FR 3 019 422 B1, le deuxième microphone 14 n'est pas disposé dans le boîtier de protection 18, mais est disposé dans un autre boîtier additionnel, l'autre boîtier additionnel étant relié à l'un des deux modules acoustique 22 par deux bras de liaison. Le transducteur à excitation mécanique osseuse du deuxième microphone est alors disposé dans l'autre boîtier additionnel. Cet autre boîtier additionnel est de préférence destiné à être appliqué au contact du côté droit du crâne de l'utilisateur, et est alors de préférence relié au module acoustique 22 droit.
- [0060] En variante encore, comme illustré dans l'exemple de la [Fig.1] du document FR 3 019 422 B1, le premier microphone 12 comporte une protubérance, par exemple venue de matière avec le boîtier de protection 18. Selon cette variante, le deuxième microphone 14, en particulier son transducteur à excitation mécanique osseuse, est disposé à l'intérieur du boîtier de protection 18.
- [0061] Le dispositif électronique de traitement 20 comprend un module d'hybridation 30 connecté au premier microphone 12 et au deuxième microphone 14 ; un module d'estimation 32 connecté au module d'hybridation 30 ; et un module de réduction de bruit 34 connecté au module d'hybridation 30 et au module d'estimation 32, comme représenté sur la [Fig.2].
- [0062] En complément facultatif, le dispositif électronique de traitement 20 comprend en outre un module de détection d'activité vocale 36 connecté au module d'hybridation 30.
- [0063] Dans l'exemple de la [Fig.1], le dispositif électronique de traitement 20 comprend une unité de traitement d'informations 40 formée par exemple d'une mémoire 42 et d'un processeur 44 associé à la mémoire 42.
- [0064] Dans l'exemple de la [Fig.1], le module d'hybridation 30, le module d'estimation 32, le module de réduction de bruit 34, ainsi qu'en complément facultatif le module de détection d'activité vocale 36, sont réalisés chacun sous forme d'un logiciel, ou d'une

brique logicielle, exécutable par le processeur 44. La mémoire 42 du dispositif de traitement 20 est alors apte à stocker un logiciel d'hybridation des premier et deuxième signaux analogiques en un signal hybride, un logiciel d'estimation du bruit dans le signal hybride, et un logiciel de réduction du bruit dans le signal hybride, ainsi qu'en complément facultatif un logiciel de détection d'activité vocale dans le signal hybride. Le processeur 44 est alors apte à exécuter chacun des logiciels parmi le logiciel d'hybridation, le logiciel d'estimation et le logiciel de réduction de bruit, ainsi qu'en complément facultatif le logiciel de détection d'activité vocale.

- [0065] En variante non représentée, le module d'hybridation 30, le module d'estimation 32, le module de réduction de bruit 34, ainsi qu'en complément facultatif le module de détection d'activité vocale 36, sont réalisés chacun sous forme d'un composant logique programmable, tel qu'un FPGA (de l'anglais *Field Programmable Gate Array*), ou encore d'un circuit intégré, tel qu'un ASIC (de l'anglais *Application Specific Integrated Circuit*).
- [0066] Lorsque le dispositif électronique de traitement 20 est réalisé sous forme d'un ou plusieurs logiciels, c'est-à-dire sous forme d'un programme d'ordinateur, également appelé produit programme d'ordinateur, il est en outre apte à être enregistré sur un support, non représenté, lisible par ordinateur. Le support lisible par ordinateur est par exemple un medium apte à mémoriser des instructions électroniques et à être couplé à un bus d'un système informatique. A titre d'exemple, le support lisible est un disque optique, un disque magnéto-optique, une mémoire ROM, une mémoire RAM, tout type de mémoire non-volatile (par exemple EPROM, EEPROM, FLASH, NVRAM), une carte magnétique ou une carte optique. Sur le support lisible est alors mémorisé un programme d'ordinateur comprenant des instructions logicielles.
- [0067] Le module d'hybridation 30 est configuré pour calculer le signal hybride à partir des premier et deuxième signaux analogiques.
- [0068] Le module d'hybridation 30 est par exemple configuré pour obtenir un premier signal filtré en appliquant au premier signal un premier filtre associé à une première plage de fréquences ; pour obtenir un deuxième signal filtré en appliquant au deuxième signal un deuxième filtre associé à une deuxième plage de fréquences ; puis pour calculer le signal hybride en sommant le premier signal filtré et le deuxième signal filtré, la deuxième plage de fréquences étant distincte de la première plage de fréquences.
- [0069] La première plage de fréquences comporte typiquement des fréquences supérieures à celles de la deuxième plage de fréquences ; les première et deuxième plages de fréquences étant par exemple disjointes.
- [0070] Le premier filtre est typiquement un filtre passe-haut avec une fréquence de coupure f_c sensiblement égale à 1000 Hz, le filtre passe-haut étant par exemple un filtre passe-haut gaussien. Le deuxième filtre est typiquement un filtre passe-bas avec une

fréquence de coupure également sensiblement égale à 1000 Hz, le filtre passe-bas étant par exemple un filtre passe-bas gaussien. Autrement dit, la première plage de fréquences est alors la plage des fréquences supérieures à 1000 Hz, et la deuxième plage de fréquence est celle des fréquences inférieures à 1000 Hz.

- [0071] En complément, le module d'hybridation 30 est configuré pour convertir le premier signal analogique en un premier signal numérique, au fur et à mesure de la réception du premier signal analogique, et pour générer des premiers tronçons successifs à partir du premier signal numérique.
- [0072] Selon ce complément, le module d'hybridation 30 est également configuré pour convertir le deuxième signal analogique en un deuxième signal numérique, au fur et à mesure de la réception du deuxième signal analogique, et pour générer des deuxièmes tronçons successifs à partir du deuxième signal numérique.
- [0073] Selon ce complément facultatif, le module d'hybridation 30 est alors configuré pour calculer des tronçons hybrides du signal hybride au fur et à mesure, à partir des premiers et deuxièmes tronçons générés ; le signal corrigé étant ensuite calculé à partir desdits tronçons hybrides.
- [0074] Dans l'exemple de la [Fig.2], le module d'hybridation 30 comporte un premier convertisseur analogique-numérique 50, connecté au premier microphone 12 aérien et configuré pour convertir le premier signal analogique issu du premier microphone 12 en un premier signal numérique x_k^{aer} , avec une fréquence d'échantillonnage f_e par exemple sensiblement égale à 22 kHz. En complément, le premier convertisseur analogique-numérique 50 est configuré pour découper le premier signal numérique x_k^{aer} , converti et échantillonné, en premiers tronçons successifs, chaque premier tronçon comportant par exemple un nombre N d'échantillons. Le nombre N d'échantillons dans chaque premier tronçon est par exemple sensiblement égal à 512. L'homme du métier observera alors qu'avec la fréquence d'échantillonnage f_e sensiblement égale à 22 kHz et le nombre N d'échantillons sensiblement égal à 512, la durée de chaque premier tronçon est d'environ 20 ms, et typiquement sensiblement égale à 23 ms.
- [0075] Dans l'exemple de la figure 2, le module d'hybridation 30 comporte en outre un premier convertisseur temporel-fréquentiel 52, connecté en sortie du premier convertisseur analogique-numérique 50 et configuré pour calculer un premier spectre $\tilde{X}_k^{aer}$ du premier signal numérique x_k^{aer} , typiquement via une transformée de Fourier, telle qu'une transformée de Fourier rapide, également notée FFT (de l'anglais *Fast Fourier Transform*). Le module hybridation 30 comporte ensuite une première unité de filtrage 54, connectée en sortie du premier convertisseur temporel-fréquentiel 52 et configurée pour appliquer le premier filtre, typiquement le filtre passe-haut gaussien de fréquence de coupure f_c sensiblement égale à 1000 Hz, pour obtenir le premier signal filtré $\tilde{X}_k^{aerHF}$.

- [0076] Dans l'exemple de la [Fig.2], le module d'hybridation 30 comporte un deuxième convertisseur analogique-numérique 60, connecté au deuxième microphone 14 ostéophonique et configuré pour convertir le deuxième signal analogique issu du deuxième microphone 14 en un deuxième signal numérique x_k^{ost} , avec la fréquence d'échantillonnage f_e . En complément, le deuxième convertisseur analogique-numérique 60 est configuré pour découper le deuxième signal numérique x_k^{ost} , converti et échantillonné, en deuxièmes tronçons successifs, chaque deuxième tronçon comportant par exemple le nombre N d'échantillons. L'homme du métier observera alors qu'avec la fréquence d'échantillonnage f_e sensiblement égale à 22 kHz et le nombre N d'échantillons sensiblement égal à 512, la durée de chaque deuxième tronçon est d'environ 20 ms, et typiquement sensiblement égale à 23 ms.
- [0077] Dans l'exemple de la figure 2, le module d'hybridation 30 comporte en outre un deuxième convertisseur temporel-fréquentiel 62, connecté en sortie du deuxième convertisseur analogique-numérique 60 et configuré pour calculer un deuxième spectre $\tilde{X}_k^{ost}$ du deuxième signal numérique x_k^{ost} , typiquement via une transformée de Fourier, telle que la transformée de Fourier rapide, ou FFT. Le module hybridation 30 comporte ensuite une deuxième unité de filtrage 64, connectée en sortie du deuxième convertisseur temporel-fréquentiel 62 et configurée pour appliquer le deuxième filtre, typiquement le filtre passe-bas gaussien de fréquence de coupure f_c sensiblement égale à 1000 Hz, pour obtenir le deuxième signal filtré $\tilde{X}_k^{ostBF}$.
- [0078] Par convention, dans la présente description, pour un signal noté x, sa forme continue dans le temps est notée x(t), et sa forme discrétisée est notée x[n] où n est un entier naturel, n formant alors une variable représentant le temps discrétisé. Dans le domaine fréquentiel, m représente la variable de fréquence discrète, comprise entre 0 et N/2, où N représente le nombre d'échantillons par tronçon, par exemple égal à 512.
- [0079] La forme discrétisée de chaque signal vérifie alors l'équation suivante :
- [0080] [Math.1]
- $$x[n] = x(n \times T_e)$$
- [0081] où n est la variable entière représentant le temps discrétisé, et
- [0082] T_e est un pas de discrétisation temporelle vérifiant l'équation suivante :
- [0083] [Math.2]
- $$T_e = \frac{1}{f_e}$$
- [0084] où f_e est la fréquence d'échantillonnage, par exemple sensiblement égale à 22 kHz.
- [0085] La variable de fréquence discrète m est typiquement associée à un vecteur fréquence f[m] vérifiant l'équation suivante :

[0086] [Math.3]

$$f[m] = m \times \frac{f_c}{N}$$

[0087] où N est le nombre d'échantillons compris dans un tronçon,

[0088] m est la variable de fréquence discrète, et

[0089] f_c est la fréquence d'échantillonnage.

[0090] La fréquence varie alors typiquement entre 0 Hz et $f_c/2$ Hz, avec un pas fréquentiel égal à f_c/N .

[0091] Par convention, le $k^{\text{ème}}$ tronçon du signal x est noté x_k ou $x_k[n]$, et $\tilde{X}_k[m]$ dans le domaine fréquentiel avec :

[0092] [Math.4]

$$\tilde{X}_k[m] = FFT(x_k[n])$$

[0093] où FFT représente l'opérateur numérique permettant d'estimer la transformée de Fourier discrète d'un signal, par exemple mis en œuvre via le convertisseur temporel-fréquentiel 52, 62 respectif.

[0094] La soustraction spectrale décrite par la suite ne nécessite de travailler que sur le spectre en amplitude du signal, la phase étant conservée et inchangée tout au long du processus, avec $|\tilde{X}_k[m]|$ représentant le spectre en amplitude et $\varphi(\tilde{X}_k[m])$ représentant le spectre en phase de $x_k[n]$ respectivement. Par convention, le spectre sans autre précision désignera alors par la suite le spectre en amplitude.

[0095] Dans l'exemple de la figure 2, le module hybridation 30 comporte également un sommateur 70, également appelé additionneur, connecté en sortie d'une part de la première unité de filtrage 54, et d'autre part de la deuxième unité de filtrage 64, et configuré pour sommer le premier signal filtré $\tilde{X}_k^{aerHF}$ et le deuxième signal filtré $\tilde{X}_k^{ostBF}$ afin d'obtenir le signal hybride $\tilde{X}_k^{hyb}$.

[0096] Le module hybridation 30 est alors par exemple configuré pour calculer le signal hybride $\tilde{X}_k^{hyb}$ en sommant le premier signal filtré $\tilde{X}_k^{aerHF}$ et le deuxième signal filtré $\tilde{X}_k^{ostBF}$ via l'équation suivante :

[0097] [Math.5]

$$\tilde{X}_k^{hyb} = \alpha \tilde{X}_k^{aerHF} + \beta \tilde{X}_k^{ostBF}$$

[0098] où α et β sont des constantes.

[0099] Les valeurs des constantes α et β sont de préférence réglables permettant d'avoir un signal de sortie au niveau équivalent à celui d'entrée du premier microphone 12 aérien. En outre, cela permet de donner une éventuelle prépondérance au signal aérien, ou respectivement au signal ostéophonique.

- [0100] En complément facultatif, le module d'hybridation 30 est configuré, lors de la génération des premiers tronçons successifs, pour générer chaque nouveau premier tronçon avec des échantillons d'un premier tronçon précédent et de nouveaux échantillons du premier signal numérique.
- [0101] Selon ce complément facultatif, le module d'hybridation 30 est configuré de manière analogue, lors de la génération des deuxièmes tronçons successifs, pour générer chaque nouveau deuxième tronçon avec des échantillons d'un deuxième tronçon précédent et de nouveaux échantillons du deuxième signal numérique.
- [0102] Il y a alors un chevauchement entre les premiers tronçons successifs ainsi générés, c'est-à-dire d'un premier tronçon généré au suivant ; et de manière analogue entre les deuxièmes tronçons successifs ainsi générés, c'est-à-dire d'un deuxième tronçon généré au suivant.
- [0103] Un taux de chevauchement correspond alors à un ratio, au sein de chaque nouveau premier tronçon, entre le nombre d'échantillons du premier tronçon précédent utilisés et le nombre total d'échantillons du premier tronçon, c'est-à-dire du nouveau premier tronçon généré ; ou respectivement au ratio, au sein de chaque nouveau deuxième tronçon, entre le nombre d'échantillons du deuxième tronçon précédent utilisés et le nombre total d'échantillons du deuxième tronçon. Le taux de chevauchement est par exemple compris entre 50 % et 75 %, c'est-à-dire entre 0,5 et 0,75. Autrement dit, au sein de chaque nouveau premier tronçon, entre la moitié et trois-quarts des derniers échantillons du premier tronçon précédent sont utilisés ; et de manière analogue au sein de chaque nouveau deuxième tronçon, entre la moitié et trois-quarts des derniers échantillons du deuxième tronçon précédent sont utilisés. Ce chevauchement entre tronçons est illustré à la [Fig.3].
- [0104] Sur la [Fig.3], les tronçons qui seraient obtenus par un simple découpage (i.e. sans chevauchement) du signal issu du premier convertisseur analogique-numérique 50, respectivement du deuxième convertisseur analogique-numérique 60, sont notés x_i , qu'il s'agisse des premiers ou des deuxièmes tronçons, où i est un indice prenant les valeurs successives $k-2$, $k-1$ et k dans cet exemple. Ces tronçons x_i qui seraient obtenus par simple découpage et sans chevauchement sont également appelés tronçons physiques. Les autres tronçons, représentés à la [Fig.3] et illustrant ce chevauchement, sont également appelés tronçons chevauchés et notés x'_i , avec i égal à $k-1$ ou k dans cet exemple.
- [0105] Dans l'exemple de la [Fig.3], l'homme du métier observera que le taux de chevauchement est sensiblement égal à 50 %, et que le tronçon x'_{k-1} comporte alors 50 % d'échantillons issus du tronçon précédent, correspondant à la dernière moitié du tronçon x_{k-2} dans cet exemple ; et 50 % de nouveaux échantillons, correspondant à la

première moitié du tronçon x_{k-1} dans cet exemple.

[0106] Sur la [Fig.3], les tronçons obtenus après réduction de bruit par le module de réduction de bruit 34 sont notés y_i lorsqu'ils résultent de tronçons physiques x_i , et respectivement y'_i lorsqu'ils résultent de tronçons chevauchés x'_i , avec i égal à $k-1$ ou k dans cet exemple.

[0107] Dans le cas d'un chevauchement à 50% le tronçon de sortie y_k^{out} vérifie alors typiquement l'équation suivante :

[0108] [Math.6]

$$y_k^{out} = \frac{1}{2}y'_{k-1} \left[\frac{N}{2} : N \right] + \frac{1}{2}y_{k-1} [0 : N] + \frac{1}{2}y'_k \left[0 : \frac{N}{2} \right]$$

[0109] où N représente le nombre d'échantillons par tronçon, par exemple égal à 512,

[0110] y_i représente un tronçon obtenu après réduction de bruit à partir d'un tronçon physique x_i , et

[0111] y'_i représente un tronçon obtenu après réduction de bruit à partir d'un tronçon chevauché x'_i .

[0112] Le module d'estimation 32 est configuré pour estimer un bruit dans le signal hybride.

[0113] Lorsqu'en complément facultatif, le module de détection d'activité vocale 36 est configuré pour déterminer une présence de voix ou une absence de voix dans chaque tronçon du signal hybride, le module d'estimation 32 est alors configuré pour estimer le bruit dans le signal hybride en fonction de chaque tronçon avec une absence déterminée de voix.

[0114] Autrement dit, lorsque le module de détection d'activité vocale 36 détermine une présence de voix dans un tronçon donné, le spectre du bruit n'est pas mis à jour. Au contraire, lorsque le module de détection d'activité vocale 36 détermine une absence de voix dans un tronçon donné, le spectre du bruit de fond est mis à jour. Cette mise à jour du spectre du bruit de fond est alors effectuée lorsque le tronçon n'est pas de la voix et que la probabilité que cela soit du bruit est élevée. La robustesse du module de détection d'activité vocale 36 permettra d'avoir autant plus de précision sur l'estimation et la poursuite du bruit.

[0115] Selon ce complément facultatif, le module d'estimation 32 est typiquement configuré pour mettre à jour le spectre du bruit de fond $|\tilde{N}_k|$ selon l'équation suivante :

[0116] [Math.7]

$$\begin{cases} |\tilde{N}_k| = p \times |\tilde{N}_{k-1}| + (1-p) \times |\tilde{X}_k^{hyb}| & \text{si } DAV = 0 \\ |\tilde{N}_k| = |\tilde{N}_{k-1}| & \text{si } DAV = 1 \end{cases}$$

[0117] où p est un facteur d'oubli, de valeur par exemple égale à 0,95 ;

[0118] DAV est un indicateur d'activité vocale issu du module de détection d'activité vocale 36, DAV étant égal à 1 si une présence de voix est déterminée, et à 0 sinon, i.e. si une

absence de voix est déterminée ;

[0119] $|\tilde{X}_k^{hyb}|$ représente le spectre du signal hybride $\tilde{X}_k^{hyb}$

[0120] $|\tilde{N}_{k-1}|$, et resp. $|\tilde{N}_k|$, représentent les spectres du bruit de fond pour le tronçon d'indice k-1, et resp. d'indice k.

[0121] Le module de réduction de bruit 34 est configuré pour calculer le signal corrigé en appliquant un algorithme de soustraction spectrale généralisée au signal hybride et en fonction du bruit estimé.

[0122] Dans l'exemple de la [Fig.2], le module de réduction de bruit 34 comporte une unité de soustraction spectrale généralisée 80, également appelée unité SSG 80, apte à mettre en œuvre l'algorithme de soustraction spectrale généralisée.

[0123] L'algorithme de soustraction spectrale généralisée vérifie par exemple l'équation suivante :

[0124] [Math.8]

$$\begin{cases} |\tilde{Y}_k[m]|^\gamma = |\tilde{X}_k[m]|^\gamma - \alpha_k[m] \times \delta[m] \times |\tilde{N}_k[m]|^\gamma & \text{si } |\tilde{X}_k[m]|^\gamma - \alpha_k[m] \times \delta[m] \times |\tilde{N}_k[m]|^\gamma \geq \beta |\tilde{N}_k[m]|^\gamma \\ |\tilde{Y}_k[m]|^\gamma = \beta |\tilde{N}_k[m]|^\gamma & \text{sinon} \end{cases}$$

[0125] $|\tilde{Y}_k[m]|$ représente le spectre du signal débruité pour le tronçon d'indice k ;

[0126] $|\tilde{X}_k[m]|$ représente le spectre du signal hybride pour le tronçon d'indice k ;

[0127] $|\tilde{N}_k[m]|$ représente le spectre du bruit de fond pour le tronçon d'indice k ;

[0128] α_k représente un coefficient de surestimation du bruit pour le tronçon d'indice k ;

[0129] δ représente un coefficient de correction ;

[0130] β représente un coefficient de réintroduction du bruit ; et

[0131] γ représente un coefficient de puissance, typiquement égal à 1 ou 2.

[0132] L'algorithme de soustraction spectrale généralisée se calcule par exemple en amplitude, et le coefficient de puissance γ est alors égal à 1 ; ou encore en puissance, et le coefficient de puissance γ est alors égal à 2.

[0133] Dans le cas d'un calcul en amplitude de la soustraction spectrale généralisée, avec $\gamma=1$, peu de bruit musical sera produit, mais le signal de voix estimé pourra être plus ou moins distordu en fonction du rapport signal sur bruit. Le bruit musical est un ensemble d'artefacts produits lors de la soustraction spectrale, constitué de tonales courtes en temps et produisant un bruit relativement désagréable.

[0134] Dans le cas d'un calcul en puissance de la soustraction spectrale généralisée, avec $\gamma=2$, peu de distorsion sera créée, mais une quantité non négligeable de bruit musical pourra être générée.

[0135] Le coefficient de surestimation de bruit α est de préférence recalculé à chaque tronçon d'indice k, et est alors noté α_k . Ce coefficient permet d'éviter la génération d'une quantité trop importante de bruit musical. Pour maximiser son efficacité, son calcul s'effectue par bandes de fréquences et dépend du rapport signal sur bruit sur

chacune de ces bandes.

- [0136] Les spectres $|\tilde{X}_k[m]|$ et $|\tilde{N}_k[m]|$ sont d'abord découpés en sous-spectres notés $|\tilde{X}_k^j[m]|$ et $|\tilde{N}_k^j[m]|$, où j représente le numéro de la bande de fréquence. Ainsi, j valeurs du rapport signal sur bruit, notées RSB_k^j , chacune associée à une bande de fréquence d'indice j , sont typiquement calculées selon l'équation suivante :

- [0137] [Math.9]

$$RSB_k^j = 10 \times \log_{10} \left(\frac{\sum_{m=0}^{N_j} |\tilde{X}_k^j[m]|^2}{\sum_{m=0}^{N_j} |\tilde{N}_k^j[m]|^2} \right)$$

- [0138] où RSB_k^j représente le rapport signal sur bruit pour le tronçon d'indice k et la bande de fréquence d'indice j ,

- [0139] N_j représente le nombre d'échantillons fréquentiels contenus dans la bande d'indice j ;

- [0140] $|\tilde{X}_k[m]|$ représente le spectre du signal hybride pour le tronçon d'indice k ; et

- [0141] $|\tilde{N}_k[m]|$ représente le spectre du bruit de fond pour le tronçon d'indice k .

- [0142] Puis, pour chaque valeur de rapport signal sur bruit, le coefficient de surestimation du bruit α_k vérifie par exemple l'équation suivante :

- [0143] [Math.10]

$$\begin{cases} \alpha_k^j = 4.75 \text{ si } RSB_k^j < -5 \text{ dB} \\ \alpha_k^j = 4 - \frac{3}{20} \times RSB_k^j \text{ si } -5 \text{ dB} \leq RSB_k^j \leq 20 \text{ dB} \\ \alpha_k^j = 1 \text{ si } RSB_k^j > 20 \text{ dB} \end{cases}$$

- [0144] Globalement, ce calcul du coefficient de surestimation de bruit α permet de surestimer le bruit lorsque le rapport signal sur bruit est faible, et de réduire l'introduction d'artefacts de type bruit musical.

- [0145] Le coefficient de surestimation du bruit α_k^j est ensuite converti pour pouvoir être ré-introduit dans l'équation (8), par exemple selon l'équation suivante :

- [0146] [Math.11]

$$\alpha_k[m] = \alpha_k^j \forall m \in [f_j ; f_{j+1}]$$

- [0147] où l'intervalle $[f_j ; f_{j+1}]$ correspond à toutes les fréquences de la $j^{\text{ème}}$ bande de fréquences. Typiquement, à chaque tronçon la fonction $\alpha_k[m]$ sera une fonction constante par morceaux, où chaque morceau correspondra à une bande de fréquences déterminée par l'utilisateur.

- [0148] Le coefficient de correction δ est un coefficient de correction fréquentiel calculé une seule fois, typiquement au début de l'algorithme, et n'évoluant pas au cours du temps.

- [0149] Ce coefficient est un simple pré-facteur dépendant de la fréquence, afin de maximiser certaines bandes de fréquences de manière adaptée à la captation de voix.
- [0150] Le coefficient de correction δ est par exemple une fonction constante par morceaux, vérifiant l'équation suivante :
- [0151] [Math.12]
- $$\begin{cases} \delta[m] = 1 \quad \forall f[m] < 1000 \text{ Hz} \\ \delta[m] = 2.5 \quad \forall f[m] \in [1000, 2000[\text{ Hz} \\ \delta[m] = 1.5 \quad \forall f[m] \in [2000, 4000[\text{ Hz} \\ \delta[m] = 1 \quad \forall f[m] \geq 4000 \text{ Hz} \end{cases}$$
- [0152] Compte tenu des calculs effectués avec les spectres en amplitude, il ne faut pas que l'estimation $|\tilde{Y}_k[m]|^p$ soit négative car cela n'aurait pas de sens mathématiquement. C'est pourquoi l'équation (8) comporte une condition pour éviter les valeurs négatives.
- [0153] Le coefficient de réintroduction du bruit β permet alors de choisir si l'on réintroduit du bruit ou non en cas de valeurs potentiellement négatives. Lorsque le coefficient de réintroduction du bruit β est choisi égal à 0, toute soustraction conduisant à une valeur négative est remplacée par la valeur nulle. En revanche pour toute valeur supérieure à 0, on réintroduit du bruit. Cela permet de conserver une partie du bruit qui peut être perçu comme un bruit de confort masquant une partie du bruit musical lorsqu'il y en a qui est créé.
- [0154] Le coefficient de réintroduction du bruit β vaut généralement quelques pourcents. Le coefficient de réintroduction du bruit β est par exemple sensiblement égal à 0,05, soit une réintroduction de 5% du bruit de fond dans le signal de sortie. Cette valeur est un paramètre prédéfini.
- [0155] Il est à noter que plus le rapport signal sur bruit est faible ou mauvais, moins l'estimation du signal débruité est efficace et plus la voix sera altérée. Il est donc intéressant de mettre une valeur du coefficient de réintroduction du bruit β plus élevée dans le cas d'un mauvais rapport signal sur bruit, afin de recapter quelques harmoniques de la voix dans le bruit de fond qui seraient perdues dans la soustraction spectrale autrement.
- [0156] Dans l'exemple de la [Fig.2], le module de réduction de bruit 34 comporte en outre un convertisseur fréquentiel-temporel 82, connecté en sortie de l'unité de soustraction spectrale généralisée 80, et configuré pour calculer un signal temporel à partir du signal fréquentiel issu de l'unité SSG 80, typiquement via une transformée de Fourier inverse, telle qu'une transformée de Fourier rapide inverse, également notée IFFT (de l'anglais *Inverse Fast Fourier Transform*).
- [0157] Comme indiqué précédemment, les calculs dans le domaine fréquentiel ont été effectués avec l'amplitude du spectre du signal du tronçon. La phase de celui-ci, qui

demeure non modifiée, est alors réintégrée au signal avant la transformée de Fourier inverse permettant de revenir dans le domaine temporel, par exemple selon l'équation suivante :

[0158] [Math.13]

$$y_k[n] = IFFT\left(|\tilde{Y}_k[m]| \times e^{j\varphi(\tilde{X}_k[m])}\right)$$

[0159] où $y_k[n]$ représente le signal de sortie débruité pour le tronçon d'indice k ;

[0160] IFFT représente l'opérateur numérique de transformée de Fourier inverse ;

[0161] $|\tilde{Y}_k[m]|$, et resp. $\varphi(\tilde{X}_k[m])$, représentent le spectre en amplitude, et resp. en phase, du signal débruité pour le tronçon d'indice k .

[0162] Dans l'exemple de la [Fig.2], le module de réduction de bruit 34 comporte ensuite un convertisseur numérique-analogique 84, connecté en sortie du convertisseur fréquentiel-temporel 82 et configuré pour fournir le signal corrigé $y(t)$ sous forme analogique. Le signal débruité y_k^{hyb} issu du convertisseur fréquentiel-temporel 82 est alors resynthétisé en le signal corrigé $y(t)$ via le convertisseur numérique-analogique 84, avec synthèse des tronçons chevauchés le cas échéant, puis délivré en sortie du dispositif de traitement 20.

[0163] Le module de détection d'activité vocale 36 est configuré pour déterminer une présence de voix ou une absence de voix dans chaque tronçon du signal hybride.

[0164] Le module de détection d'activité vocale 36 est par exemple configuré pour déterminer la présence de voix ou l'absence de voix à partir du deuxième signal issu du transducteur à excitation mécanique osseuse ; et de préférence uniquement à partir dudit deuxième signal, sans prise en compte du premier signal.

[0165] Le deuxième microphone 14, ostéophonique ou solidien, est apte à mesurer les vibrations de la peau et du visage liée à la sollicitation des cordes vocales, et permet de capter la partie voisée d'un signal vocal tout en étant très peu sensible au bruit de fond (qui *a priori* ne fait pas suffisamment vibrer la peau de l'utilisateur pour être captée).

[0166] L'intérêt d'utiliser le deuxième microphone 14 ostéophonique réside dans son immunité au bruit de fond. Cette immunité est encore plus grande dans la partie basse fréquence du signal acquis.

[0167] Avantagusement, la détection d'activité vocale est alors effectuée après un filtrage dans le domaine fréquentiel (fonctionnant également dans le domaine temporel) du signal solidien. Le module de détection d'activité vocale 36 est alors de préférence configuré pour déterminer la présence de voix ou l'absence de voix à partir du deuxième signal filtré issu du deuxième signal filtré $\tilde{X}_k^{\text{ostBF}}$ issu de la deuxième unité de filtrage 64.

[0168] En complément facultatif, le module de détection d'activité vocale 36 est configuré pour calculer une valeur RMS pour chaque tronçon du deuxième signal, i.e. pour

chaque deuxième tronçon ; puis pour déterminer la présence de voix ou l'absence de voix en fonction de valeurs RMS respectives.

[0169] Le traitement est basé sur le calcul de l'énergie du signal tronçon par tronçon. Cependant ici, grâce au caractère immune au bruit du signal du microphone solidien filtré, l'énergie de la voix émergera tout le temps de l'énergie plancher du bruit. Le calcul du niveau RMS permet alors de connaître l'énergie du signal.

[0170] Comme connu en soi, la valeur efficace, dite aussi valeur RMS (de l'anglais *Root Mean Square*, signifiant moyenne quadratique), d'un signal périodique est la racine carrée de la moyenne du carré de cette grandeur, sur un intervalle de temps donné ou la racine carrée du moment d'ordre deux (ou variance) du signal.

[0171] Pour un tronçon temporel $x_k[n]$ de N échantillons, le calcul de la valeur RMS s'effectue alors typiquement via l'équation suivante :

[0172] [Math.14]

$$RMS_k = \sqrt{\frac{\sum_{n=0}^N x_k[n]^2}{N}}$$

[0173] où RMS_k représente la valeur RMS pour le tronçon d'indice k ;

[0174] $x_k[n]$ représente le signal pour le tronçon d'indice k ;

[0175] N représente le nombre d'échantillons dudit tronçon.

[0176] Or, dans le domaine fréquentiel, grâce à l'identité de Parseval selon laquelle l'énergie est égale dans les domaines fréquentiel et temporel, on obtient l'équation suivante :

[0177] [Math.15]

$$RMS_k = \frac{1}{\sqrt{N}} \sqrt{\sum_{m=-\frac{N}{2}}^{\frac{N}{2}} |\tilde{X}_k[m]|^2}$$

[0178] où RMS_k représente la valeur RMS pour le tronçon d'indice k ;

[0179] $|\tilde{X}_k[m]|$ représente le spectre du signal hybride pour le tronçon d'indice k ; et

[0180] N représente le nombre d'échantillons dudit tronçon.

[0181] Cette valeur du niveau RMS est optionnellement convertie en une valeur dBFS à partir de l'équation suivante :

[0182] [Math.16]

$$RMS_k^{dB} = 20 \times \log_{10} (RMS_k)$$

[0183] où $\log_{10}$ représente l'opérateur logarithme décimal, ou encore logarithme de base 10.

[0184] Cette valeur dBFS est typiquement comprise entre -94 dBFS au minimum (dans le cas d'une résolution dynamique de 16 bits) et 0 dBFS au maximum (pour un signal constant qui vaudrait 1).

[0185] En complément facultatif encore, le module de détection d'activité vocale 36 est configuré pour déterminer la présence de voix ou l'absence de voix en fonction d'une valeur moyenne de M dernières valeurs RMS calculées, également appelée RMS lissé,

et/ou d'une variation de valeur RMS entre une valeur RMS courante et une valeur RMS précédente, également appelée taux de variation du niveau RMS, avec M un nombre entier supérieur ou égal à 1.

[0186] Selon ce complément facultatif encore, le module de détection d'activité vocale 36 est par exemple configuré pour déterminer la présence de voix si ladite valeur moyenne est supérieure ou égale à un seuil prédéfini de moyenne A ou si ladite variation de valeur RMS est supérieure ou égale à un seuil prédéfini de variation B.

[0187] La valeur du niveau RMS est susceptible de varier dans le temps, et de subir des brusques variations lorsque le microphone concerné, en particulier le deuxième microphone 14, capte une vibration importante. Ce complément facultatif permet alors d'améliorer la précision et de réduire les erreurs de l'algorithme, avec un moyennage sur les M dernières valeurs calculées du niveau RMS (lors des M derniers tronçons). Ceci est par exemple mis en œuvre via un buffer circulaire qui à chaque nouveau tronçon vient ajouter la nouvelle valeur RMS calculée, supprime la M^{ième} dernière, puis moyenne l'ancienne. Le niveau RMS lissé au k^{ième} tronçon, noté $\bar{RMS}_k^{dB}$, vérifie par exemple l'équation suivante :

[0188] [Math.17]

$$\bar{RMS}_k^{dB} = \frac{1}{M} \times \sum_{j=0}^{M-1} RMS_{k-j}^{dB}$$

[0189] Le suivi de la valeur de $\bar{RMS}_k^{dB}$ au cours du temps permet de repérer les zones de voix lorsque celui-ci dépasse un certain seuil. Néanmoins, dû au lissage, ce niveau peut dépasser le seuil choisi légèrement en retard. Avantageusement, une deuxième métrique liée au niveau au RMS, à savoir le taux de variation du niveau RMS noté ΔRMS_k^{dB} , est alors calculée pour mieux détecter l'apparition de la voix, par exemple via l'équation suivante :

[0190] [Math.18]

$$\Delta RMS_k^{dB} = \frac{(\bar{RMS}_k^{dB} - \bar{RMS}_{k-1}^{dB})}{dt}$$

[0191] où ΔRMS_k^{dB} représente le taux de variation du niveau RMS pour le tronçon d'indice k ;

[0192] $\bar{RMS}_{k-1}^{dB}$, resp. $\bar{RMS}_k^{dB}$, représente le niveau RMS lissé pour le tronçon d'indice k-1, et resp. d'indice k ;

[0193] dt représente un delta de temps entre deux tronçons successifs.

[0194] La valeur dt peut correspondre exactement au delta de temps entre deux tronçons successifs, et la variation du niveau RMS sera alors exprimée en dB.s⁻¹, mais celui-ci peut prendre des valeurs très importantes.

[0195] En variante, et par commodité, la valeur dt est choisie égale à 1. Le cas échéant,

ΔRMS_k^{dB} est un taux de variation exprimé en dB.tronçon⁻¹. Cette grandeur est pertinente car au moment où un interlocuteur se met à parler, le niveau RMS augmente brutalement, se traduisant par un ΔRMS_k^{dB} positif et supérieur à 1 dB.tronçon⁻¹. Cette grandeur variant vite, elle permet de détecter la voix très rapidement, évitant ainsi de louper le début d'une phrase.

[0196] La prise de décision pour la détection d'activité vocale instantanée est alors définie par exemple par l'équation suivante :

[0197] [Math.19]

$$\begin{cases} \text{Si } \tilde{RMS}_k^{dB} \geq A \text{ alors } DAV_k = 1 \\ \text{Ou si } \Delta RMS_k^{dB} \geq B \text{ alors } DAV_k = 1 \\ \text{Sinon } DAV_k = 0 \end{cases}$$

[0198] où $\tilde{RMS}_k^{dB}$ représente le niveau RMS lissé pour le tronçon d'indice k ;

[0199] ΔRMS_k^{dB} représente taux de variation du niveau RMS pour le tronçon d'indice k ;

[0200] DAV_k est un indicateur d'activité vocale pour le tronçon d'indice k, cet indicateur étant égal à 1 si une présence de voix est déterminée, et à 0 sinon ;

[0201] A représente le seuil prédéfini de moyenne et B représente le seuil prédéfini de variation, correspondant respectivement aux seuils de niveau et du taux de variation à dépasser pour considérer que le tronçon est parlé.

[0202] Ces valeurs de seuil A et B sont prédéfinies en fonction de la dynamique de l'appareil acoustique 10, par exemple en fonction du gain du microphone concerné, en particulier du deuxième microphone 14, etc.

[0203] Le calcul de la détection d'activité vocale décrit ci-dessus donne une valeur instantanée pour chaque tronçon successif (qu'il soit chevauché ou non). Se baser uniquement sur une valeur instantanée peut conduire à des erreurs, par exemple un micro-silence dans la voix pourrait créer un passage à 0 non souhaité de l'indicateur d'activité vocale DAV. Au contraire, un bruit impulsionnel très court peut conduire à un indicateur d'activité vocale DAV égal à 1 pour un seul tronçon, avant de repasser à 0. En fonction de l'utilisation du module de détection d'activité vocale 36 (avec un mode où le canal n'est ouvert que si $DAV = 1$ par exemple), ce comportement peut provoquer des artefacts désagréables. C'est pourquoi le calcul de la détection d'activité vocale est avantageusement lissé afin d'éviter ces artefacts.

[0204] Ce lissage est par exemple réalisé à partir de l'utilisation d'un temps d'attaque et d'un temps de relâche. Lorsqu'un indicateur d'activité vocale DAV instantané DAV_{inst}^k est égal à 1 au moins aussi longtemps que le temps d'attaque (ou le nombre de tronçon(s) équivalent), alors un indicateur d'activité vocale DAV lissé DAV_{lisse}^k devient égal à 1. Au contraire, lorsque l'indicateur d'activité vocale DAV instantané DAV_{inst}^k est égal à 0 au moins aussi longtemps que le temps de relâche, alors

l'indicateur d'activité vocale DAV lissé DAV_{lisse}^k repasse à 0. Dans tous les autres cas, l'indicateur d'activité vocale DAV lissé DAV_{lisse}^k conserve la valeur qu'il avait au tronçon précédent. Pour la mise en œuvre de ce lissage, un compteur C_k est par exemple utilisé. La modification de ce compteur C_k est typiquement régie par le tableau 1 ci-après pour chaque tronçon courant d'indice k , en fonction de l'indicateur d'activité vocale DAV instantané DAV_{inst}^k et de la valeur du compteur C_{k-1} au tronçon précédent d'indice $k-1$:

[0205] [Tableaux1]

ET	$C_{k-1} \geq 0$	$C_{k-1} < 0$
$DAV_{inst}^k = 0$	Réinitialisation du compteur : $C_k = 0$	$C_k = C_{k-1} - 1$
$DAV_{inst}^k = 1$	$C_k = C_{k-1} + 1$	Réinitialisation du compteur : $C_k = 0$

[0206] La prise de décision pour la détection d'activité vocale lissée est alors définie par exemple par l'équation suivante :

[0207] [Math.20]

$$\begin{cases} \text{Si } C_k > t_{atk} \text{ alors } DAV_{lisse}^k = 1 \\ \text{Si } C_k < -t_{rel} \text{ alors } DAV_{lisse}^k = 0 \\ \text{Sinon } DAV_{lisse}^k = DAV_{lisse}^{k-1} \end{cases}$$

[0208] où DAV_{lisse}^k est l'indicateur d'activité vocale lissé pour le tronçon d'indice k , cet indicateur étant égal à 1 si une présence de voix est déterminée, et à 0 sinon ;

[0209] C_k est le compteur pour le tronçon d'indice k ;

[0210] t_{atk} représente le temps d'attaque ; et

[0211] t_{rel} représente le temps de relâche.

[0212] Le fonctionnement de l'appareil acoustique 10, et en particulier du dispositif de traitement 20, selon l'invention va être à présent décrit en regard de la [Fig.4] représentant un organigramme du procédé de traitement selon l'invention.

[0213] Le traitement appliqué au signal pour réduire le bruit est effectué de manière numérique et en temps réel. En effet, lorsque l'opérateur utilise l'appareil acoustique 10, le signal doit être débruité et envoyé à son interlocuteur le plus rapidement possible, en cherchant à diminuer au maximum la latence, avec une valeur souhaitée de 20 à 30 ms. Pour permettre un débruitage qualitatif, il faut disposer d'un minimum d'informations à analyser avant de pouvoir réduire le bruit efficacement. Le traitement effectué est alors un traitement par bloc, appliqué tronçon par tronçon au signal d'entrée. Comme indiqué précédemment, les tronçons sont typiquement chacun d'une

durée d'environ 20 ms. En effet, sur cette durée, la voix a un comportement quasi stationnaire, alors que le bruit l'est sur des durées bien plus importantes.

- [0214] Afin d'optimiser la consommation électrique, la fréquence d'échantillonnage est de préférence inférieure à 22 050 Hz, permettant une bande passante comprise dans l'intervalle [0 ; 11 025 Hz]. En conséquence pour avoir des tronçons de signal d'environ 20 ms à cette fréquence d'échantillonnage, ceux-ci devront contenir typiquement 512 échantillons.
- [0215] Le traitement appliqué au signal pour réduire le bruit est en grande partie effectué dans le domaine fréquentiel, qui est plus adapté au débruitage du fait que le but est de réduire le niveau dans les bandes de fréquences contenant le plus de bruit. Néanmoins, du fait de travailler par tronçons en fréquentiel, des problèmes de discontinuités et d'imprécisions peuvent apparaître d'un tronçon à un autre, et un chevauchement des tronçons, avec un taux de chevauchement de préférence supérieur à 50%, idéalement égal à 75%, tel que décrit ci-dessus, est alors avantageusement mis en œuvre pour les atténuer.
- [0216] Lors d'une étape initiale 100, le dispositif de traitement 20 calcule alors, via son module d'hybridation 30, le signal hybride à partir des premier et deuxième signaux analogiques, issus des premier et deuxième microphones 12, 14, de la manière décrite précédemment.
- [0217] Lors d'une étape optionnelle suivante 110, le dispositif de traitement 20 détermine, via son module de détection d'activité vocale 36, une présence de voix ou une absence de voix dans chaque tronçon du signal hybride, de la manière décrite précédemment.
- [0218] Le dispositif de traitement 20 estime ensuite, lors de l'étape suivante 120 et via son module d'estimation 32, le bruit dans le signal hybride, obtenu précédemment lors de l'étape d'hybridation 100, de la manière décrite précédemment.
- [0219] Lorsqu'optionnellement une présence de voix ou une absence de voix dans chaque tronçon du signal hybride a été déterminée lors de l'étape de détection d'activité vocale 110, le bruit est alors, lors de l'étape d'estimation 120, estimé dans le signal hybride en fonction de chaque tronçon avec une absence déterminée de voix, de la manière décrite précédemment.
- [0220] Enfin, lors de l'étape suivante 130, le dispositif de traitement 20 applique, via son module de réduction de bruit 34, l'algorithme de soustraction spectrale généralisée au signal hybride et en fonction du bruit estimé, afin de calculer le signal corrigé.
- [0221] Comme indiqué précédemment, le procédé de traitement est en temps réel ou en quasi-temps réel, avec une latence d'environ 20 à 30 ms, et un traitement par bloc, appliqué tronçon par tronçon au signal d'entrée.
- [0222] Aussi, à l'issue de l'étape 130, le procédé de traitement retourne à l'étape initiale 100, et plus généralement, chacune des étapes 100, optionnellement 110, 120 et 130 est

réitérée régulièrement afin d'être mise en œuvre pour chaque tronçon successif de signal.

[0223] Sur la [Fig.5], la courbe 200 représente alors un exemple avec un signal provenant d'un enregistrement aérien d'un locuteur s'exprimant dans un environnement fortement bruité (bruit véhicule à plus de 90 db(A)). La courbe 250 à la [Fig.5] présente le même signal après la mise en œuvre du dispositif de traitement 20 selon l'invention. On constate que le bruit est fortement atténué avec le dispositif de traitement 20 selon l'invention, tout en observant que les parties correspondant à la voix sont bien visibles et présentent alors une bonne intelligibilité.

[0224] La [Fig.6] présente un exemple de détection d'activité vocale utilisée sur un signal de voix enregistré par un microphone aérien classique pour différentes phases successives de bruit, d'une absence de bruit jusqu'à un bruit fort. La courbe 300 est la représentation temporelle de ce signal sur laquelle est superposée la décision prise par la détection d'activité vocale, où les zones grisées 310 correspondent à des zones pour lesquelles une présence de voix a été déterminée, i.e. $DAV = 1$; les autres zones, non grisées ou blanches, correspondant à des zones pour lesquelles une absence de voix a été déterminée, i.e. $DAV = 0$. Sur la [Fig.6], la courbe 320 représente le niveau RMS de ce signal issu du microphone aérien au cours du temps avec le niveau seuil à dépasser pour la prise de décision, le niveau seuil étant représenté par la droite horizontale 330 en trait pointillé. La courbe 340 correspond à l'estimation par l'algorithme du niveau RMS du bruit de fond dans les phases où la détection d'activité vocale a déterminé une absence de voix.

[0225] Dans cet exemple de la [Fig.6], le niveau seuil a été choisi volontairement bas, avec une valeur sensiblement égale à -40 dBFS pour permettre une bonne détection de la voix en l'absence de bruit. En effet, on constate que dans la phase sans bruit, pour la période temporelle comprise entre les instants temporels 0s et 15s, la voix émerge bien du bruit et le niveau RMS moyenné dépasse bien le seuil à chaque fois que l'utilisateur parle. La détection d'activité vocale classique est donc correcte sur la partie silencieuse. Cependant, dès que le bruit présente un niveau modéré, le niveau RMS moyenné est systématiquement au-dessus du seuil fixé, puisque trop bas. En conséquence, cela aboutit à une détermination erronée d'une présence de voix pendant toute la suite du signal : la détection d'activité vocale devient alors inopérante, car incapable de séparer la contribution du bruit de celle de la voix. La détection d'activité vocale donnant une réponse toujours positive, l'estimation du niveau RMS du bruit est par la même également totalement faussée, et reste sur la valeur prise lors de l'absence de bruit.

[0226] La [Fig.7] est analogue à la [Fig.6], à la différence que le seuil de détection a été remonté à une valeur sensiblement égale à -20 dBFS. La courbe 400 est la repré-

sentation temporelle de ce signal sur laquelle est superposée la décision prise par la détection d'activité vocale, où les zones grisées 410 correspondent à des zones pour lesquelles une présence de voix a été déterminée, i.e. $DAV = 1$; les autres zones, non grisées ou blanches, correspondant à des zones pour lesquelles une absence de voix a été déterminée, i.e. $DAV = 0$. Sur la [Fig.7], la courbe 420 représente le niveau RMS de ce signal issu du microphone aérien au cours du temps avec le niveau seuil à dépasser pour la prise de décision, le niveau seuil étant représenté par la droite horizontale 430 en trait pointillé. La courbe 440 correspond à l'estimation par l'algorithme du niveau RMS du bruit de fond dans les phases où la détection d'activité vocale a déterminé une absence de voix.

- [0227] Sur la [Fig.7], l'homme du métier constatera alors que la détection de voix dans la partie à bruit modéré, entre les instants temporels 15s et 30s environ, est plutôt correcte. Le niveau RMS, aux moments où il y a de la voix, permet de discriminer celle-ci du bruit. Cependant, dès que l'on augmente encore le niveau de bruit, ce seuil ne permet plus de bien distinguer la voix du bruit, et de nombreuses zones sont considérées comme exclusivement parlées, entre les instants temporels 34s et 42s par exemple, alors qu'il y a en réalité des moments d'absence de voix dans ces zones. Pire encore, en raison du seuil trop haut, dans la partie sans bruit, la détection d'activité vocale de l'état de la technique confond plusieurs fois la voix avec du bruit et manque certaines détections ou les coupe trop tôt. Cela détériore alors gravement le signal de voix. De plus, cela fausse totalement l'estimation du niveau de bruit, correspondant à la courbe 440, qui est artificiellement augmentée lorsque la personne parle.
- [0228] Finalement, au travers de ces deux exemples des figures 6 et 7 illustrant l'état de la technique, l'homme du métier comprendra qu'il faudrait que le seuil varie automatiquement (bas pour les phases de silence, plus haut pour les phases de bruit) pour permettre de bons résultats de la détection d'activité vocale de l'état de la technique avec un microphone aérien. En effet, avec la détection d'activité vocale classique, un réglage fixe du seuil ne peut correspondre correctement à la fois à un environnement bruyé et à un environnement calme, notamment en raison de la forte sensibilité des microphones aérien à l'environnement.
- [0229] La [Fig.8] illustre la mise en œuvre du dispositif de traitement 20 selon l'invention, et notamment la détection d'activité vocale selon l'invention à partir du deuxième signal issu du transducteur à excitation mécanique osseuse, ceci sur le même enregistrement que celui utilisé pour les exemples des figures 6 et 7, mais avec le deuxième microphone 14 ostéophonique, et ensuite l'utilisation de l'algorithme de soustraction spectrale généralisée.
- [0230] La courbe 500 est la représentation temporelle de ce signal sur laquelle est superposée la décision prise par la détection d'activité vocale, où les zones grisées 510

correspondent à des zones pour lesquelles une présence de voix a été déterminée, i.e. $DAV = 1$; les autres zones, non grisées ou blanches, correspondant à des zones pour lesquelles une absence de voix a été déterminée, i.e. $DAV = 0$. Sur la [Fig.8], la courbe 520 représente le niveau RMS de ce signal issu du deuxième microphone 14 ostéophonique au cours du temps avec le niveau seuil à dépasser pour la prise de décision, le niveau seuil étant représenté par la droite horizontale 530 en trait pointillé. La courbe 540 correspond à l'estimation par l'algorithme du niveau RMS du bruit de fond dans les phases où la détection d'activité vocale a déterminé une absence de voix.

- [0231] Avec le dispositif de traitement 20 selon l'invention, un premier élément marquant est que la forme d'onde associée à cet enregistrement ostéophonique filtré (filtre passe-bas) est beaucoup moins marquée par le bruit. Quel que soit le niveau de bruit, la voix émerge très facilement de celui-ci. Cet effet est encore plus visible sur la représentation du niveau RMS du signal filtré au cours du temps, il y a près de 40 dB de différence entre les pics liés à la voix et le bruit de fond. En conséquence, le choix de la valeur seuil devient plus aisé et offre une plus grande latitude qu'avec le dispositif de traitement de l'état de la technique. Ce seuil a par exemple été fixé arbitrairement ici à -35dBFS, tout en observant qu'une valeur de seuil à -25dBFS ou à -45dBFS aurait donné des résultats semblables. Grâce à cette émergence naturelle, l'algorithme de soustraction spectrale généralisée est particulièrement efficace et repère aussi bien la voix dans trois zones de bruits différents.
- [0232] Enfin, grâce à ses performances, le dispositif de traitement 20 selon l'invention est capable de détecter précisément les périodes temporelles en présence de bruit uniquement. De cette façon, le moyennage du niveau RMS du microphone aérien uniquement aux moments où $DAV = 0$, permet d'obtenir une bonne estimation du niveau du bruit de fond, représenté par la courbe 540.
- [0233] Ces résultats montrent bien l'intérêt du dispositif de traitement 20 selon l'invention de par le gain important en performance et en coût de calcul, par rapport au dispositif de traitement de l'état de la technique.
- [0234] Ainsi, lorsque l'utilisateur se trouve dans un environnement bruyant, et qu'il utilise l'appareil acoustique 10, par exemple avec une radio, pour communiquer avec un interlocuteur à distance, le signal envoyé à l'interlocuteur serait, sans mise en œuvre de l'invention, altéré par la captation non souhaitée d'une portion de bruit de fond. Le dispositif électronique de traitement 20 selon l'invention permet de réduire la présence de ce bruit de fond dans le signal envoyé à l'interlocuteur, et en particulier de filtrer la voix de ce bruit, afin de viser à n'envoyer que le signal utile à l'interlocuteur via la radio.
- [0235] Les résultats obtenus avec le dispositif électronique de traitement 20 selon l'invention, notamment ceux présentés ci-dessus en regard des figures 5 et 8, montrent

en outre la synergie entre la détection d'activité vocale basée sur la captation d'un signal via le deuxième microphone 14 ostéophonique et la réduction de bruit via l'algorithme de soustraction spectrale généralisée. Cette synergie permet d'avoir une très bonne précision quant à l'activité vocale, ce qui permet de mettre à jour le spectre du bruit de façon efficace. Les résultats obtenus avec l'algorithme de soustraction spectrale généralisée sont alors améliorés, tout en utilisant un nombre limité d'opérations de calcul.

[0236] On conçoit ainsi que le dispositif électronique de traitement 20, et le procédé de traitement associé, permettent d'améliorer encore la réduction du bruit dans le signal délivré en sortie de l'appareil acoustique 10.

Revendications

- [Revendication 1] Dispositif électronique de traitement (20) pour un appareil acoustique (10),
 l'appareil acoustique (10) comprenant un premier microphone (12) comportant un transducteur électroacoustique apte à recevoir des ondes sonores acoustiques d'un signal sonore issu de cordes vocales d'un utilisateur et à transformer lesdites ondes acoustiques en un premier signal analogique ; et un deuxième microphone (14) comportant un transducteur à excitation mécanique osseuse apte à recevoir par conduction osseuse des oscillations vibratoires dudit signal sonore et à transformer lesdites oscillations vibratoires en un deuxième signal analogique,
 le dispositif électronique de traitement (20) étant configuré pour être connecté aux premier et deuxième microphones (12,14), pour recevoir en entrée les premier et deuxième signaux analogiques et pour délivrer en sortie un signal corrigé,
 le dispositif électronique de traitement (20) comprenant :
 - un module d'hybridation (30) configuré pour calculer un signal hybride à partir des premier et deuxième signaux analogiques ;
 caractérisé en ce qu'il comprend en outre :
 - un module d'estimation (32) connecté au module d'hybridation (30) et configuré pour estimer un bruit dans le signal hybride ;
 - un module de réduction de bruit (34) connecté au module d'hybridation (30) et au module d'estimation (32), le module de réduction de bruit (34) étant configuré pour calculer le signal corrigé en appliquant un algorithme de soustraction spectrale généralisée au signal hybride et en fonction du bruit estimé.
- [Revendication 2] Dispositif (20) selon la revendication 1, dans lequel le signal hybride comporte plusieurs tronçons successifs, et le dispositif (20) comprend en outre un module de détection d'activité vocale (36) connecté au module d'hybridation (30) et configuré pour déterminer une présence de voix ou une absence de voix dans chaque tronçon du signal hybride ; le module d'estimation (32) étant alors configuré pour estimer le bruit dans le signal hybride en fonction de chaque tronçon avec une absence déterminée de voix.
- [Revendication 3] Dispositif (20) selon la revendication 2, dans lequel le module de détection d'activité vocale (36) est configuré pour déterminer la

présence de voix ou l'absence de voix à partir du deuxième signal issu du transducteur à excitation mécanique osseuse ;

le module de détection d'activité vocale (36) étant de préférence configuré pour déterminer la présence de voix ou l'absence de voix uniquement à partir du deuxième signal, sans prise en compte du premier signal.

[Revendication 4] Dispositif (20) selon la revendication 3, dans lequel le deuxième signal comporte plusieurs tronçons successifs, et le module de détection d'activité vocale (36) est configuré pour calculer une valeur RMS pour chaque tronçon du deuxième signal, puis pour déterminer la présence de voix ou l'absence de voix en fonction de valeur(s) RMS respective(s).

[Revendication 5] Dispositif (20) selon la revendication 4, dans lequel le module de détection d'activité vocale (36) est configuré pour déterminer la présence de voix ou l'absence de voix en fonction d'une valeur moyenne de M dernière(s) valeur(s) RMS calculée(s) et/ou d'une variation de valeur RMS entre une valeur RMS courante et une valeur RMS précédente, M étant un nombre entier supérieur ou égal à 1 ; le module de détection d'activité vocale (36) étant de préférence configuré pour déterminer la présence de voix si ladite valeur moyenne est supérieure ou égale à un seuil prédéfini de moyenne (A) ou si ladite variation de valeur RMS est supérieure ou égale à un seuil prédéfini de variation (B).

[Revendication 6] Dispositif (20) selon l'une quelconque des revendications précédentes, dans lequel le module d'hybridation (30) est configuré pour convertir le premier signal analogique en un premier signal numérique, au fur et à mesure de la réception du premier signal analogique, et pour générer des premiers tronçons successifs à partir du premier signal numérique, chaque nouveau premier tronçon généré comportant des échantillons d'un premier tronçon précédent et de nouveaux échantillons du premier signal numérique ; et le module d'hybridation (30) est configuré pour convertir le deuxième signal analogique en un deuxième signal numérique, au fur et à mesure de la réception du deuxième signal analogique, et pour générer des deuxièmes tronçons successifs à partir du deuxième signal numérique, chaque nouveau deuxième tronçon généré comportant des échantillons d'un deuxième tronçon précédent et de nouveaux échantillons du deuxième signal numérique ; des tronçons hybrides du signal hybride étant alors calculés au fur et à

mesure à partir des premiers et deuxièmes tronçons générés ; le signal corrigé étant ensuite calculé à partir desdits tronçons hybrides.

[Revendication 7]

Dispositif (20) selon l'une quelconque des revendications précédentes, dans lequel le module d'hybridation (30) est configuré pour obtenir un premier signal filtré en appliquant au premier signal un premier filtre associé à une première plage de fréquences ; pour obtenir un deuxième signal filtré en appliquant au deuxième signal un deuxième filtre associé à une deuxième plage de fréquences ; puis pour calculer le signal hybride en sommant le premier signal filtré et le deuxième signal filtré, la deuxième plage de fréquences étant distincte de la première plage de fréquences ;

la première plage de fréquences comportant de préférence des fréquences supérieures à celles de la deuxième plage de fréquences ; les première et deuxième plages de fréquences étant de préférence encore disjointes.

[Revendication 8]

Appareil acoustique (10) comprenant :

- un premier microphone (12) comportant un transducteur électroacoustique apte à recevoir des ondes sonores acoustiques d'un signal sonore issu de cordes vocales d'un utilisateur et à transformer lesdites ondes acoustiques en un premier signal analogique ;
- un deuxième microphone (14) comportant un transducteur à excitation mécanique osseuse apte à recevoir par conduction osseuse des oscillations vibratoires dudit signal sonore et à transformer lesdites oscillations vibratoires en un deuxième signal analogique ;
- un dispositif électronique de traitement (20) connecté aux premier et deuxième microphones (12,14), le dispositif électronique de traitement (20) étant configuré pour recevoir en entrée les premier et deuxième signaux analogiques, puis pour délivrer en sortie un signal corrigé ; caractérisé en ce que le dispositif électronique de traitement (20) est selon l'une quelconque des revendications précédentes.

[Revendication 9]

Procédé de traitement, le procédé étant mis en œuvre par un dispositif électronique de traitement (20) connecté à des premier et deuxième microphones (12,14), le premier microphone (12) comportant un transducteur électroacoustique apte à recevoir des ondes sonores acoustiques d'un signal sonore issu de cordes vocales d'un utilisateur et à transformer lesdites ondes acoustiques en un premier signal analogique ; et le deuxième microphone (14) comportant un transducteur à excitation mécanique osseuse apte à recevoir par

conduction osseuse des oscillations vibratoires dudit signal sonore et à transformer lesdites oscillations vibratoires en un deuxième signal analogique, le dispositif électronique de traitement (20) étant configuré pour recevoir en entrée les premier et deuxième signaux analogiques et pour délivrer en sortie un signal corrigé,

le procédé de traitement comprenant :

- une étape d'hybridation (100) comportant le calcul d'un signal hybride à partir des premier et deuxième signaux analogiques ;

caractérisé en ce qu'il comprend en outre :

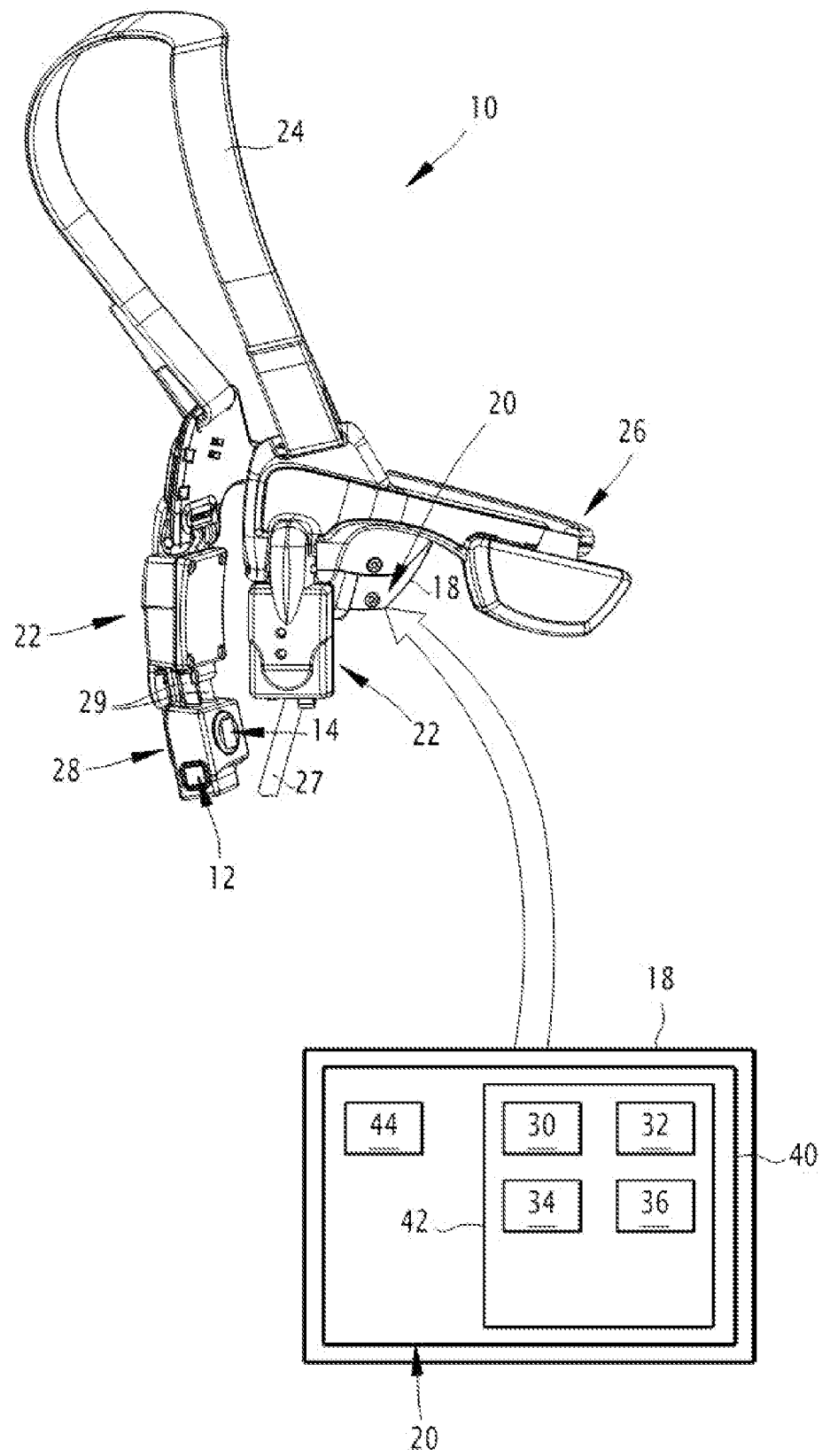
- une étape d'estimation (120) d'un bruit dans le signal hybride ; et

- une étape de réduction de bruit (130) comportant le calcul du signal corrigé en appliquant un algorithme de soustraction spectrale généralisée au signal hybride et en fonction du bruit estimé.

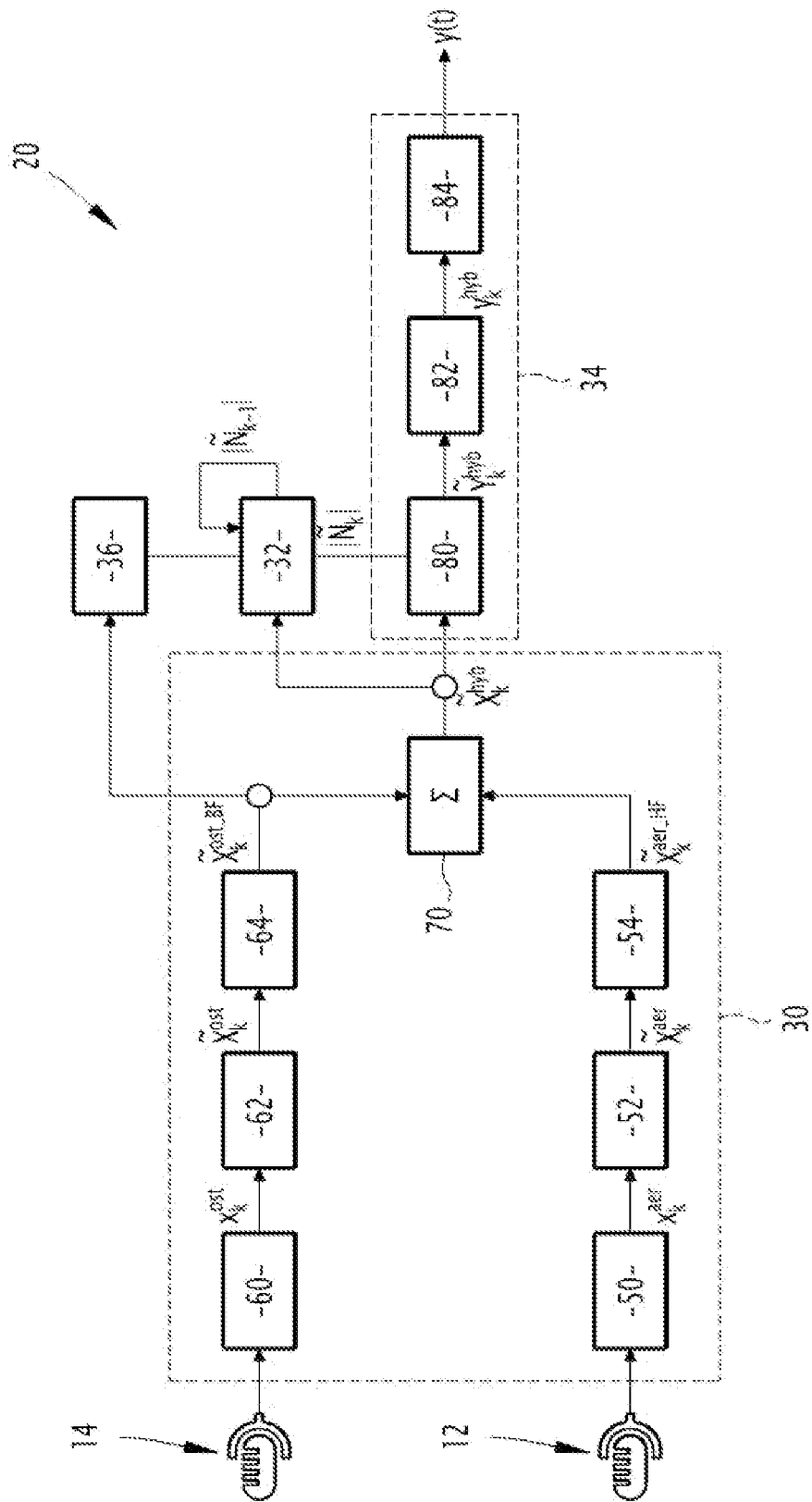
[Revendication 10]

Programme d'ordinateur comportant des instructions logicielles qui, lorsqu'elles sont exécutées par un ordinateur, mettent en œuvre un procédé selon la revendication précédente.

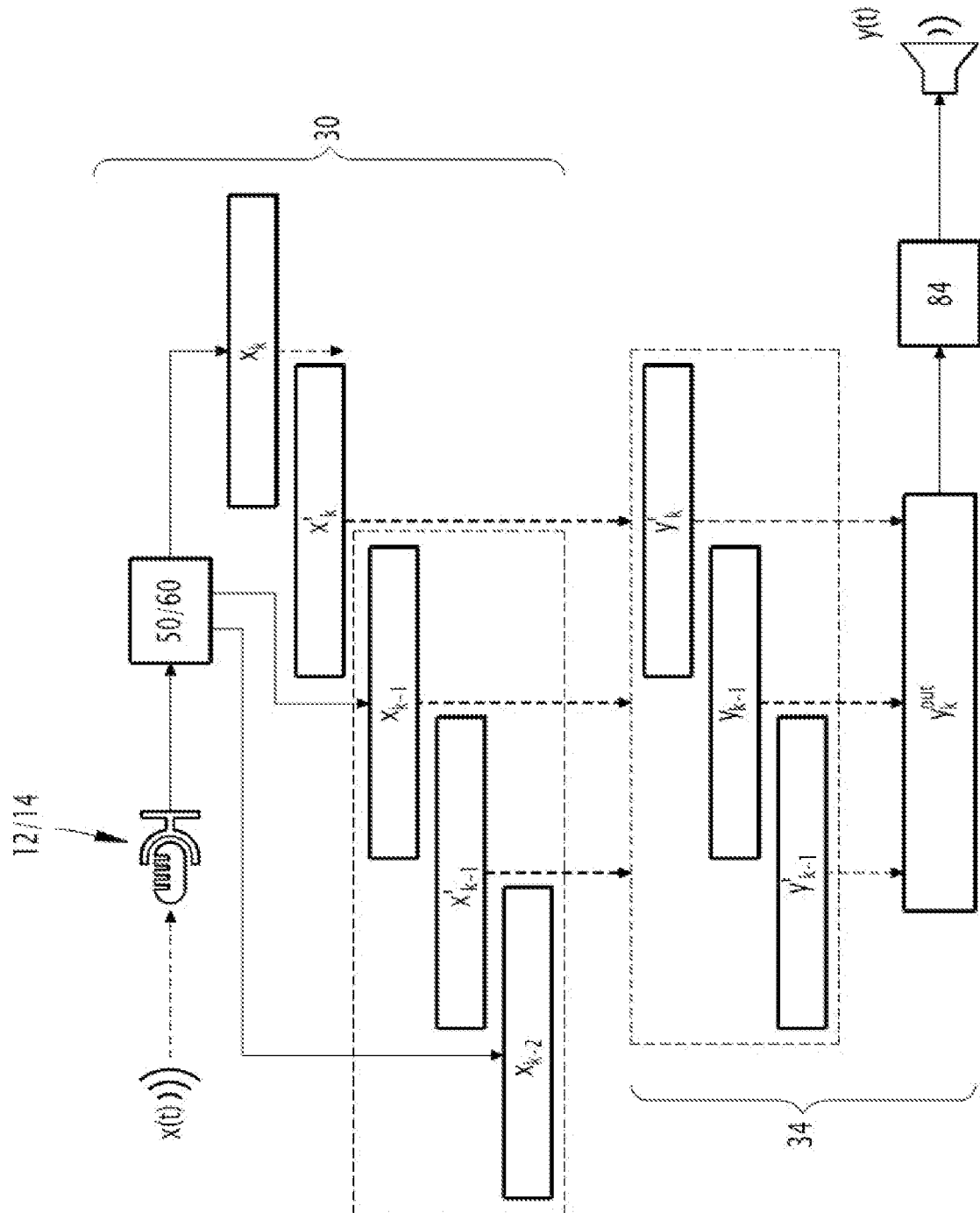
[Fig. 1]



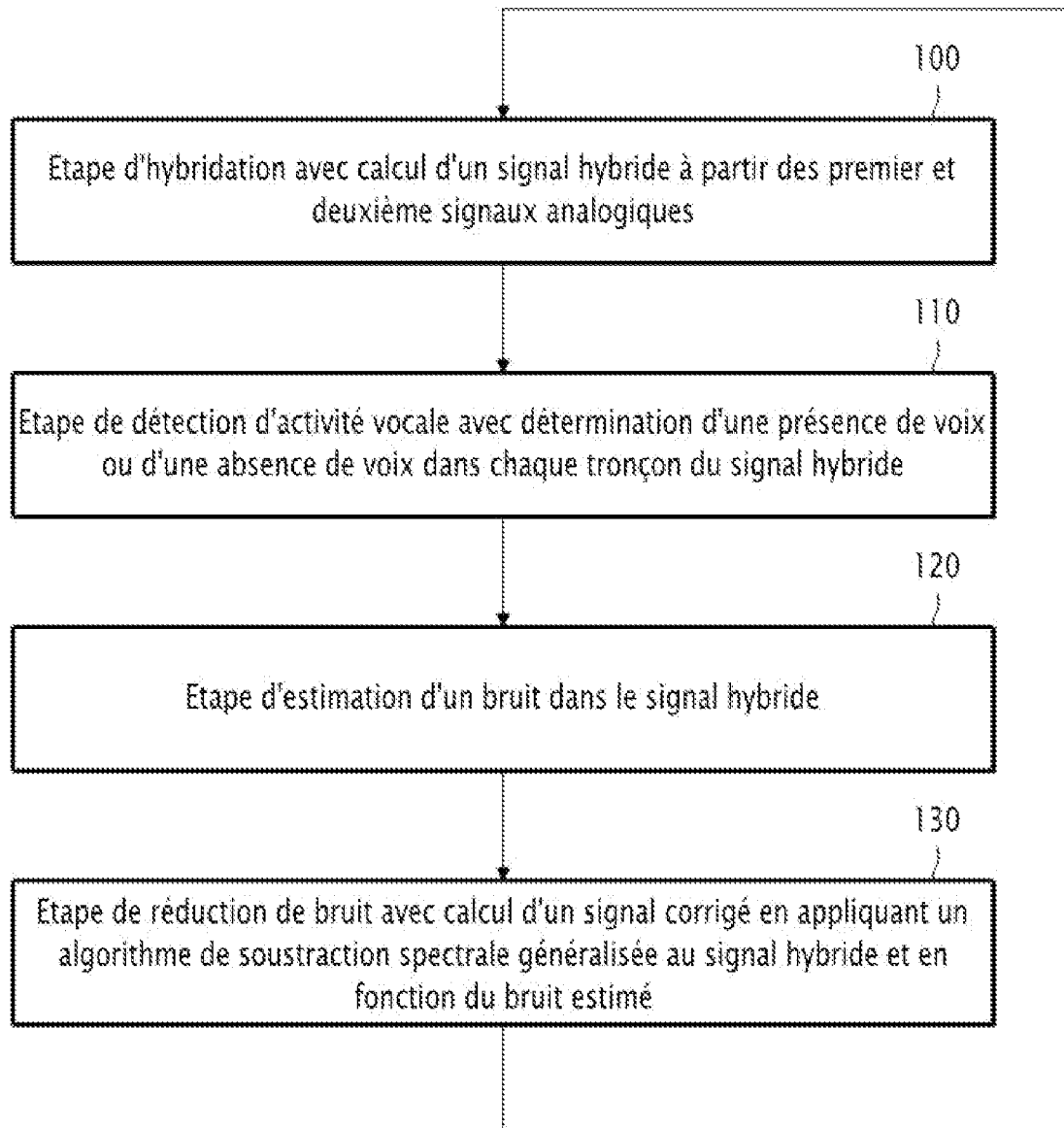
[Fig. 2]



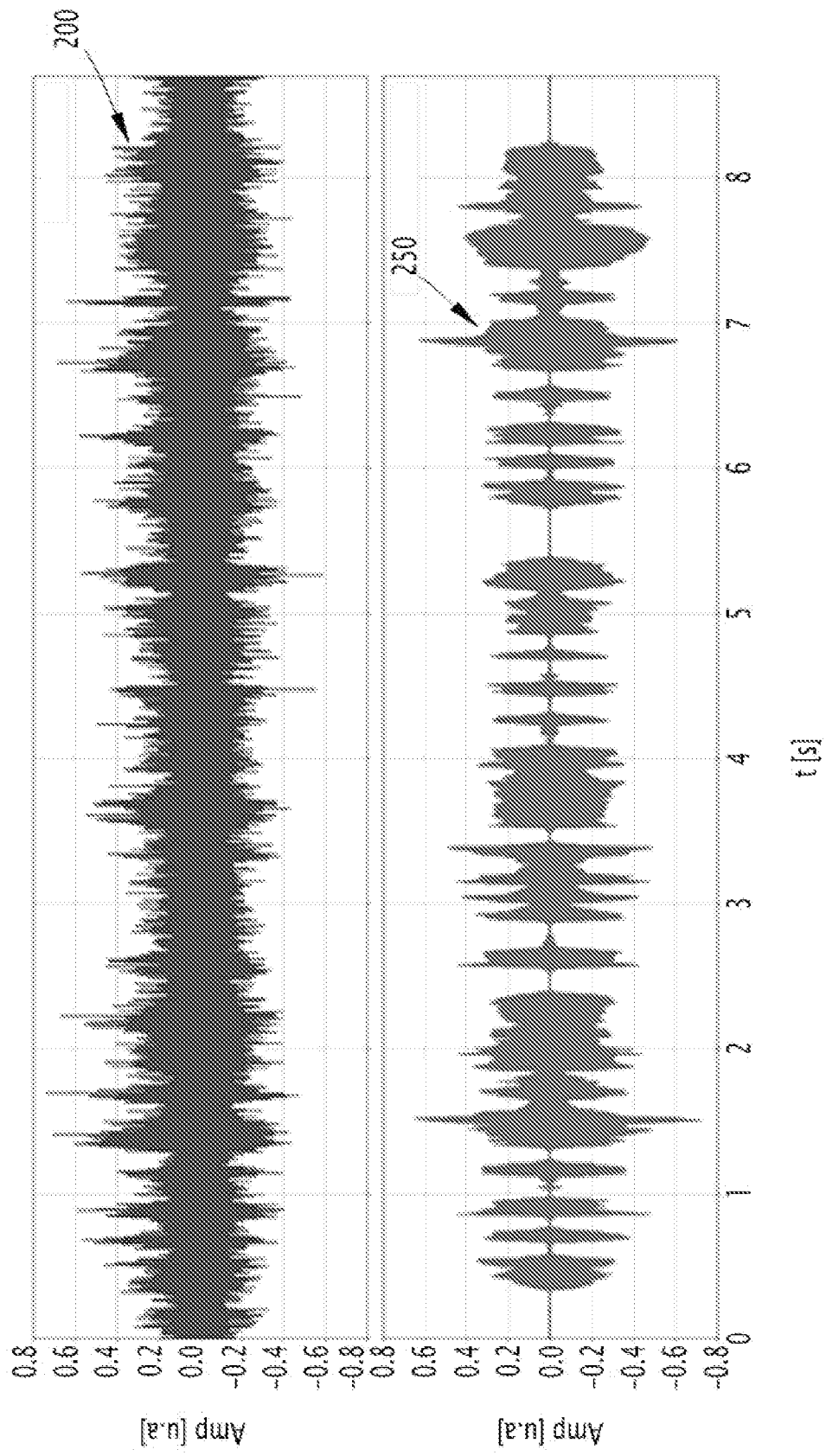
[Fig. 3]



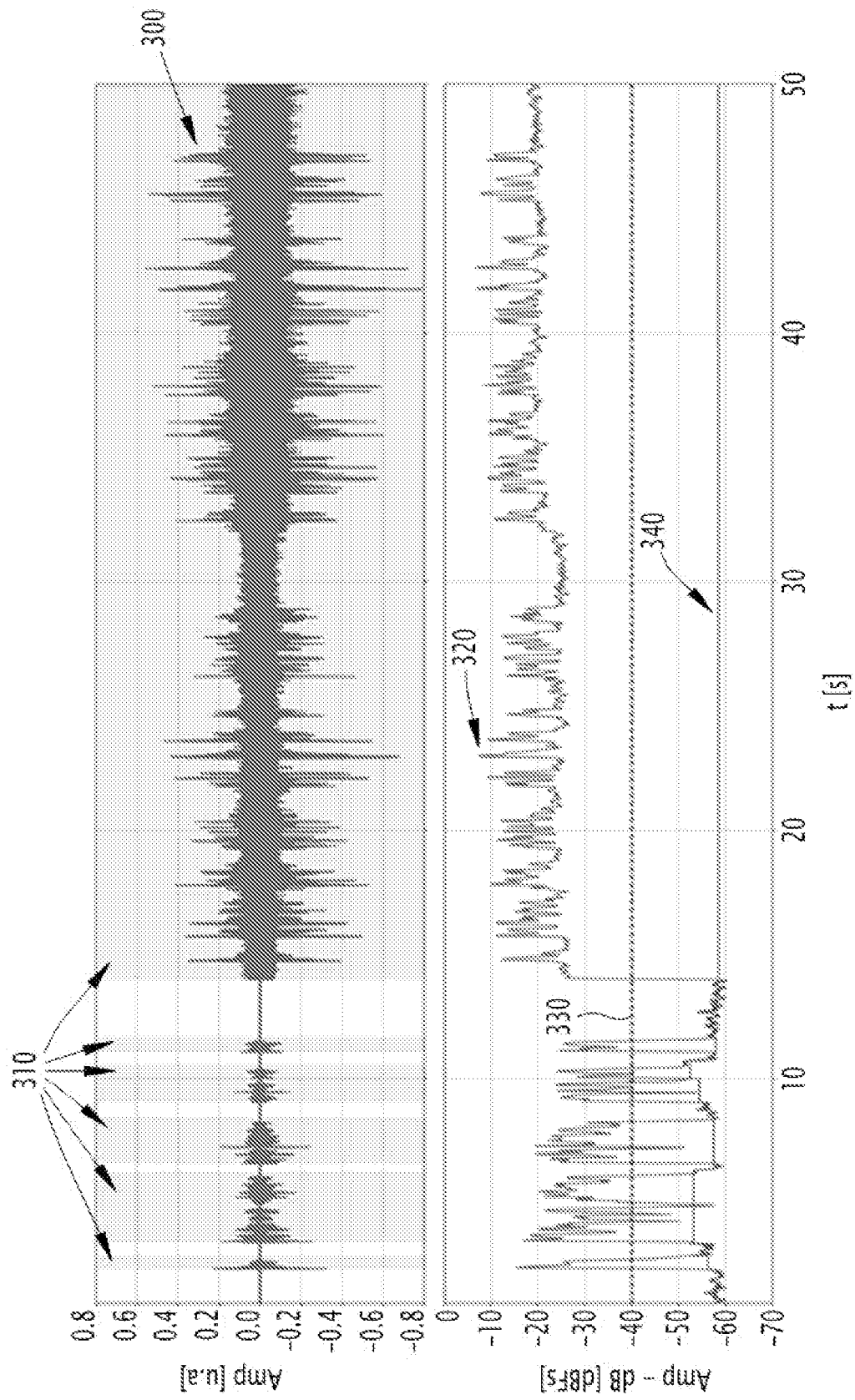
[Fig. 4]



[Fig. 5]

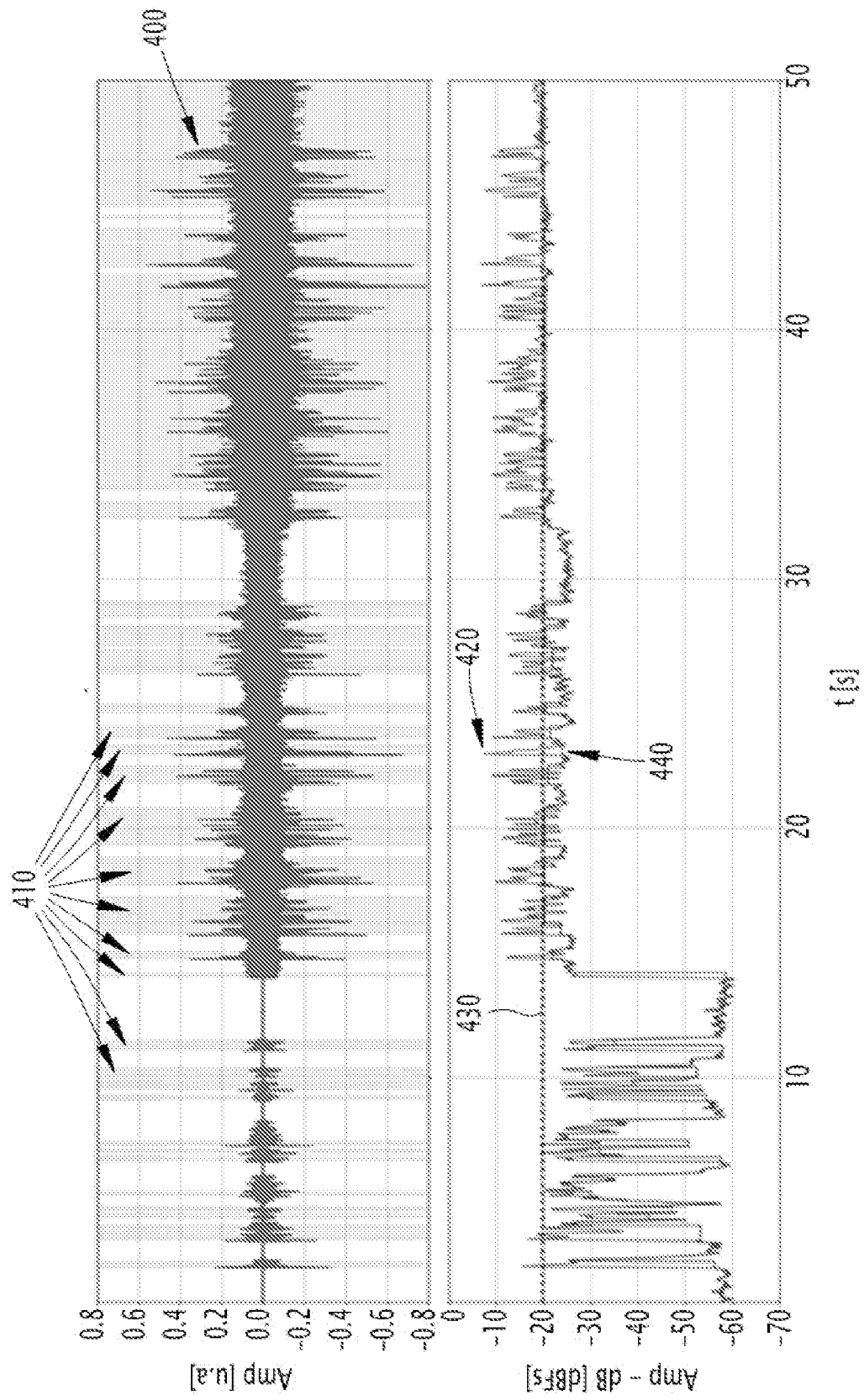


[Fig. 6]



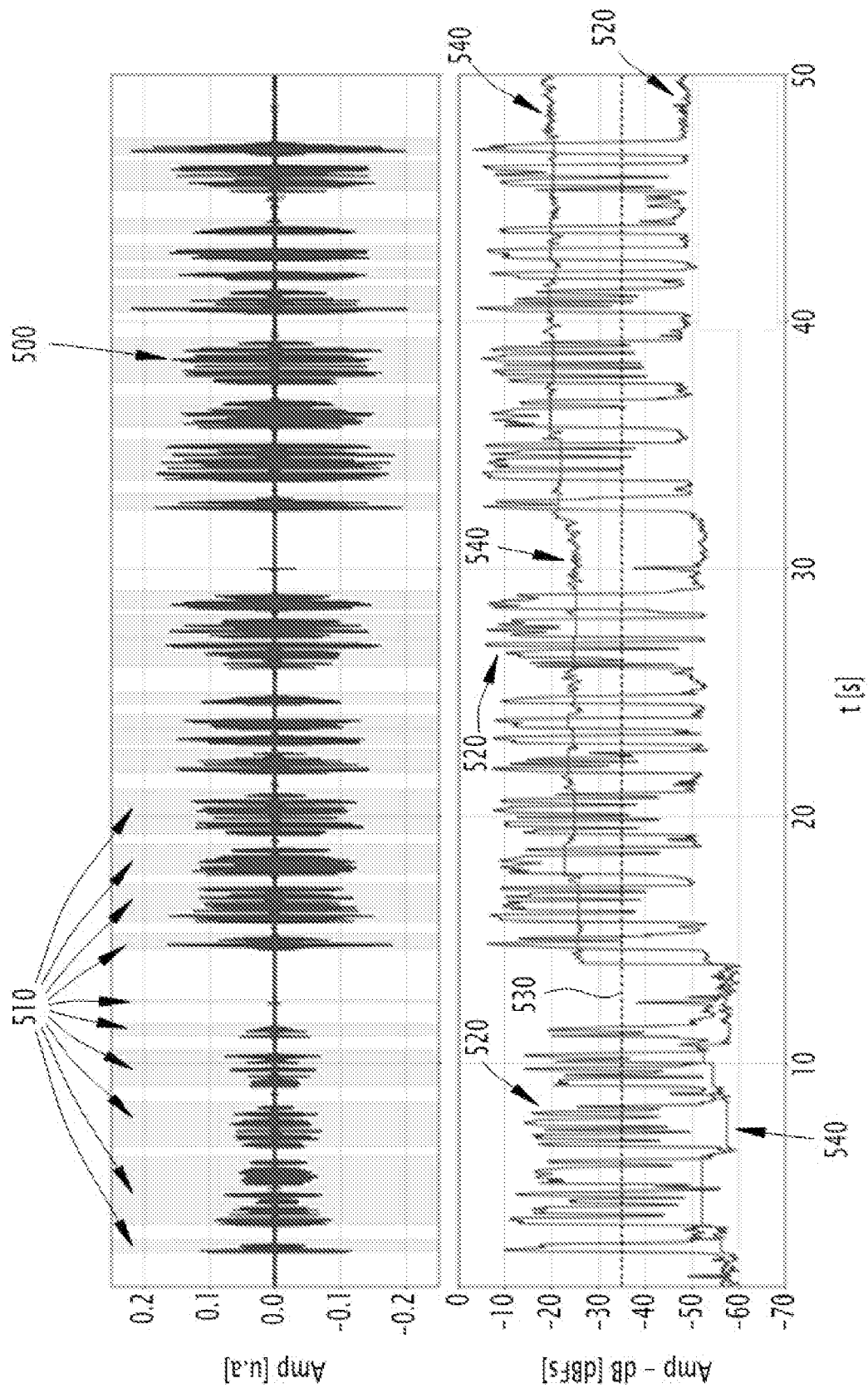
(ART ANTERIEUR)

[Fig. 7]



(ART ANTERIEUR)

[Fig. 8]



RAPPORT DE RECHERCHE PRÉLIMINAIRE

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

N° d'enregistrement
national

FA 907699
FR 2205151

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X	US 2022/150627 A1 (ZHOU MEILIN [CN] ET AL) 12 mai 2022 (2022-05-12) * alinéas [0061] - [0066], [0078] - [0085], [0093] - [0119]; figures 1, 4A, 5 * -----	1, 6, 8-10	G10K11/16 H04M1/19 H04R1/10 H04R3/00
X	US 2014/270231 A1 (DUSAN SORIN V [US] ET AL) 18 septembre 2014 (2014-09-18) * alinéa [0022] - alinéa [0049]; figures 2-5 * -----	1-10	
			DOMAINES TECHNIQUES RECHERCHÉS (IPC)
			G10L H04R
Date d'achèvement de la recherche		Examineur	
12 janvier 2023		Joder, Cyril	
CATÉGORIE DES DOCUMENTS CITÉS X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant			

ANNEXE AU RAPPORT DE RECHERCHE PRÉLIMINAIRE **RELATIF A LA DEMANDE DE BREVET FRANÇAIS NO. FR 2205151 FA 907699**

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche préliminaire visé ci-dessus.
 Les dits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du **12-01-2023**
 Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets, ni de l'Administration française

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 2022150627 A1	12-05-2022	BR 112022004158 A2	31-05-2022
		CN 114424581 A	29-04-2022
		EP 4005226 A1	01-06-2022
		JP 2022547525 A	14-11-2022
		KR 20220062598 A	17-05-2022
		US 2022150627 A1	12-05-2022
		WO 2021046796 A1	18-03-2021

US 2014270231 A1	18-09-2014	AUCUN	
