



(11)

EP 1 710 788 A1

(12)

DEMANDE DE BREVET EUROPEEN

(43) Date de publication:
11.10.2006 Bulletin 2006/41

(51) Int Cl.:
G10L 21/02 (2006.01)

(21) Numéro de dépôt: 05102714.2

(22) Date de dépôt: 07.04.2005

(84) Etats contractants désignés:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR
HU IE IS IT LI LT LU MC NL PL PT RO SE SI SK TR
Etats d'extension désignés:
AL BA HR LV MK YU

(71) Demandeur: CSEM Centre Suisse d'Electronique
et de
Microtechnique SA Recherche et Développement
2007 Neuchâtel (CH)

(72) Inventeurs:
• Vetter, Rolf
1025 Monnas (CH)

• Cornuz, Jérôme
1110 Morges (CH)
• Vuadens, Philippe
1400 Yverdon (CH)
• Sola I Caros, Josep
2035 Corcelles (CH)
• Renevey, Philippe
1005 Lausanne (CH)

(74) Mandataire: GLN et al
Rue du Puits-Godet 8a
2000 Neuchâtel (CH)

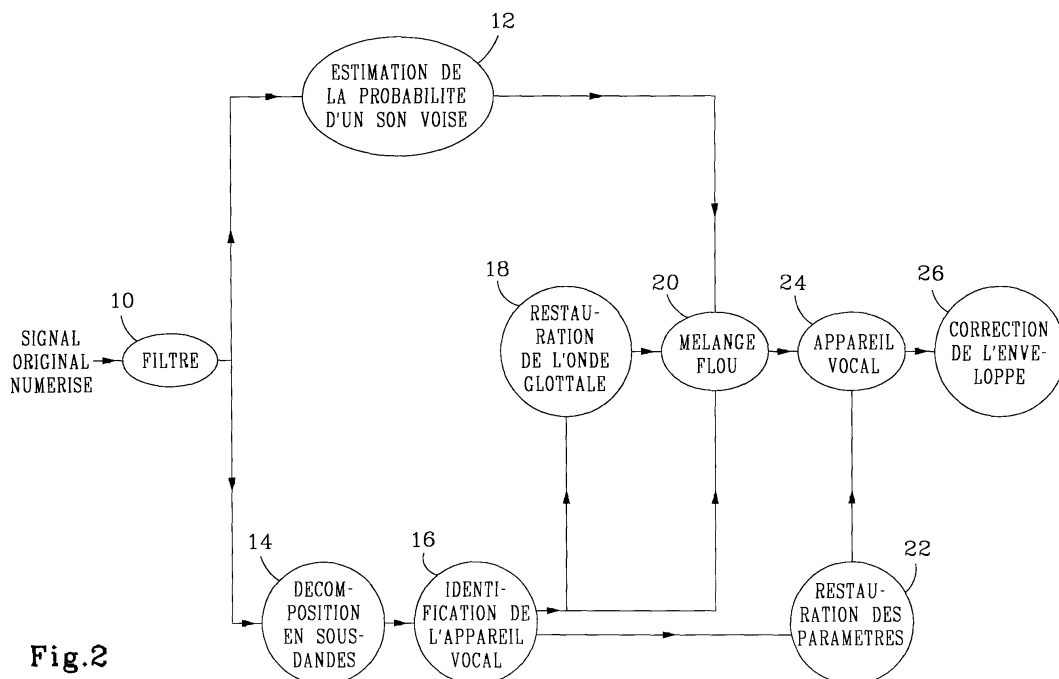
(54) Procédé et dispositif de conversion de la voix

(57) L'invention concerne un procédé de restauration d'une élocution altérée par traitement d'un signal électrique, provenant d'un microphone, représentatif de cette élocution, et comportant les principales étapes suivantes:

- extraction et séparation de l'excitation originale et des paramètres de l'articulation,

- restauration de l'excitation originale à partir d'éléments issus de l'élocution altérée et d'éléments issus d'une élocution saine, et

- reconstruction d'un signal de parole restauré à partir des paramètres de l'articulation et de l'excitation originale restaurée afin de produire un signal acoustique représentatif de l'élocution reconstruite.



Description

[0001] La présente invention concerne un procédé et un dispositif d'aide vocale pour des personnes ayant subi une ablation du larynx.

[0002] La laryngectomie, opération souvent pratiquée en cas de cancer du larynx, prive le patient de sa faculté de parler. En effet, l'enlèvement partiel ou total des cordes vocales empêche la génération d'une excitation harmonique nécessaire pour la production de voyelles ou de consonnes voisées.

[0003] Traditionnellement, une rééducation longue et pénible est nécessaire pour permettre au patient de remplacer la source d'excitation manquante, c'est-à-dire le larynx. Dans ce cas, il doit apprendre à emmagasiner de l'air dans l'oesophage, puis à le libérer. L'oesophage et le pharynx produisent ainsi une vibration basse fréquence remplaçant celle du larynx. L'élocution oesophagienne obtenue se caractérise toutefois par une faible intensité, une médiocre stabilité en fréquence et une fréquence fondamentale anormalement basse comparée à celle de l'élocution laryngienne. En outre, les difficultés d'apprentissage sont importantes, et les efforts pour parler, considérables.

[0004] Une solution à ces problèmes est le recours à un larynx artificiel. Une source d'excitation synthétique est alors créée par un dispositif inséré dans l'oesophage du patient ou porté sur la gorge. L'utilisation de ce type de larynx artificiel réduit, certes, l'effort à fournir et la durée d'apprentissage, mais elle résulte en une voix artificielle où toutes les caractéristiques de la voix humaine ont été éliminées, rendant impossible l'expression des divers sentiments et émotions. Une telle excitation électronique n'est donc pas idéale.

[0005] La présente invention pallie cet inconvénient, en proposant un procédé et un dispositif de traitement de l'élocution altérée, permettant, d'une part, le rétablissement d'une voix perçue comme naturelle et, d'autre part, la diminution des efforts pour la générer. L'invention permet, par ailleurs, d'éviter toute intervention chirurgicale supplémentaire, et propose un dispositif compact, portable, permettant son utilisation en toutes circonstances.

[0006] De façon plus précise, l'invention concerne un procédé de restauration d'une élocution altérée par traitement d'un signal électrique, provenant d'un microphone, représentatif de ladite élocution, caractérisé en ce qu'il comporte les principales étapes suivantes:

- extraction et séparation de l'excitation originale et des paramètres de l'articulation,
- restauration de l'excitation originale à partir d'éléments issus de ladite élocution altérée et d'éléments issus d'une élocution saine, et
- reconstruction de l'élocution à partir desdits paramètres de l'articulation et de ladite excitation originale restaurée afin de produire un signal acoustique représentatif de l'élocution reconstruite.

[0007] De façon avantageuse, le procédé selon l'invention comporte encore :

- une étape d'estimation de la probabilité d'un son voisé, effectuée en parallèle avec les étapes d'extraction de l'excitation et des paramètres de l'articulation, et de restauration de l'excitation originale, et
- une étape de mélange de l'excitation originale restaurée et de l'excitation originale à partir de ladite probabilité d'un son voisé.

[0008] L'invention concerne également un dispositif de restauration d'une élocution altérée, caractérisé en ce qu'il comprend :

- un système d'acquisition de cette élocution altérée produisant un signal électrique qui la représente,
- un processeur de ce signal, programmé de façon à mettre en oeuvre le procédé de restauration d'une élocution altérée, et
- un système de restitution du son transformant le signal délivré par le processeur en signal acoustique.

[0009] D'autres caractéristiques de l'invention apparaîtront plus clairement à la lecture de la description qui suit, faite en référence aux dessins annexés, dans lesquels :

- la figure 1 illustre un modèle de production de la parole,
- la figure 2 illustre schématiquement les étapes du procédé de restauration de l'élocution selon l'invention,
- la figure 3 détaille une étape clé de ce procédé, et
- la figure 4 représente schématiquement un mode de réalisation du système d'aide vocale selon l'invention.

[0010] La présente invention repose sur la supposition que le sujet est, de préférence, déjà capable, après avoir effectué l'entraînement nécessaire, de produire, par modulation du flux d'air émis, une élocution altérée, par exemple de type oesophagienne, exprimant le message qu'il souhaite transmettre. Les sons qu'il émet sont captés par un microphone extérieur dont les signaux de sortie sont traités par un circuit de traitement chargé de restaurer la voix originale, émise alors, par exemple, par un haut-parleur. Il convient de noter qu'il existe d'autres types d'élocution altérée que l'élocution oesophagienne. On peut mentionner, à titre d'exemple, l'élocution trachéo-oesophagienne, l'élocution supraglottale, etc. Dans la suite de la description, la désignation d'élocution oesophagienne sera supposée désigner tout autre type d'altération de l'élocution résultant de l'ablation total ou partielle du larynx.

[0011] Le principe, selon l'invention, de restauration de l'élocution oesophagienne repose sur un modèle connu de production de la parole, schématisé sur la figure 1 et décrit en détail dans l'article de G. Fant (Q. Prog. Status Rep. Speech Transmiss. Lab.1, 21-37). Selon ce modèle, la génération d'un son est décomposée en deux blocs distincts. Ceux-ci sont constitués, d'une part, de l'excitation 1 générée par la source, les poumons et le larynx chez le sujet sain, ou l'oesophage et le pharynx chez le sujet laryngectomisé, et, d'autre part, de l'articulation 2 réalisée par l'appareil vocal constitué de la glotte, l'épiglotte, le pharynx, la bouche et le nez.

[0012] Chez un sujet sain, l'excitation 1 est de deux types, en fonction de la configuration du larynx.

[0013] En configuration ouverte, l'air génère des turbulences dans les cavités vocales, sans faire vibrer les cordes vocales. L'excitation produite s'apparente à un signal bruité 3 et est utilisée pour générer des consonnes non voisées (p, t, s, etc...).

[0014] En configuration tendue, les vibrations du larynx excitent les cordes vocales qui produisent une onde acoustique harmonique, appelée encore onde glottale 4, dont le motif et la fréquence fondamentale moyenne, située entre 100 et 250Hz, sont caractéristiques de chaque individu. Cette excitation permet de produire les voyelles et consonnes voisées (b, d, z, etc...).

[0015] Chez un sujet sain, l'excitation 1 se présente donc sous la forme d'une suite de signaux harmoniques 4 alternant ou se mêlant avec des signaux bruités 3. En outre, l'expression des émotions, du stress et de divers sentiments, se traduit par de subtiles modulations de la fréquence fondamentale et de l'amplitude du signal harmonique 4 émis.

[0016] Chez un sujet ayant subi une laryngectomie, le premier type d'excitation, produit par des turbulences d'air dans les cavités vocales, n'est pas affecté. L'élocution de consonnes non voisées, à l'aide d'un signal bruité 3, est, par conséquent, satisfaisante.

[0017] Il en va tout autrement de l'élocution des voyelles et consonnes voisées. En effet, l'excitation par le larynx en configuration tendue est remplacée par un enchaînement d'éruptions générées par l'oesophage. L'onde acoustique produite présente une fréquence fondamentale particulièrement instable, déplacée vers les valeurs basses, puisqu'elle se situe autour de 70Hz, et dont le motif est déformé. La production de voyelles et consonnes voisées en est fortement perturbée.

[0018] Il en résulte que l'excitation, chez un sujet sans larynx, se présente sous la forme d'une suite de signaux quasi-harmoniques alternant ou se mêlant avec des signaux bruités.

[0019] Par ailleurs, l'appareil vocal d'une personne ayant subi une ablation du larynx étant fonctionnel, l'articulation 2, qui constitue le deuxième bloc de la génération de la parole, est faiblement altérée. On notera également que l'élocution altérée, par exemple de type oesophagienne, bien que déficiente au niveau de la production d'une onde acoustique harmonique, est capable

de réaliser les subtiles modulations de fréquence et d'amplitude caractéristiques de la voix humaine.

[0020] Il s'ensuit que la restauration de l'élocution altérée de type oesophagienne requiert principalement la restauration de l'excitation durant les alternances voisées de la parole, c'est à dire la restauration de l'onde glottale 4, les alternances non voisées et l'articulation n'étant relativement pas affectées par l'ablation du larynx.

[0021] On se référera maintenant à la figure 2 qui illustre le procédé de reconstruction de la voix selon l'invention, basé sur le modèle exposé ci-dessus. Il comprend diverses étapes de traitement du signal délivré par le microphone, qui sont menées en parallèle ou en série, et permettent de passer de la voix oesophagienne, à une voix de type laryngienne exprimée par le haut-parleur.

[0022] Le signal entrant est un signal électrique numérisé représentant la voix oesophagienne. Dans un premier bloc 10, ce signal original subit un traitement de rehaussement du signal (voir en particulier le brevet européen EP 1'253'581 de la demanderesse intitulé "Method and System for Enhancing Speech in a Noisy Environment"), puis le signal obtenu est dirigé en parallèle vers un bloc d'estimation de la probabilité d'un son voisé 12 et une série de trois blocs 14, 16 et 18 visant à reconstituer les deux types d'excitations selon le modèle d'élocution décrit en regard de la figure 1.

[0023] Comme exposé plus haut, un son d'origine laryngienne est de type voisé ou non voisé ou un mélange des deux. Dans le cas de l'élocution oesophagienne, il est parfois difficile de distinguer clairement entre les deux types de sons. C'est pourquoi, le bloc 12 réalise une estimation de la probabilité d'un son voisé, à l'aide de modules de classification automatique connus de l'homme de métier et tels que décrits dans l'article de J.M. Solà et al ("Environmental Robust Features for Speech Detection", Proc. INTERSPEECH-ICSLP'04, Jeju Island, Korea, October 2004). Cette estimation, sous la forme d'un nombre compris entre zéro et un, est ensuite dirigée vers un bloc 20 destiné à effectuer un mélange entre excitation harmonique et excitation bruitée, comme explicité ultérieurement.

[0024] Parallèlement à l'étape décrite précédemment, les blocs en série 14 et 16, d'une part, et 14, 16 et 18, d'autre part, visent à produire une excitation respectivement originale et harmonique.

[0025] Les blocs 14 et 16, communs aux deux chemins, réalisent respectivement la décomposition en sous-bandes et l'identification des paramètres de l'articulation, opérations connues de l'homme de métier.

[0026] La décomposition en sous-bandes consiste en un découpage en bandes du spectre en fréquences obtenu par transformée de Fourier du signal, et un rééquilibrage de ce spectre par l'augmentation en amplitude des bandes les moins bruitées par rapport aux autres. Cette opération, bien qu'efficace pour filtrer le signal, est facultative.

[0027] L'identification des paramètres de l'articulation,

par contre, est une étape cruciale visant à séparer, au sein du signal entrant, l'excitation des paramètres de l'articulation. Une description de la méthode utilisée est donnée dans un article de Yingyong Qi (J. Acoust. Soc. Am. 88 (3), September 1990). En sortie du bloc 16, on trouve donc, d'une part, l'excitation originale sous forme d'un signal amplitude en fonction du temps et, d'autre part, les paramètres de l'articulation sous forme d'un vecteur.

[0028] L'excitation originale est ensuite dirigée, d'une part, vers le bloc de restauration de l'onde glottale 18 et, d'autre part, vers le bloc 20 de mélange des deux excitations.

[0029] Le bloc 18, qui constitue le coeur de l'invention, est dissocié en différents blocs représentés en figure 3 et fait l'objet d'une description complète ultérieure. Son rôle est de restaurer l'excitation harmonique fortement altérée chez le sujet ayant subi une laryngectomie. Une fois cette restauration réalisée, le signal est dirigé vers le bloc de mélange flou 20.

[0030] Dans le bloc 20, les excitations harmonique restaurée et originale inchangée sont mélangées dans des proportions fixées par l'estimation de probabilité d'un son voisé. Si l'excitation est estimée purement bruitée, seule l'excitation originale sera conservée en sortie du bloc 20. En revanche, si l'excitation est estimée purement harmonique, seule l'excitation harmonique restaurée sera conservée en sortie du bloc 20. Dans les cas intermédiaires, le bloc 20 effectuera un mélange entre excitation harmonique restaurée et excitation originale inchangée, le dernier signal n'étant pas un signal purement bruité mais une superposition des signaux bruités et quasi-harmoniques. En sortie du bloc 20, l'excitation restaurée est de type laryngienne.

[0031] En parallèle à ces opérations de restauration de l'excitation, le vecteur de paramètres d'articulation est, à la sortie du bloc d'identification de l'appareil vocal 16, vers un bloc 22 destiné à la restauration de ces paramètres. En effet, bien que satisfaisante chez un sujet laryngectomisé, l'articulation doit être quelque peu corrigée pour correspondre à celle d'un sujet sain. Une description de cette opération est donnée dans l'article "Replacing tracheoesophageal voicing sources using LPC synthesis", Yingyong Qi, J. Acoust. Soc. Am. 88 (3), September 1990.

[0032] Les paramètres de l'articulation ainsi restaurés et l'excitation reconstituée sont dirigés vers un bloc 24 qui procède à une convolution de l'excitation par les paramètres de l'articulation afin de reconstituer la parole.

[0033] Enfin, une dernière correction est apportée par un bloc 26 qui effectue un redressement du signal afin de corriger une décroissance de la puissance observée sur les voyelles longues.

[0034] Ainsi est proposé un procédé de reconstitution de la parole pour un sujet laryngectomisé. On notera que cet algorithme peut être également utile à d'autres sujets, présentant des pathologies différentes, mais ayant des difficultés à émettre un son voisé.

[0035] On se référera maintenant à la figure 3 qui il-

lustre en détail les opérations de restauration de l'excitation harmonique, réalisées par le bloc 18 et constituant le coeur de la présente invention.

[0036] Le bloc 18 est constitué de 4 sous-blocs 18a, 18b, 18c et 18d assurant les différentes fonctions nécessaires à la restauration de l'excitation harmonique. Comme déjà exposé, l'excitation harmonique se présente sous la forme d'un signal sensiblement périodique possédant un motif de base, dont la forme dépend uniquement de la physiologie de la source d'excitation, et une fréquence fondamentale moyenne, caractéristiques d'un individu. Chez un sujet laryngectomisé, la restauration du signal glottal consiste principalement en une restauration du motif de base, fortement déformé, un décalage vers le haut et une stabilisation de la fréquence fondamentale moyenne. Toutefois, ces trois opérations, à elles seules, ne suffisent pas à restituer une voix devant être perçue comme humaine. En effet, chez un sujet sain, le caractère humain de la voix, par opposition à une voix robotisée, est donné par ses imperfections et par l'émotion qu'elle véhicule. Au niveau du signal acoustique, les particularités de la voix humaine se traduisent par plusieurs éléments.

- Premièrement, la variabilité de la fréquence fondamentale moyenne, c'est-à-dire les très faibles variations du laps de temps entre deux motifs de base consécutifs. Un écart constant entre deux motifs de base résulte en une voix robotisée.
- Deuxièmement, la variabilité du motif, c'est-à-dire les très faibles variations de forme et d'amplitude du motif de base au cours du temps. Comme précédemment, la reproduction à l'identique d'un motif de base conduit à une voix robotisée.
- Troisièmement, la modulation en amplitude et en fréquence du signal harmonique, qui introduisent de l'émotion dans la voix. On notera que la modulation en fréquence et en amplitude n'est pas la variabilité de ces grandeurs, mais est une variation, plus lente dans le temps et contrôlée, de la fréquence fondamentale moyenne et de l'amplitude moyenne.

[0037] Les trois éléments énumérés ci-dessus doivent se retrouver sur l'excitation harmonique restaurée afin de donner l'illusion d'une voix humaine. Les opérations de restauration de l'excitation harmonique selon l'invention visent donc à restaurer le motif de base, tout en introduisant une variabilité, restaurer la fréquence fondamentale moyenne, en introduisant également une variabilité, et moduler le signal obtenu en amplitude et en fréquence.

[0038] Une première opération, réalisée par le bloc d'identification des paramètres harmoniques 18a, consiste donc à estimer et extraire du signal entrant la fréquence fondamentale moyenne et la puissance moyenne instantanées, calculées sur un intervalle de temps

donné, par exemple 20ms. Cette opération est plus complexe que chez un sujet sain, en raison de la déformation de l'onde quasi-harmonique. Une méthode, basée sur l'histogramme des enveloppes supérieures et inférieures détectées, est utilisée pour déterminer, à intervalle régulier, par exemple toutes les 8ms, cette fréquence fondamentale moyenne instantanée. Pour plus de précisions sur la méthode utilisée, on se référera à l'article de V. Parsa et al. (Journal of Speech, Language and Hearing Research, Vol42, 112-126, February 1999).

[0039] Puis, le bloc 18b procède à la correction de ces paramètres harmoniques selon le processus suivant:

- La puissance moyenne instantanée est conservée sans être modifiée puisque le sujet ayant subi une ablation du larynx est capable d'effectuer des modulations en amplitude exprimant ses émotions.
- Puis, un décalage de la fréquence fondamentale moyenne instantanée vers les valeurs hautes est réalisé afin de la ramener dans la fourchette standard.
- Enfin, une variabilité issue d'une voix saine est introduite au niveau de cette même fréquence.

[0040] Pour ce faire, une table de référence a été constituée au préalable à partir d'une voix saine, de préférence celle du sujet avant son opération. Cette table, figurée par le bloc 18c, contient un grand nombre de motifs de base enregistrés au cours de la parole et présentant une variabilité caractéristique de la voix humaine. Elle contient aussi des statistiques sur la variabilité de la fréquence fondamentale moyenne, également caractéristique de la voix humaine enregistrée, et calculées à partir de l'enregistrement. Le bloc 18b est ainsi connecté au bloc 18c de façon à recevoir les informations sur la variabilité en fréquence fondamentale telles que contenues dans la table de référence. Le signal issu du bloc 18b contient donc toutes les données nécessaires pour reconstituer un signal harmonique modulé en amplitude, dont la fréquence fondamentale moyenne instantanée et la variabilité en fréquence fondamentale correspondent à celles d'un sujet sain. A ce stade de la restauration, par contre, le motif de base du signal est encore caractéristique d'une élocution oesophagienne.

[0041] Les signaux issus des blocs 18b et 18c sont alors dirigés vers le bloc 18d de reconstruction de l'onde glottale. Là, une onde glottale possédant toutes les caractéristiques de l'onde glottale humaine est reconstituée à partir des motifs de base sains et des paramètres de l'onde harmonique, corrigés ou non, au niveau du bloc 18b. La perception de l'excitation émise à l'issue du bloc 18d est celle d'une voix humaine, où les émotions sont exprimées grâce à la modulation en amplitude et en fréquence, réalisée par le sujet laryngectomisé.

[0042] Ainsi est proposé un algorithme particulièrement efficace de reconstruction d'une onde glottale humaine expressive.

[0043] L'invention porte également sur le dispositif

d'aide vocale permettant la mise en oeuvre du procédé décrit précédemment. Ce système, représenté schématiquement en figure 4, comporte essentiellement un dispositif d'acquisition de la voix 30, tel qu'un microphone, destiné à capter le signal acoustique oesophagien émis par le patient et le transformer en un signal électrique. Ce microphone est relié à un premier module d'amplification 32 chargé de régler l'échelle dynamique et lui-même connecté en sortie à un module A/D 34 de conversion du signal analogique en signal digital.

[0044] En sortie du convertisseur 34, le signal est appliqué à un processeur de signal digital DSP 36 (de l'anglais Digital Signal Processing) qui effectue les différentes opérations de restauration de la voix décrites précédemment.

[0045] Les signaux digitaux issus du processeur 36 sont reçus par un module D/A 38 de transformation du signal digital en signal analogique, lui-même relié à un deuxième module d'amplification 40. En sortie de ce dernier, un haut-parleur 42 transforme le signal électrique en signal acoustique. Il convient de noter que tout type d'appareil de restitution du son est également applicable. En particulier, le signal peut être traité par un appareil téléphonique adapté.

[0046] Le microphone 30 peut être de type portable, pour une utilisation quotidienne, ou fixe, par exemple pour un discours en public. Les modules 32 à 40 sont, par exemple, intégrés dans un seul et même boîtier, portable ou non, et le haut-parleur 42 peut être fixé aux épaules du patient ou en toute autre position stratégique.

[0047] L'utilisation d'un haut-parleur 42 en association avec un processeur de signal digital 36 permet, en outre, la compensation du signal acoustique émis par le sujet. Cette possibilité peut s'avérer très utile pour une discussion en petit comité, dans laquelle la voix du laryngectomisé se superpose à la voix corrigée par le dispositif. Un module actif d'annulation de la première voix, faisant partie du processeur de signal digital, permet alors d'éliminer la gêne causée par cette superposition, et accroît le confort des auditeurs comme de l'orateur.

Revendications

1. Procédé de restauration d'une élocution altérée par traitement d'un signal électrique, provenant d'un microphone, représentatif de ladite élocution, **caractérisé en ce qu'il** comporte les principales étapes suivantes:

- extraction et séparation (16) de l'excitation originale et des paramètres de l'articulation,
- restauration (18) de l'excitation originale à partir d'éléments issus de ladite élocution altérée et d'éléments issus d'une élocution saine, et
- reconstruction de l'élocution (24) à partir desdits paramètres de l'articulation et de ladite excitation originale restaurée afin de produire un

signal acoustique représentatif de l'élocution re-construite.

2. Procédé selon la revendication 1, **caractérisé en ce qu'il** comporte, en outre :

- une étape d'estimation de la probabilité d'un son voisé (12), effectuée en parallèle avec les étapes d'extraction de l'excitation et des paramètres de l'articulation (16), et de restauration de l'excitation originale (18), et
- une étape de mélange de l'excitation originale restaurée et de l'excitation originale (20) à partir de ladite probabilité d'un son voisé.

3. Procédé selon l'une des revendications 1 et 2, **caractérisé en ce que** :

- les éléments issus de l'élocution altérée sont l'amplitude moyenne instantanée et la fréquence fondamentale moyenne instantanée, et
- les éléments issus d'une élocution saine sont la variabilité en fréquence, la forme du motif et sa variabilité.

4. Procédé selon la revendication 3, **caractérisé en ce que** l'étape de restauration de l'excitation originale comprend les opérations suivantes :

- calcul de l'amplitude moyenne instantanée et de la fréquence fondamentale moyenne instantanée de l'excitation originale (18a),
- décalage de la fréquence fondamentale moyenne instantanée et introduction d'une variabilité en fréquence fondamentale issue d'une élocution saine (18c), et
- reconstruction d'une excitation harmonique (18d) présentant le motif issu d'une élocution saine et sa variabilité caractéristique, l'amplitude moyenne instantanée calculée, la fréquence fondamentale moyenne instantanée décalée et la variabilité de fréquence fondamentale issue d'une élocution saine.

5. Dispositif de restauration d'une élocution altérée, **caractérisé en ce qu'il** comprend :

- un système d'acquisition de cette élocution altérée (30) produisant un signal électrique qui la représente,
- un processeur (36) de ce signal, programmé de façon à mettre en oeuvre le procédé selon l'une des revendications 1 à 4, et
- un système de haut-parleur (42) transformant le signal délivré par le processeur en signal acoustique.

6. Dispositif selon la revendication 5, **caractérisé en**

ce qu'il comprend, en outre :

- un premier module d'amplification (32) en sortie du système d'acquisition de la parole (30),
- un module de transformation du signal analogique en signal digital (34) entre le premier module d'amplification (32) et le processeur de signal (36),
- un module de transformation du signal digital en signal analogique (38) en sortie du processeur de signal (36), et
- un deuxième module d'amplification (40) entre le module de transformation du signal digital en signal analogique (38) et le système de haut-parleur (42).

7. Dispositif selon l'une des revendications 5 et 6, **caractérisé en ce qu'il** comporte, en outre, un module de compensation de l'élocution altérée.

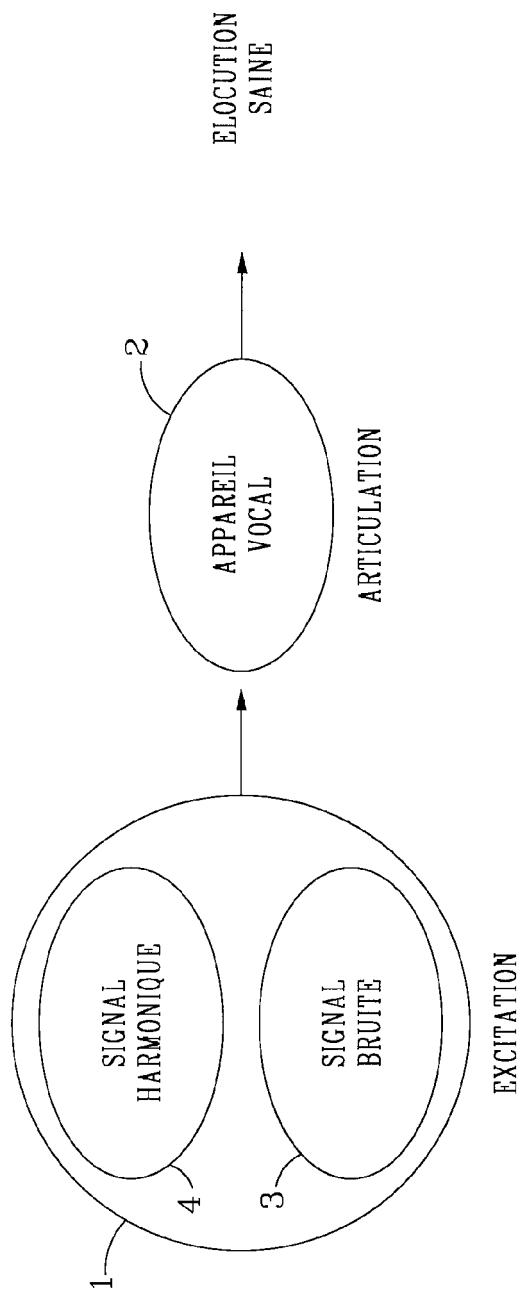


Fig.1

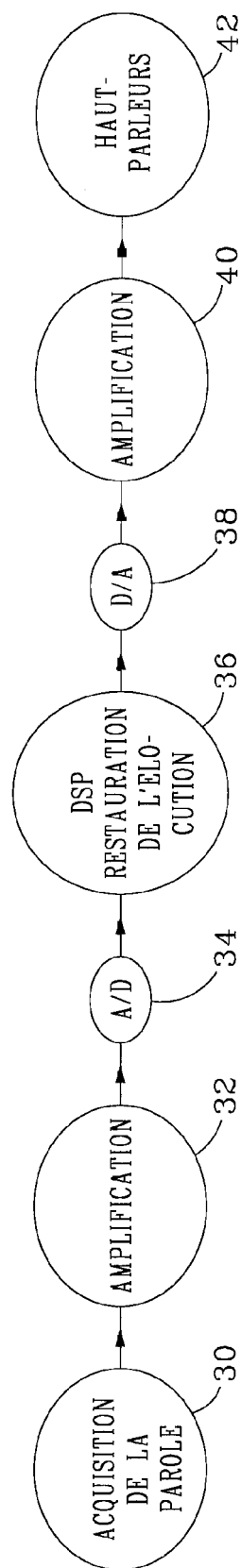


Fig.4

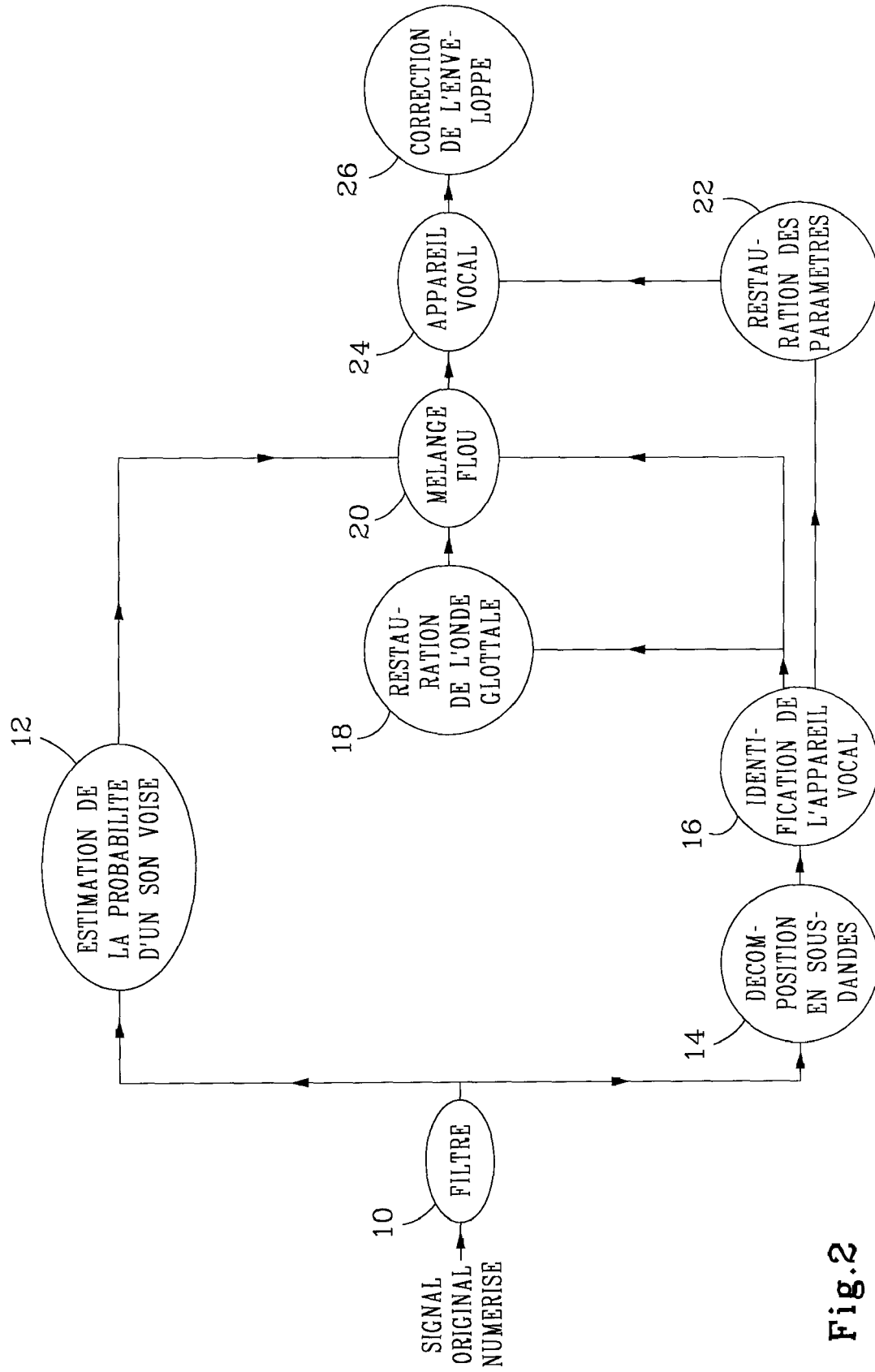


Fig.2

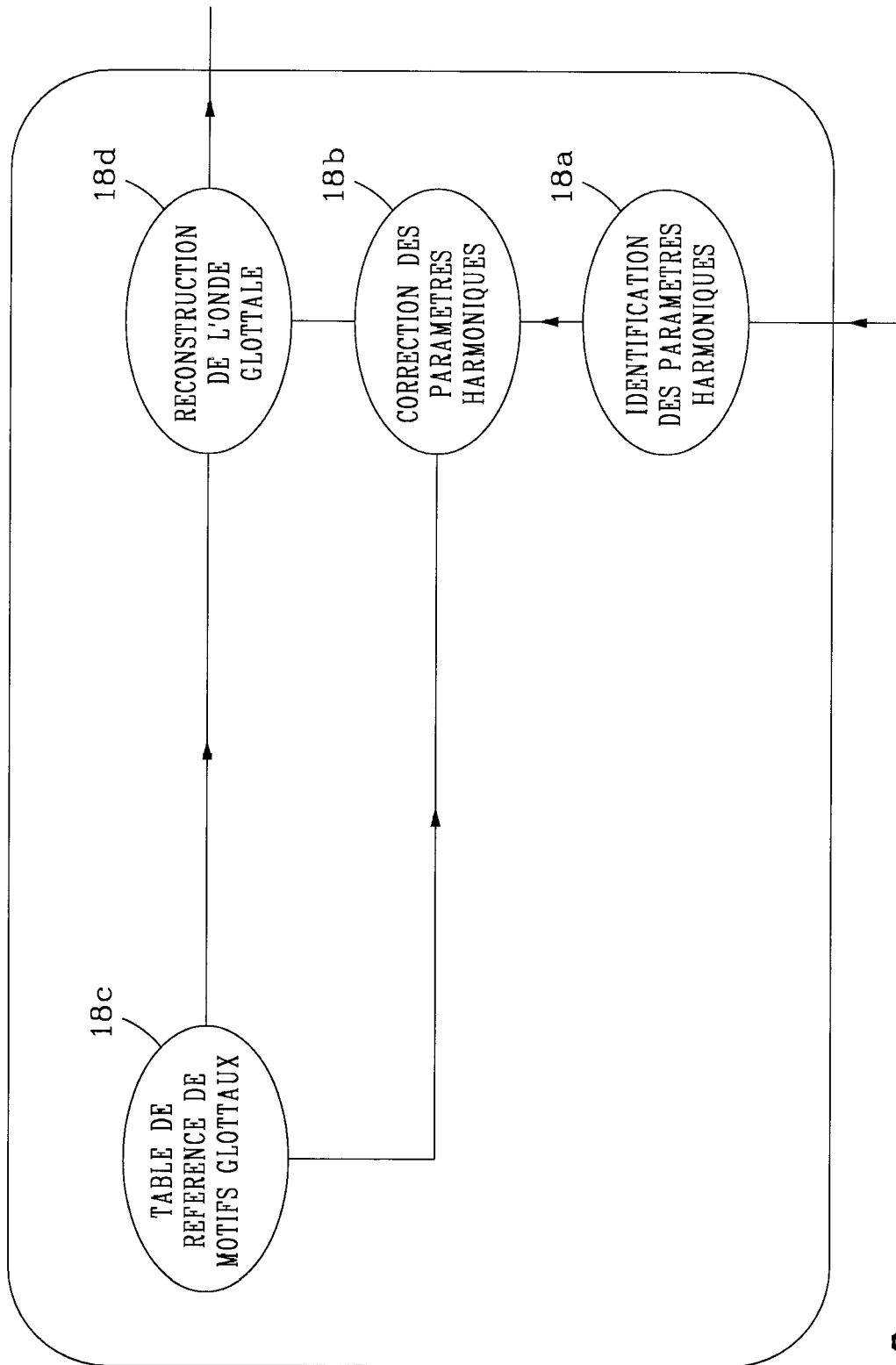


Fig.3



DOCUMENTS CONSIDERES COMME PERTINENTS			
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (IPC)
D,Y	YINGYONG QI: "REPLACING TRACHEOESOPHAGEAL VOICING SOURCES USING LPC SYNTHESIS" JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, AIP / ACOUSTICAL SOCIETY OF AMERICA, MELVILLE, NY, US, vol. 88, no. 3, 1 septembre 1990 (1990-09-01), pages 1228-1235, XP000162105 ISSN: 0001-4966 * page 1228, colonne 2, ligne 16 - ligne 30 *	1,5	G10L21/02
Y	RENTZOS D ET AL: "Transformation of speaker craracteristics for voice conversion" AUTOMATIC SPEECH RECOGNITION AND UNDERSTANDING, 2003. ASRU '03. 2003 IEEE WORKSHOP ON ST. THOMAS, VI, USA NOV. 30-DEC. 3, 2003, PISCATAWAY, NJ, USA,IEEE, 30 novembre 2003 (2003-11-30), pages 706-711, XP010713233 ISBN: 0-7803-7980-2 * alinéa [0002]; composé 2 *	1,5	DOMAINES TECHNIQUES RECHERCHES (IPC) G10L
A	US 2004/260552 A1 (NAVRATIL JIRI ET AL) 23 décembre 2004 (2004-12-23) * alinéas [0037], [0045] * * figure 4 *	1,5	
A	US 6 336 092 B1 (GIBSON BRIAN CHARLES ET AL) 1 janvier 2002 (2002-01-01) * abrégé * * colonne 2, ligne 28 - ligne 31 *	1,5	
Le présent rapport a été établi pour toutes les revendications			
Lieu de la recherche Munich		Date d'achèvement de la recherche 24 novembre 2005	Examineur Krembel, L
CATEGORIE DES DOCUMENTS CITES 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 antérieur, mais publié à la date de dépôt ou après cette date D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant	

1

EPC FORM 1503 03.02 (P04C02)



DOCUMENTS CONSIDERES COMME PERTINENTS			
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (IPC)
A	SAVIC M ET AL: "VOICE PERSONALITY TRANSFORMATION" DIGITAL SIGNAL PROCESSING, ACADEMIC PRESS, ORLANDO, FL, US, vol. 1, no. 2, 1 avril 1991 (1991-04-01), pages 107-110, XP000393609 ISSN: 1051-2004 * figures 3,6 *	1,5	
D,A	----- US 6 359 988 B1 (ESPY-WILSON CAROL) 19 mars 2002 (2002-03-19) * abrégé * -----	1,5	
			DOMAINES TECHNIQUES RECHERCHES (IPC)
Le présent rapport a été établi pour toutes les revendications			
Lieu de la recherche Munich		Date d'achèvement de la recherche 24 novembre 2005	Examineur Krembel, L
CATEGORIE DES DOCUMENTS CITES 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 antérieur, mais publié à la date de dépôt ou après cette date D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant	

EPO FORM 1503 03.82 (P04C02)

**ANNEXE AU RAPPORT DE RECHERCHE EUROPEENNE
RELATIF A LA DEMANDE DE BREVET EUROPEEN NO.**

EP 05 10 2714

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche européenne visé ci-dessus.

Lesdits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du

Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets.

24-11-2005

Document brevet cité au rapport de recherche		Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 2004260552	A1	23-12-2004	AUCUN	

US 6336092	B1	01-01-2002	AT 233424 T	15-03-2003
			AU 7024798 A	24-11-1998
			WO 9849670 A1	05-11-1998
			DE 69811656 D1	03-04-2003
			DE 69811656 T2	16-10-2003
			EP 0979503 A1	16-02-2000
			JP 2001522471 T	13-11-2001

US 6359988	B1	19-03-2002	AU 8036600 A	10-04-2001
			WO 0118791 A1	15-03-2001

EPO FORM P0460

Pour tout renseignement concernant cette annexe : voir Journal Officiel de l'Office européen des brevets, No.12/82

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.

Littérature non-brevet citée dans la description

- **J.M. SOLÀ et al.** Environmental Robust Features for Speech Detection. *Proc. INTERSPEECH-ICSLP'04*, Octobre 2004 [0023]
- **YINGYONG QI.** *J. Acoust. Soc. Am.*, Septembre 1990, vol. 88 (3 [0027]
- **YINGYONG QI.** Replacing tracheoesophageal voicing sources using LPC synthesis. *J. Acoust. Soc. Am.*, Septembre 1990, vol. 88 (3 [0031]
- **V. PARSA et al.** *Journal of Speech, Language and Hearing Research*, Février 1999, vol. 42, 112-126 [0038]