

⑫ DEMANDE DE BREVET D'INVENTION

A1

②2 Date de dépôt : 27.12.12.

③0 Priorité :

④3 Date de mise à la disposition du public de la
demande : 04.07.14 Bulletin 14/27.

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 :

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

⑦2 Inventeur(s) : ROUAUD CHARLES.

⑦3 Titulaire(s) : LIPEO Société par actions simplifiée.

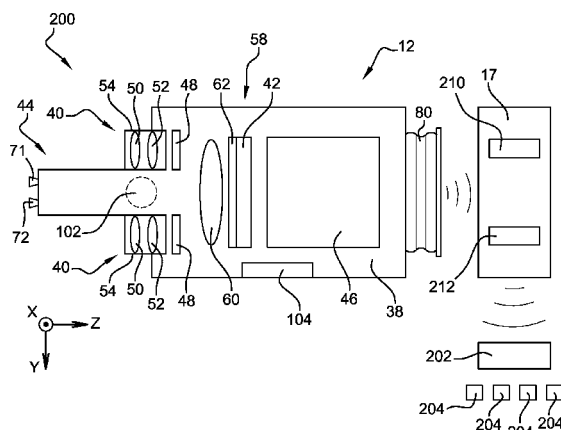
⑦4 Mandataire(s) : CABINET LAVOIX.

⑤4 PROCÉDE DE COMMUNICATION ENTRE UN LOCUTEUR ET UN APPAREIL ELECTRONIQUE ET APPAREIL
ELECTRONIQUE ASSOCIE.

⑤7 L'invention se rapporte à un appareil électronique
(200) caractérisé en ce que :

- le dispositif (12) d'acquisition de données comprend,
en outre, un capteur (102) de souffle différentiel adapté à
mesurer le gradient de température et/ou de pression entre
l'air expiré par le nez du locuteur et l'air expiré par la bouche
du locuteur, et

- l'unité de traitement (17) est propre à déterminer des
données à communiquer également à partir des données is-
sues du capteur (102) de pression différentielle.



PROCEDE DE COMMUNICATION ENTRE UN LOCUTEUR ET UN APPAREIL
ELECTRONIQUE ET APPAREIL ELECTRONIQUE ASSOCIE

La présente invention concerne un appareil électronique. L'invention se rapporte également à un procédé de communication entre un locuteur et l'appareil électronique.

Il est souhaitable de pouvoir contrôler un appareil électronique avec la voix. Ceci implique de pouvoir reconnaître avec précision et rapidité les ordres émis par le locuteur.

Il est ainsi connu du document US-A-2010/0036657 un système de reconnaissance vocale comprenant un transmetteur transmettant un signal de test, un récepteur pour recevoir le signal de test et une unité de reconnaissance vocale pour reconnaître de la voix à partir du signal reçu. Le transmetteur transmet le signal de test vers les organes vocaux du locuteur. Le récepteur reçoit le signal de test qui a été réfléchi par les organes vocaux. L'unité de reconnaissance vocale reconnaît la voix ou les formes d'onde de la voix sur la base de la forme de l'onde de réflexion du signal de test reçu par le récepteur.

Mais, le système de reconnaissance vocale proposé dans ce document s'avère peu fiable. Pour la communication d'ordres à un appareil électronique, cela résulte en une mauvaise interprétation des ordres du locuteur par l'appareil électronique

Il existe donc un besoin pour un appareil électronique pouvant être commandé par un locuteur de manière fiable.

Selon l'invention, ce but est atteint par un procédé de communication entre un locuteur et un appareil électronique. Le procédé comprend les étapes d'émission de données par le locuteur et d'acquisition d'au moins une image d'une partie de la langue du locuteur simultanément à l'étape d'émission. Le procédé comporte aussi une étape de mesure du gradient de température et/ou de pression entre l'air expiré par le nez du locuteur et l'air expiré par la bouche du locuteur simultanément à l'étape d'émission. Le procédé comprend également une étape de détermination de données à communiquer à partir des données acquises et mesurées. Le procédé comprend aussi une étape de commande de l'appareil électronique à partir des données déterminées.

Suivant des modes de réalisation particuliers, le procédé de communication comprend une ou plusieurs des caractéristiques suivantes, prise(s) isolément ou suivant toute combinaison techniquement possible :

- l'étape de détermination est une étape de reconnaissance vocale
- l'étape de reconnaissance vocale comprend une première étape de reconnaissance vocale sur la base des données acquises seulement, la première étape de reconnaissance vocale permettant d'obtenir des premières données de

reconnaissance et un premier taux de probabilité d'erreur entre les premières données et les données émises par le locuteur.

- l'étape de reconnaissance vocale comprend une étape de comparaison du premier taux de probabilité d'erreur à une valeur seuil.

5 - lorsque le premier taux de probabilité est inférieur ou égal à la valeur seuil, les données à communiquer sont les premières données à communiquer,

10 - l'étape de reconnaissance vocale comprend lorsque le premier taux de probabilité est supérieur à une valeur seuil, une deuxième étape de reconnaissance vocale sur la base des données acquises et mesurées seulement, la deuxième étape de reconnaissance vocale permettant d'obtenir des deuxièmes données, les données à communiquer étant les deuxièmes données.

- le procédé comprend également une étape d'enregistrement de données de son par le microphone.

15 - l'étape de détermination de données à communiquer prend également en compte les données enregistrées.

20 - l'étape de détermination comprend une première étape de reconnaissance vocale sur la base des données enregistrées seulement, la première étape de reconnaissance vocale permettant d'obtenir des premières données de reconnaissance et un premier taux de probabilité d'erreur entre les premières données de reconnaissance et les données émises par le locuteur.

25 - l'étape de détermination comprend une étape de comparaison du premier taux de probabilité d'erreur à une première valeur seuil telle que lorsque le premier taux de probabilité d'erreur est inférieur ou égal à la première valeur seuil, les données à communiquer sont les premières données de reconnaissance.

30 - l'étape de détermination comprend une deuxième étape de reconnaissance vocale sur la base des données acquises et enregistrées lorsque le premier taux de probabilité d'erreur est supérieur à la première valeur seuil, la deuxième étape de reconnaissance vocale permettant d'obtenir des deuxièmes données de reconnaissance et un deuxième taux de probabilité d'erreur entre les deuxième données de reconnaissance et les données émises par le locuteur.

- l'étape de détermination comprend une étape de comparaison du deuxième taux de probabilité d'erreur à une deuxième valeur seuil telle que lorsque le deuxième taux de probabilité est inférieur ou égal à la deuxième valeur seuil, les données à communiquer sont les deuxièmes données de reconnaissance.

35 - l'étape de détermination comprend une troisième étape de reconnaissance vocale sur la base des données acquises, mesurées et enregistrées lorsque le deuxième

taux de probabilité est supérieur à une deuxième valeur seuil, la troisième étape de reconnaissance vocale permettant d'obtenir des troisièmes données de reconnaissance, les données à communiquer étant les troisièmes données de reconnaissance.

- le premier taux de probabilité d'erreur et le deuxième taux de probabilité d'erreur sont égaux.

Il est également proposé un système de détermination de la position dans l'espace de la langue d'un locuteur comportant un dispositif d'acquisition de données du locuteur. Le dispositif comprend une unité d'illumination d'au moins une partie de la langue du locuteur et d'au moins une partie supérieure d'une dent inférieure, et une unité d'imagerie adaptée à la prise d'image de la partie de langue du locuteur et de la partie supérieure de la dent inférieure du locuteur illuminées par l'unité d'illumination. Le système comprend un support du dispositif d'acquisition de données. Le dispositif d'acquisition d'images comprend en outre un organe de maintien sur au moins du dent du locuteur.

Suivant des modes de réalisation particuliers, le système comprend une ou plusieurs des caractéristiques suivantes, prise(s) isolément ou suivant toute combinaison techniquement possible :

- l'organe de maintien est une lamelle d'appui dental adaptée à prendre appui sur au moins une dent du locuteur.

- la lamelle d'appui dental présente une largeur d'au moins 15 millimètres.

- la lamelle d'appui dental comprend une face avant destinée à prendre appui sur la moins une dent du locuteur, la face avant ayant une surface d'au moins 5 millimètres carrés.

- l'organe de maintien comprend deux ventouses adaptées à être appliquées sur au moins une dent du locuteur.

- l'unité d'imagerie comprend une caméra et un système optique d'imagerie, le système optique d'imagerie étant agencé pour que la caméra puisse prendre une image de trois zones distinctes de la langue du locuteur simultanément.

- l'unité d'illumination comporte au moins une source laser émettant un faisceau laser présentant une extension longitudinale et une extension transverse, la ou chaque source laser étant munie chacune d'un système optique réduisant le rapport entre l'extension transverse et l'extension longitudinale d'un facteur compris entre 2 et 5.

- le système de détermination de la position comprend une unité de traitement déterminant la position dans l'espace de la langue du locuteur à partir d'au moins une image fournie par l'unité d'imagerie.

- le support est un casque.

La présente invention se rapporte à un procédé de détermination de la position dans l'espace de la langue d'un locuteur à l'aide du système de détermination de la position tel que précédemment décrit, le procédé comprenant les étapes de positionnement de l'organe de maintien sur au moins une dent du locuteur et de prise d'au moins une image par le dispositif d'acquisition de données.

Selon un mode de réalisation, la ou chaque dent du locuteur font partie du groupe constitué des canines et des incisives de la mâchoire supérieure du locuteur.

Il est également proposé un module de reconnaissance vocale comprenant un dispositif d'acquisition de données d'un locuteur comprenant une unité d'illumination d'au moins une partie de la langue du locuteur et d'au moins une partie supérieure d'une dent inférieure du locuteur, et une unité d'imagerie adaptée à la prise d'image de la partie de langue du locuteur et de la partie supérieure de la dent inférieure du locuteur illuminées par l'unité d'illumination. Le dispositif d'acquisition de données comprend, en outre, un capteur de souffle différentiel adapté à mesurer le gradient de pression et/ou de température entre l'air expiré par le nez du locuteur et l'air expiré par la bouche du locuteur.

Suivant des modes de réalisation particuliers, le module comprend une ou plusieurs des caractéristiques suivantes, prise(s) isolément ou suivant toute combinaison techniquement possible :

- le dispositif d'acquisition de données comprend, en outre, un organe de maintien sur au moins une dent du locuteur.

- le dispositif d'acquisition de données comprend, en outre, un microphone.

- le module comprend, en outre, une unité de traitement propre à déterminer des données à retranscrire à partir des données issues du dispositif d'acquisition et une unité de retranscription des données à retranscrire déterminées par l'unité de traitement.

- les données à retranscrire font partie d'un groupe constitué de la fréquence, l'intensité, l'attaque et la tombée des sons émis par le locuteur.

- les données à retranscrire sont au moins un morphème et/ou au moins un phonème émis par le locuteur.

- l'unité de traitement comporte une sous-unité de reconnaissance d'au moins un morphème et/ou au moins un phonème émis par le locuteur, la sous-unité de reconnaissance étant adaptée à reconnaître les morphèmes et phonèmes émis dans une première langue et une sous-unité de conversion des morphèmes et/ou phonèmes dans une deuxième langue, la deuxième langue étant différente de la première langue, les données à retranscrire étant les au moins un morphème et/ou au moins un phonème convertis par la sous-unité de conversion.

La présente invention se rapporte également à une utilisation du module de reconnaissance vocale tel que précédemment décrit pour produire de la musique.

La présente invention se rapporte aussi à une utilisation du module de reconnaissance vocale tel que précédemment décrit comme prothèse pour un handicapé privé de la parole.

La présente invention concerne aussi une utilisation du module de reconnaissance vocale tel que précédemment décrit pour de la traduction automatique.

La présente invention se rapporte également à une utilisation du module de reconnaissance vocale tel que précédemment décrit comme interface homme-machine de console de jeu vidéo.

Il est également proposé un appareil électronique comprenant au moins un actionneur, une unité de contrôle contrôlant l'au moins un actionneur et un dispositif d'acquisition de données d'un locuteur. Le dispositif comprend une unité d'illumination d'au moins une partie de la langue du locuteur et d'au moins une partie supérieure d'une dent inférieure, et une unité d'imagerie adaptée à la prise d'image de la partie de langue du locuteur et de la partie supérieure de la dent inférieure illuminées par l'unité d'illumination. L'appareil électronique comporte aussi une unité de traitement configurée pour déterminer des données à communiquer à partir des données issues de l'unité d'imagerie et transmettre les données à communiquer à l'unité de contrôle. Le dispositif d'acquisition de données comprend, en outre, un capteur de souffle différentiel adapté à mesurer le gradient de température et/ou de pression entre l'air expiré par le nez du locuteur et l'air expiré par la bouche du locuteur. L'unité de traitement est propre à déterminer des données à communiquer également à partir des données issues du capteur de pression différentielle.

Suivant des modes de réalisation particuliers, l'appareil électronique comprend une ou plusieurs des caractéristiques suivantes, prise(s) isolément ou suivant toute combinaison techniquement possible :

- le dispositif d'acquisition de données comprend, en outre, un organe de maintien sur au moins une dent du locuteur.

- le dispositif d'acquisition de données comprend, en outre, un microphone.

- l'unité de traitement est propre à déterminer des données à communiquer également à partir des données issues du microphone.

- l'appareil électronique est choisi dans le groupe constitué d'une console de jeux vidéo, d'un téléphone mobile, d'une interface homme-machine, d'un ordinateur, d'un robot, et d'un fauteuil pour handicapé.

D'autres caractéristiques et avantages de l'invention apparaîtront à la lecture de la description qui suit de modes de réalisation de l'invention, donnés uniquement à titre d'exemple et en référence aux dessins qui sont :

- figure 1, une vue de côté du système de détermination de la position selon l'invention ;

- figure 2, une vue schématique de dessus du dispositif d'acquisition du système de détermination selon l'invention de la figure 1 ;

- figure 3, une vue de côté du dispositif d'acquisition du système de détermination selon l'invention de la figure 1 ;

- figure 4, une vue schématique d'une partie d'un module de reconnaissance vocale selon l'invention,

- figure 5, une vue schématique en section du dispositif d'acquisition du module de reconnaissance selon l'invention de la figure 4 en position sur les dents du locuteur et

- figure 6, une vue schématique d'une partie d'un appareil selon l'invention.

Pour l'ensemble des figures, il est défini un plan horizontal par rapport à un locuteur, par exemple le plan de Francfort.

Dans la suite, il est défini un axe X vertical perpendiculaire à ce plan de Francfort passant par le plan de symétrie bilatérale du locuteur. Cela permet également de définir des termes « haut » et « bas », sachant que la bouche du locuteur est située en bas par rapport au nez du locuteur. Les termes « gauche » et « droit » sont définis de sorte que lorsque l'on regarde le locuteur, la partie gauche est à gauche de l'axe X et la partie droite est à droite de l'axe X.

Il est également défini un axe Y transversal qui est parallèle à l'axe reliant les deux yeux du locuteur et perpendiculaire à l'axe X.

Il est également défini un axe Z horizontal comme un axe perpendiculaire aux axes transversal Y et vertical X. Cela permet également de définir des termes « arrière » et « avant », sachant que le nez du locuteur est situé en avant par rapport aux tempes du locuteur.

Il est, en outre, défini un axe a-a' passant par le milieu (point A) d'une incisive supérieure du locuteur et faisant un angle de 13° par rapport à l'axe horizontal Z.

Il est aussi défini l'axe b-b' passant un point B situé 12 mm en bas du point A et faisant un angle de 22° par rapport à l'axe horizontal Z.

Un système 10 de détermination de la position dans l'espace de la langue du locuteur suivant l'invention est représenté à la figure 1.

Le système 10 de détermination comprend un dispositif 12 d'acquisition de données du locuteur, un support 14 du dispositif 12 d'acquisition de données et un organe

16 de liaison reliant le support 14 au dispositif 12 et une unité de traitement 17 en communication avec le dispositif 12 d'acquisition de données.

Le dispositif 12 d'acquisition de données du locuteur se présente sous la forme d'un corps 38 muni d'éléments internes, éléments qui sont détaillés dans la suite de la description. Le corps 38 permet de supporter et de protéger les différents éléments internes.

Dans le cas de la figure 1, le support est une monture 14 adaptée pour être positionnée sur la tête du locuteur. Cette monture 14 comprend deux articulations latérales à hauteur des joues, reliées à deux arceaux passant par le sommet et l'arrière de la tête.

L'organe 16 de liaison comprend deux bras télescopiques assurant la liaison avec le dispositif au moyen de deux fixations latérales.

Le corps 38 selon l'exemple de la figure 1 a sensiblement une forme de parallélépipède rectangle.

Le corps 38 est fixé à l'organe 16 de liaison de sorte qu'au moins une face du corps 38 s'étende sensiblement dans un plan contenant l'axe a-a'.

Comme visible notamment à la figure 2, le dispositif 12 d'acquisition comprend deux unités 40 d'illumination, une unité 42 d'imagerie, un organe 44 de maintien sur au moins du dent du locuteur sous la forme d'une lamelle 44 d'appui dental et une unité 46 de prétraitement.

Chaque unité 40 d'illumination comprend une source 48 d'émission de lumière et un système optique 50.

A titre d'exemple, dans le cas de la figure 2, les deux sources 48 d'émission de lumière sont des lasers.

Un laser (acronyme de l'anglais « light amplification by stimulated emission of radiation », ce qui signifie en français « amplification de la lumière par émission stimulée de rayonnement ») est un appareil qui produit une lumière spatialement et temporellement cohérente basée sur l'effet laser.

Selon l'exemple de la figure 2, les sources 48 d'émission de lumière sont propres à émettre des faisceaux de lumière dont la longueur d'onde est la même.

Par exemple, les sources 48 d'émission de lumière sont propres à émettre des faisceaux de lumière dont la longueur d'onde est comprise entre 500 nanomètres (nm) et 580 nm, ce qui correspond pour l'œil humain à la couleur verte. Une telle longueur permet de réduire la diffusion du faisceau laser par l'élément illuminé par le faisceau.

En variante, pour augmenter encore cet effet, les sources 48 d'émission de lumière sont propres à émettre des faisceaux de lumière dont la longueur d'onde est comprise dans le bleu (entre 400 nm et 500 nm).

5 Les faisceaux émis par les sources 48 d'émission de lumière sont sensiblement monochromatiques. Dans le cadre de cette invention, un faisceau est considéré comme monochromatique lorsque la dispersion en longueur d'onde du faisceau est inférieure à 1 nm.

10 Chaque source 48 d'émission de lumière émet un faisceau ayant une puissance optique de 0.5 milliWatt (mW). Une telle puissance non focalisée en un point, mais répartie sur une ligne d'au moins 15 millimètres (mm) permet de rendre le dispositif d'acquisition compatible avec les normes de sécurité laser.

De préférence, la source 48 d'émission de lumière est une source de classe 1 selon la norme NF EN 60825-1/A2.

15 Selon l'exemple de la figure 1, les lasers utilisés sont des lasers à semi-conducteurs qui présentent l'avantage d'être compacts et peu onéreux. En particulier, selon un mode de réalisation préféré, les lasers sont des diodes laser.

Chaque système optique 50 est propre à réduire l'extension transverse du faisceau laser par rapport à l'extension longitudinale du faisceau laser.

20 Par exemple, chaque système optique 50 effectue une anamorphose qui réduit le rapport entre l'extension transverse du faisceau laser et l'extension longitudinale du faisceau laser d'un facteur compris entre 2 et 5 et de préférence d'un facteur 3. Ainsi, il peut être considéré que le faisceau laser en sortie de chaque système optique 50 est un faisceau laser sensiblement plan.

25 Un faisceau laser sensiblement plan permet d'obtenir une image en trois dimensions du profil de la langue ainsi que de l'une des incisives inférieures.

A titre d'illustration, chaque système optique 50 comprend deux lentilles 52, 54 : une lentille sphérique 52 et une lentille cylindrique 54.

30 Une lentille cylindrique est un cylindre dont la base est par exemple une demi-ellipse. Dans une telle situation, la lentille cylindrique comprend une face plane et le faisceau laser en sortie du système optique s'étend sensiblement selon un plan normal à la face plane.

Les unités 40 d'illumination sont propres à illuminer au moins une partie de la langue du locuteur et au moins une partie supérieure d'une dent inférieure du locuteur.

35 Les deux unités 40 d'illumination sont agencées sur le corps 38 de manière à illuminer respectivement la partie gauche de la cavité buccale et la partie droite de la cavité buccale.

Par exemple, les deux unités 40 d'illumination sont positionnées symétriquement par rapport à l'axe vertical X. L'unité 40 d'illumination positionnée sur la droite du corps 38 du dispositif 16 d'acquisition permet d'illuminer la partie gauche de la cavité buccale tandis que l'unité 40 d'illumination positionnée sur la gauche du dispositif permet

Selon l'exemple de la figure 1, chaque unité 40 d'illumination émet vers le bas d'un angle de 45° par rapport à l'axe horizontal et d'un angle de 30° par rapport à l'axe vertical, l'angle étant orienté vers la gauche pour une des deux unités 40 d'illumination et vers la droite pour l'autre unité 40 d'illumination.

Ainsi, dans l'exemple de la figure 1, l'angle entre les faisceaux laser émis par les deux unités 40 d'illumination est de 45° .

De manière plus générale, les faisceaux laser émis par les deux unités 40 d'illumination forment entre eux un angle compris entre 40° et 80° .

L'unité 42 d'imagerie comprend une caméra 56 et un système optique d'imagerie 58.

La caméra 56 comprend une matrice de photodétecteurs, la matrice permettant d'obtenir une résolution de 640 par 480 pixels.

Les photodétecteurs sont notamment sensibles aux longueurs d'onde des faisceaux émis par les unités 40 d'illumination.

La caméra 56 est, par exemple, une caméra de type CMOS (acronyme de l'expression anglaise « Complementary Metal Oxide Semiconductor »).

Le système optique d'imagerie 58 comprend une lentille de champ 60, un filtre 62, une première lentille 64, une deuxième lentille 66 et un miroir 68 de renvoi ainsi que visible aux figures 2 et 3.

La lentille de champ 60 est positionnée pour que, dans une partie du champ de la lentille de champ 60, la caméra 56 puisse imager la voûte palatine.

Dans le cas particulier présenté, la lentille de champ 60 a un diamètre de 9 millimètres (mm) et une focale de moins 12 mm.

Par exemple, la lentille de champ 60 est en méthacrylate.

La caméra 56 avec un axe optique orienté selon l'axe b-b' a un champ s'étendant de 20 mm selon la direction horizontale et de 15 mm selon la direction verticale.

Le filtre 62 est un filtre interférentiel centré sur la longueur d'onde d'émission des unités 48 d'illumination. Cela permet de limiter les incidences lumineuses parasites. La détection du signal utile s'en trouve améliorée.

La première lentille 64 est agencée pour qu'en combinaison avec la lentille de champ 60, la caméra 56 puisse acquérir au moins une image de la langue du locuteur en

position relevée du locuteur. La première lentille 64 est donc positionnée plus proche des dents avant de la mâchoire supérieure du locuteur que la lentille de champ 60.

La deuxième lentille 66 et le miroir 68 de renvoi sont agencés pour qu'en combinaison avec la lentille de champ 60, la caméra puisse acquérir au moins une image de la langue en position abaissée du locuteur et des dents avant (incisives et/ou canines) de la mâchoire inférieure du locuteur.

La lamelle 44 est en saillie par rapport au corps 38 du dispositif d'acquisition 12.

De préférence, la lamelle 44 présente une face d'appui destinée à être en contact avec les dents supérieures sur une surface d'appui d'au moins 6 millimètres carrés (mm^2) au niveau de la ou des dents du locuteur lorsque la lamelle 44 d'appui dental est en appui sur au moins une dent du locuteur.

Cela permet d'assurer que la lamelle 44 réduit les possibilités de mouvement du dispositif d'acquisition 12 en position.

Pour améliorer encore cet effet, la surface d'appui est d'au moins 16 mm^2 .

De plus, l'appui se fait sur deux incisives supérieures au milieu des deux dents à équidistance entre la gencive et l'extrémité basse de chaque dent.

Selon un mode de réalisation préférée, la surface d'appui est inférieure à 50 mm^2 . Cela permet d'éviter de trop gêner le locuteur lorsque la lamelle 44 d'appui dental est en appui sur au moins une dent du locuteur.

Une telle surface d'appui avec les dents est obtenue en choisissant la forme adéquate de la face d'appui. Notamment, différentes valeurs pour les extensions le long des axes X, Y et Z de la face d'appui sont envisageables tout en permettant d'obtenir une telle surface d'appui.

Un exemple est détaillé plus précisément en référence à la figure 2.

Dans ce cas, la lamelle 44 a sensiblement la forme d'un parallélépipède rectangle.

La lamelle 44 a ainsi une face rectangulaire dans un plan parallèle aux axes Y et Z (plan de la figure 2). Dans la suite, l'extension selon l'axe Y de la face rectangulaire est appelée la largeur l_y de la lamelle 44 et l'extension selon l'axe Z de la face rectangulaire est appelée la longueur l_z de la lamelle 44.

La largeur l_y de la lamelle 44 est comprise entre 15 millimètres et 25 millimètres. De préférence, la largeur l_y de la lamelle 44 est de 20 mm.

Selon un mode de réalisation préféré, la largeur l_y de la lamelle 44 est inférieure à 20 mm. Cela permet d'éviter une gêne lors de la prononciation des consonnes « ch » et « j ».

La longueur l_z de la lamelle 44 est comprise entre 15 millimètres et 50 millimètres. De préférence, la longueur l_z de la lamelle 44 est de 35 mm.

L'extension de la lamelle 44 selon l'axe X est appelée épaisseur l_x dans la suite. L'épaisseur l_x est comprise entre 0,5 mm et 1,5 mm. De préférence, l'épaisseur l_x de la lamelle 44 est de 1 mm.

5 La face d'appui a également une forme rectangulaire plane, la forme rectangulaire ayant comme longueur, la largeur l_y de la lamelle 44 et comme largeur, l'épaisseur l_x de la lamelle 44.

Dans cette situation, la surface d'appui correspond à la surface de la face d'appui soit le produit de la largeur l_y par l'épaisseur l_x de la lamelle.

10 Pour procurer une bonne stabilité, la surface d'appui est d'au moins 5 millimètres carrés au niveau de la ou des dents du locuteur lorsque la lamelle 44 d'appui dental est en appui sur au moins une dent du locuteur.

Pour améliorer encore cet effet, la surface d'appui est supérieure à 10 mm².

Pour limiter la gêne produite par la présente de la lamelle 44, la surface d'appui est inférieure à 20 mm².

15 L'organe 44 de maintien comprend deux ventouses 71, 72 destinées à être appliquées sur au moins une dent du locuteur.

Les deux ventouses 71, 72 prolongent la lamelle 44.

Les deux ventouses 71, 72 sont des embouts souples en silicone.

20 Les deux ventouses 71, 72 sont maintenues en position par une dépression générée par une pompe 80. La pompe 80 est une pompe manuelle à soufflet.

Selon une variante, la lamelle 44 est séparable du dispositif d'acquisition 12. Cela permet de considérer la lamelle 44 comme un consommable. Il en résulte une amélioration de la stérilité et de l'hygiène de la lamelle 44.

25 L'unité de prétraitement 46 est propre à compresser la taille des données de mesure issues de la caméra 56.

De telles compressions rendent possible une transmission sérielle des données à débit réduit depuis le dispositif d'acquisition 16 vers l'unité de traitement 17.

L'unité de traitement 17 est en liaison avec l'unité de prétraitement 46 du dispositif d'acquisition 16.

30 La liaison est une liaison sans fil, par exemple une liaison par voie hertzienne.

De préférence, il est fait usage d'une technologie ZigBee. ZigBee est un protocole de haut niveau permettant la communication à débit élevé de petites radios, à consommation réduite, basée sur la norme IEEE 802.15.4 pour les réseaux à dimension personnelle (Wireless Personal Area Networks : WPANs).

35 En variante, la liaison est de type filaire.

Selon les cas, la liaison se fait selon un canal half-duplex, également appelé à l'alternat, permettant de transporter des informations dans les deux sens, mais pas simultanément ou selon un canal full-duplex, pour lequel l'information est transportée simultanément dans chaque sens.

5 L'unité de traitement 17 est propre à déterminer par analyse d'images ou par reconnaissance de formes la position dans l'espace de la langue du locuteur à partir d'au moins une image fournie par l'unité 42 d'imagerie.

Le fonctionnement du système 10 de détermination de la position dans l'espace de la langue du locuteur va maintenant être décrit.

10 Il est proposé d'introduire la lamelle 44 du dispositif 16 en prenant appui sur une dent du locuteur tout en empêchant la fermeture centrale des lèvres.

Ainsi, la lamelle 44 est en appui sur une dent incisive supérieure en son milieu A. Ce point A constitue la référence du dispositif 12 dans le référentiel tridimensionnel du locuteur.

15 Le choix de ce point A répond à deux critères : l'obtention d'une bonne stabilité du dispositif 12 par un appui sur une partie osseuse sans interface peaussier et une tolérance à un usage prolongé du dispositif 12. La tolérance à l'usage prolongé se fait sans gêne, les possibilités d'humecter une grande partie des lèvres et de mettre en appui les dents supérieures et inférieures pour déglutir étant conservées.

20 Plus précisément, la lamelle 44 est en appui sur toute la largeur lx sur les deux canines supérieures selon l'axe Y.

Dans cette position, la lamelle 44 ne pénètre pas à l'intérieur de la cavité buccale du locuteur. En outre, la langue ainsi que les dents inférieures du locuteur n'entrent pas en contact avec la lamelle 44.

25 Ainsi, les mouvements de la mâchoire inférieure, de la langue et des incisives inférieures sont permis lorsque la lamelle 44 du dispositif 12 est en position.

Il existe en permanence un orifice entre les lèvres du locuteur. Par l'orifice ainsi généré, l'observation permanente des mouvements de la pointe et/ou du dos de la langue à l'intérieur de la cavité buccale du locuteur est possible, et ce pour beaucoup de phonèmes prononcés par le locuteur.

30 Lorsque la lamelle 44 est en position, la caméra 56 est à hauteur du bas de la cavité buccale, ce qui permet d'éviter la formation de buée sur les éléments du système optique 58.

35 Le dispositif 16 permet de faire une image de trois zones distinctes de la cavité buccale du locuteur.

La première zone est la voûte palatine, et le cas échéant une partie de la langue si la langue est proche de la voûte palatine sans la recouvrir. Les photons collectés lors de l'acquisition de l'image ont suivi le trajet optique suivant : émission par au moins une des sources laser dans la cavité buccale, diffusion par la cavité buccale, passage par le filtre 62 et la lentille de champ 60 et détection au niveau des photodétecteurs de la caméra 56.

La deuxième zone possible est le dos de la langue lorsque la langue est en position relevée dans la cavité buccale du locuteur et couvre la voûte palatine. Les photons collectés lors de l'acquisition de l'image ont suivi le trajet optique suivant : émission par au moins une des sources laser dans la cavité buccale, diffusion par la cavité buccale, passage par la première lentille 64 puis par le filtre 62 et par la lentille de champ 60 et détection au niveau des photodétecteurs de la caméra 56.

La troisième zone comprend la mâchoire inférieure du locuteur et notamment les dents avant (incisives et/ou canines). La troisième zone comprend, le cas échéant, la langue si la langue est en position abaissée dans la cavité buccale du locuteur. La caméra 56 acquiert ainsi une image permettant de déterminer la position de la langue par rapport à la mâchoire inférieure. Les photons collectés lors de l'acquisition de l'image ont suivi le trajet optique suivant : émission par au moins une des sources laser dans la cavité buccale, diffusion par la cavité buccale, réflexion au niveau du miroir 66 de renvoi, passage par la première lentille 66 puis par le filtre 62, par la lentille de champ 60 et détection au niveau des photodétecteurs de la caméra 56.

Ainsi, l'ensemble des optiques du dispositif 12 permet d'obtenir des mesures de triangulation optique.

Selon une variante, seule une des trois zones est imagée.

En outre, en variante, la lèvre inférieure est hors du champ de la caméra 58.

Les ventouses 71, 72 assurent un bon positionnement de la lamelle 44. Le bon positionnement est la position la plus proche possible de la cavité buccale du locuteur sans provoquer de gêne du locuteur.

Il peut être observé que la position du dispositif 12 centré sur la phytrium n'est pas critique. Une tolérance de + ou – 3 mm est acceptable.

L'appui sur le locuteur sans interface directement en appui sur les dents de la mâchoire supérieure assure une bonne stabilité à la lamelle 44.

En outre, du fait de cette bonne stabilité, lorsque le support 12 est en position, le support 12 constitue une référence sur la boîte crânienne du locuteur. De manière générale, le support 12 est positionné sur une partie de la tête présentant une souplesse faible pour éviter d'introduire un bougé de référence lors des mouvements du locuteur.

Dès lors, le support 12,, l'organe 16 de liaison ainsi que le dispositif 12 d'acquisition sont immobilisés en position par rapport au locuteur. Cela implique que l'unité 42 d'imagerie est positionnée de manière sensiblement fixe par rapport au locuteur, ce qui permet de générer une image plus aisée à exploiter.

5 En outre, la langue ainsi que les incisives inférieures n'entrent pas en contact avec la lamelle 44 tout en conservant toute leur liberté de mouvement. De ce fait, la gêne de l'introduction de la lamelle 44 dans la cavité buccale est minimisée.

De plus, la lamelle 44 n'empêche pas les dents d'être en contact pour la déglutition.

10 Selon une variante, la lamelle 44 est munie d'un cache dimensionné de sorte que lorsque la lamelle 44 est en position, la bouche du locuteur soit masquée depuis l'extérieur. Cela permet que les émissions vocales du locuteur se fassent en toute discrétion.

Les images prises par le système 10 de détermination de la position dans l'espace
15 de la langue du locuteur sont notamment adaptées pour une application dans le domaine de la reconnaissance vocale.

Selon un autre aspect de l'invention, il est, de ce fait, proposé à la figure 4 un module 100 de reconnaissance vocale comprenant le système 10 de détermination de la position tel que décrit en référence aux figures 1, 2 et 3. Les différents éléments décrits
20 précédemment sont identiques sauf l'unité de traitement 17 du module 100 qui est différente de l'unité de traitement 17 du système 10 de détermination de la position dans l'espace de la langue du locuteur.

Le module 100 est également équipé d'un capteur 102 de souffle différentiel, d'un microphone 104 et d'une unité 106 de retranscription de données.

25 Le capteur 102 de souffle différentiel et le microphone 104 font partie du dispositif 12 d'acquisition de données.

Le capteur 102 de souffle différentiel est, selon l'exemple de la figure 1, un ensemble de thermistances placées face à la cavité buccale du locuteur et sous les narines. L'air expulsé à la température du corps humain (environ 37°C) produit un
30 gradient de température par rapport à la température de l'environnement. Ce gradient de température est détecté par l'ensemble de thermistances.

Cet ensemble de thermistances permet de mesurer l'intensité, l'attaque et la tombée de la voix en fonction du temps. L'ensemble présente une faible inertie thermique, une haute sensibilité, une meilleure compacité que le capteur 102 de pression
35 différentielle.

Le capteur 102 de souffle différentiel est solidaire de la lamelle 44.

Dans le cas de la figure 4, le capteur 102 de souffle différentiel est agencé de manière à mesurer la répartition spatiale du souffle et le gradient entre le flux d'air inspiré par le nez du locuteur et le flux d'air expiré par la bouche du locuteur.

En variante, le capteur 102 de souffle différentiel est un capteur de pression différentielle propre à mesurer la différence de pression entre deux gaz.

En variante au lieu du capteur 102 de souffle différentiel, le dispositif est équipé d'un capteur de pression cinétique de l'air expulsé.

Le microphone 104 est propre à enregistrer des ondes sonores, et en particulier la voix du locuteur.

10 L'unité de traitement 17 du module 100 comporte une sous-unité de réception 108 et une sous-unité de reconnaissance 110 de morphème et/ou de phonème.

Dans ce mode de réalisation, l'unité de prétraitement 46 du dispositif 12 d'acquisition est propre à compresser la taille des données de mesure issues de la caméra, des données de mesure issues du capteur 102 de souffle différentiel et des
15 données de mesure issues du microphone 104.

De telles compressions rendent possible une transmission sérielle des données à débit réduit depuis le dispositif 12 d'acquisition vers la sous-unité de réception 108.

La sous-unité de réception 108 est en liaison avec l'unité de prétraitement 46.

La liaison est une liaison sans fil, par exemple une liaison par voie hertzienne.

20 En variante, la liaison est de type filaire.

Selon les cas, la liaison se fait selon un canal half-duplex, également appelé à l'alternat, permettant de transporter des informations dans les deux sens, mais pas simultanément ou selon un canal full-duplex, pour lequel l'information est transportée simultanément dans chaque sens.

25 La sous-unité de réception 108 est ainsi propre à recevoir les données de mesure compressées par l'unité de prétraitement 46.

La sous-unité reconnaissance 110 de morphème et/ou de phonème est également propre à assurer un post-traitement des données de mesures reçues par la sous-unité de réception 108. Ce post-traitement permet d'obtenir des données à retranscrire.

30 L'unité de traitement 17 envoie les données à retranscrire à l'unité 106 de retranscription.

L'unité 106 de retranscription de données est propre à retranscrire les données envoyées par l'unité de traitement 17.

Par exemple, l'unité 106 de retranscription est un générateur de son.

35 Le fonctionnement du module 100 de reconnaissance vocale va maintenant être décrit.

Le locuteur émet des sons correspondants, par exemple, à « Bonjour ».

Lorsque le locuteur émet ces données, plusieurs données sont mémorisées : au moins une image de la langue du locuteur est acquise, le gradient de pression entre l'air expiré par le nez du locuteur et l'air expiré par la bouche du locuteur est mesuré et le son émis par le locuteur est enregistré par le microphone 104.

Les données mémorisées sont compressées par l'unité de prétraitement 46 puis envoyées à la sous-unité de réception 108 de l'unité de traitement 17.

Sur la base des données reçues, la sous-unité de reconnaissance 110 reconnaît un ou plusieurs phonèmes et/ou morphèmes, en l'occurrence « bon » et « jour ».

L'unité de traitement 17 envoie alors les données « bon » et « jour » à l'unité 106 de retranscription qui émet les sons « Bonjour ».

Le module 100 de reconnaissance vocale selon l'invention permet d'obtenir une meilleure distinction des phonèmes qu'un module de reconnaissance vocale basé seulement sur une lecture labiale comme dans l'état de la technique.

Dans la suite, sont illustrés des exemples concrets dans lesquels l'utilisation du module 100 de reconnaissance vocale selon l'invention permet la reconnaissance de phonèmes que ne permet pas l'utilisation d'un module de reconnaissance vocale basé seulement sur une lecture labiale comme dans l'état de la technique.

Cela repose sur le fait que, lors de l'émission de voyelle par un locuteur, la demanderesse a constaté que les dents ne sont jamais serrées l'une contre l'autre. Il subsiste toujours un interstice entre les dents, supérieur à deux millimètres dans le plan normal à l'axe b-b'.

Dès lors, il est possible d'émettre des sons compréhensibles avec les dents serrées autour d'un crayon. Les dents sont immobiles mais les mouvements de la langue et les expirations permettent la génération de son.

Cela confirme que la lecture labiale seule est insuffisante. Il est, en effet, estimé que seulement 30% du message oral émis par ce biais sont perçus par un récepteur humain ainsi que cela est précisé dans la thèse de M. Nouredine ABOUTABIT intitulée « Reconnaissance de la langue française parlée complétée (LPC) : décodage phonétique des gestes main-lèvres » soutenue le 11 décembre 2007 à l'institut national polytechnique de Grenoble. Cela dépend de l'articulation du locuteur mais aussi de sa physiologie. Par exemple, il convient de prendre en compte si le locuteur a une forte barbe, une bouche lippue ou est atteint d'une paralysie faciale.

Par exemple, il est alors possible de discriminer entre un son [é] et un son [ai]. En français, les expressions « né » et « naît » sont en effet des homophones que seul le mouvement de la langue permet de distinguer. Les lèvres ont la même position. Ainsi, le

module 100 de reconnaissance vocale est capable de distinguer ces deux homophones sur la base de l'observation de la langue du locuteur.

Les voyelles « e » et « o » se différencient aussi par la position de la langue. Le « o » se prononce langue relevée tandis que le « e » se prononce avec la langue abaissée. Il est donc aisé de distinguer la prononciation des deux voyelles à l'aide du module 100 de reconnaissance vocale.

Comparativement, dans le cas d'une reconnaissance labiale comme dans l'état de la technique, la modification de la position des lèvres est faible. De ce fait, dans l'état de la technique, la reconnaissance des deux voyelles est difficilement obtenue.

L'emploi du capteur 102 de souffle différentiel permet aussi de différencier les sons [o] et [on]. Cela n'est pas possible avec l'observation du mouvement des lèvres ou de la langue.

Les voyelles orales, comme le son « o », se prononcent avec le voile du palais relevé, ce qui ferme le passage nasal. Les voyelles nasales, comme le son « on », se prononcent avec le voile du palais abaissé, ce qui laisse passer l'air par la bouche et par le nez. Ainsi, en mesurant la différence de température sous le nez et à l'entrée de la bouche, il est possible de déterminer quel son a été émis par le locuteur.

Comme l'émission du son « o » résulte d'une expulsion de l'air par la bouche alors que l'émission du son « on » est nasale, le capteur 102 de souffle différentiel mesure une différence de température négative si le son « o » est émis par le locuteur alors que le capteur 102 de souffle différentiel mesure une différence de température positive si le son « on » est émis par le locuteur.

Lors de la prononciation du son « i », de même, il est possible de déterminer sa prononciation par étude de l'air expulsé par la bouche lors de l'émission du son.

En variante, le microphone 104 du dispositif 16 est utilisé pour enregistrer le son émis par le locuteur. Cela permet de corréler les interprétations des morphèmes et des phonèmes du locuteur avec le son enregistré par le microphone.

Selon un mode de réalisation, cette corrélation est faite avec les images de la bouche et/ou les mesures de température différentielle du capteur 102 de souffle différentiel.

Dans tous les cas, le module 100 de reconnaissance vocale selon l'invention permet d'obtenir une meilleure distinction des phonèmes qu'un module 100 de reconnaissance vocale basé seulement sur une lecture labiale comme dans l'état de la technique.

En outre, le module 100 de reconnaissance vocale est utilisable pour un pilote d'avion, même soumis à des fortes accélérations. Par comparaison, il est à noter qu'un

simple dispositif de lecture labiale ne peut être utilisé avec un pilote d'avion du fait de la déformation des lèvres du pilote d'avion par l'accélération et/ou les vibrations de l'avion.

La même remarque s'applique également pour une personne atteinte de paralysie des lèvres ayant encore l'usage de sa langue.

5 Le module 100 de reconnaissance vocale est également utilisable dès qu'il est souhaitable de retranscrire des données issues de sons émis par le locuteur.

 Selon une application, le module 100 de reconnaissance vocale permet de suppléer l'usage de la parole pour une personne qui en serait privée. En ce sens, le module 100 de reconnaissance vocale est une prothèse pour un handicapé privé de
10 l'usage de la parole.

 Pour un champ d'application similaire, le module 100 de reconnaissance vocale est également employable dans le contexte de l'orthophonie pour corriger la mauvaise prononciation par le patient.

 Selon une autre application, le module 100 de reconnaissance vocale retranscrit
15 des instructions d'un locuteur qui ne peut pas parler librement ou dont la locution est gênée par un masque.

 C'est, par exemple, le cas d'un chirurgien dans une salle d'opérations qui a ses deux mains prises pour la manipulation des instruments. La module 100 de reconnaissance vocale lui permet d'étendre directement son champ d'action au pilotage
20 d'instruments connexes, en se passant du concours du personnel l'assistant, par exemple pour la modification du zoom d'une caméra endoscopique ou pour la modulation de l'intensité d'un bistouri électrique.

 Selon une autre application, le module 100 de reconnaissance vocale sert à produire des sons.

25 Dans ce cas, la sous-unité de reconnaissance 110 des phonèmes et morphèmes de l'unité de traitement 17 est remplacée par une sous-unité de reconnaissance de son.

 La sous-unité de reconnaissance de son est propre à déterminer la fréquence et/ou l'intensité d'un son.

 La sous-unité de reconnaissance de son est capable de déterminer la fréquence
30 par l'observation du mouvement relatif des deux mâchoires du locuteur. L'amplitude de ce mouvement est de 4 mm à 20 mm dans la direction verticale et de 2 mm en avant dans la direction horizontale et 6 mm en arrière pour la direction horizontale.

 Un mouvement de 2 mm en avant dans la direction horizontale correspond à un son plus grave tandis qu'un mouvement de 6 mm en arrière pour la direction horizontale
35 correspond à un son plus aigu.

La sous-unité de reconnaissance de son permet ainsi de quantifier la fréquence d'un son sur la base de la position relative des mâchoires l'une par rapport à l'autre.

Le module 100 de reconnaissance vocale est ainsi bien adapté pour produire de la musique.

5 Le traitement combiné de la cinématique de la langue et du souffle permet de générer des sons avec des paramètres de fréquence (hauteur de la note), d'intensité, d'attaque et de tombée de la note.

Pour une application de traduction, le module 100 de reconnaissance vocale est également bien adapté.

10 Dans ce cas, la sous-unité de reconnaissance 110 des phonèmes et morphèmes de l'unité de traitement 17 est propre à reconnaître les morphèmes et phonèmes émis dans une première langue, par exemple le français.

15 L'unité de traitement 17 du module 100 de reconnaissance vocale comporte également une sous-unité de conversion 112 des morphèmes et/ou phonèmes dans une deuxième langue différente de la première langue, par exemple l'anglais.

Dans l'exemple précité du mot « bonjour » pour de la traduction du français vers l'anglais, converti en « hello » par la sous-unité de conversion 112, l'unité 106 de retranscription retranscrit alors le son « hello ».

20 L'emploi du module 100 de reconnaissance vocale selon l'invention permet d'augmenter le pourcentage de mots émis par le locuteur qui sont reconnus, et ce indépendamment de l'environnement puisque la reconnaissance est basée sur les mouvements de la langue et les différences de température entre l'air expiré par le nez et l'air expiré par la bouche.

25 Cela rend le module 100 de reconnaissance vocale particulièrement bien adapté dans les situations de traduction simultanée.

Selon une variante, le module de reconnaissance vocale 100 est un moyen alternatif à la souris ou au clavier pour la saisie d'un microordinateur.

30 Selon encore un autre aspect de l'invention, la demanderesse a constaté que des instructions orales de contrôle d'un appareil électronique sont plus rapidement traitées que par des instructions qui seraient transmises à l'appareil électronique via un clavier. Un traitement plus rapide permet d'obtenir une meilleure précision dans l'usage de l'appareil électronique considéré.

35 Ainsi, il est proposé un appareil électronique 200 comprenant une unité de contrôle 202 et le système 10 de détermination de la position tel que décrit en référence aux figures 1, 2 et 3. Les différents éléments décrits précédemment sont identiques sauf l'unité de traitement 17 de l'appareil électronique 200 qui est différente de l'unité de

traitement 17 du système 10 de détermination de la position dans l'espace de la langue du locuteur.

L'unité de contrôle 202 est propre à contrôler l'ensemble des actionneurs 204 de l'appareil électronique 200. Dans le cadre de cette invention, il est entendu par
5 « actionneur », tout élément ayant une interaction avec un élément physique. Ainsi, un écran d'affichage sur lequel s'affiche des informations est considéré comme un actionneur 204.

L'unité de traitement 17 comporte une sous-unité de détermination des données 210 et une sous-unité de transmission 212.

10 La sous-unité de détermination des données 210 est propre à déterminer des données à communiquer à l'unité de contrôle 202 à partir des données issues de l'unité 42 d'imagerie, des données issues du capteur 102 de souffle différentiel et des données issues du microphone 104. Les données à communiquer sont notées DC dans la suite de la description.

15 La sous-unité de transmission 212 est propre à transmettre les données déterminées par la sous-unité de détermination des données 210 vers l'unité de contrôle.

Le fonctionnement de l'appareil électronique 200 va maintenant être illustré dans le cas où l'appareil électronique 200 est un téléphone mobile muni d'un actuateur qui est un terminal 204 et que l'utilisateur du téléphone mobile souhaite envoyer un SMS.

20 Le service de messagerie SMS, plus connu sous le sigle SMS (acronyme anglais pour « Short Message Service ») ou texto permet de transmettre de courts messages textuels. C'est l'un des services de la téléphonie mobile qui a été introduit par la norme GSM (acronyme anglais pour « Global System for Mobile Communications »).

A titre d'illustration, dans la suite, il est supposé que le locuteur veut envoyer un
25 SMS à Alice. Le contenu du SMS est « Où es-tu ? ».

Dans cet exemple, le locuteur émet les données « envoi SMS Alice ».

Lorsque le locuteur émet les données « envoi SMS Alice », plusieurs données sont mémorisées : au moins une images de langue du locuteur est acquise, le gradient de température entre l'air expiré par le nez du locuteur et l'air expiré par la bouche due
30 locuteur est mesuré et le son émis par le locuteur est enregistré par le microphone 104.

La sous-unité de détermination des données 210 de l'unité de traitement 17 détermine alors les données à communiquer DC à partir des données mémorisées.

Cette étape de détermination est une étape de reconnaissance vocale similaire à ce qui a été décrit précédemment.

De préférence, l'étape de détermination est mise en œuvre avec des ordres de priorité pour diminuer le temps de traitement des données mémorisées par la sous-unité de détermination des données 210.

Un exemple d'ordre de priorité est explicité dans la suite.

5 L'étape de détermination comprend alors une première étape de reconnaissance vocale R1 sur la base des données enregistrées par le microphone 104 seulement. En effet, si le locuteur émet un son (par exemple, cela lui est possible si le locuteur est seul), la reconnaissance vocale par le microphone 104 est souvent la plus rapide.

10 La première étape de reconnaissance vocale R1 permet d'obtenir des premières données de reconnaissance D1 et un premier taux T1 de probabilité d'erreur entre les premières données de reconnaissance D1 et les données émises par le locuteur.

Le premier taux T1 est représentatif de la probabilité que les premières données de reconnaissance D1 soient différentes des données émises par le locuteur.

15 Le premier taux T1 de probabilité d'erreur est, par exemple, estimé sur la base du rapport signal sur bruit des données enregistrées par le microphone 104. Plus le rapport signal sur bruit est faible, plus le premier taux T1 de probabilité d'erreur est élevé.

Le premier taux T1 de probabilité d'erreur est alors comparé à une première valeur seuil V1.

20 La première valeur seuil V1 est choisie pour que l'écart entre les premières données de reconnaissance D1 et les données émises par le locuteur soit suffisamment faible pour que l'unité de contrôle 202 donne à l'appareil électronique 200 les ordres correspondant aux données émises par le locuteur si les premières données de reconnaissance D1 sont transmises à l'unité de contrôle 202.

25 Par exemple la première valeur seuil V1 est égale à 10%, de préférence 5% et de manière préférée égale à 2,5%.

Lorsque le premier taux T1 de probabilité d'erreur est inférieur ou égal à la première valeur seuil V1, les données à communiquer DC sont les premières données de reconnaissance D1.

30 Dans ce cas, il a été évité d'étudier les données d'images et de gradient de température pour obtenir les données à communiquer, ce qui permet d'obtenir un gain de temps.

35 Lorsque le premier taux T1 de probabilité d'erreur est supérieur à la première valeur seuil V1, l'étape de détermination comprend une deuxième étape de reconnaissance vocale R2 sur la base des données d'image acquises et des données sonores enregistrées.

Selon une variante, ce sont les données de différence de température et les données sonores enregistrées qui sont prises en compte lors de la deuxième étape de reconnaissance vocale R2.

5 La deuxième étape de reconnaissance vocale R2 permet d'obtenir des deuxièmes données de reconnaissance D2 et un deuxième taux T2 de probabilité d'erreur entre les deuxièmes données de reconnaissance D2 et les données émises par le locuteur.

Le deuxième taux T2 est représentatif de la probabilité que les deuxièmes données de reconnaissance D2 soient différentes des données émises par le locuteur.

10 Le deuxième taux T2 de probabilité d'erreur est par exemple estimé sur la base du rapport signal sur bruit des données d'images acquises. Plus le rapport signal sur bruit est faible, plus le deuxième taux T2 de probabilité d'erreur est élevé.

Le deuxième taux T2 de probabilité d'erreur est alors comparé à une deuxième valeur seuil V2.

15 La deuxième valeur seuil V2 est choisie pour que l'écart entre les deuxièmes données de reconnaissance D2 et les données émises par le locuteur soit suffisamment faible pour que l'unité de contrôle 202 donne à l'appareil électronique 200 les ordres correspondant aux données émises par le locuteur si les deuxièmes données de reconnaissance D2 sont transmises à l'unité de contrôle 202.

20 Par exemple, la deuxième valeur seuil V2 est égale à 10%, de préférence 5% et de manière préférée égale à 2,5%.

Selon une variante, la deuxième valeur seuil V2 et la première valeur seuil V1 sont égales.

Lorsque le deuxième taux T2 est inférieur ou égal à la deuxième valeur seuil V2, les données à communiquer DC sont les deuxièmes données de reconnaissance D2.

25 Dans ce cas, il a été évité d'étudier les données de gradient de température pour obtenir les données à communiquer DC, ce qui permet d'obtenir un gain de temps.

30 Lorsque le deuxième taux T2 est supérieur à la deuxième valeur seuil V2, l'étape de détermination comprend une troisième étape de reconnaissance vocale R3 sur la base de toutes les données du dispositif 12 d'acquisition, à savoir les données acquises, mesurées et enregistrées.

La troisième étape de reconnaissance vocale R3 permet d'obtenir des troisièmes données de reconnaissance D3, les données à communiquer DC étant les troisièmes données de reconnaissance D3.

35 Les données à communiquer DC sont ensuite transmises à l'unité de contrôle 202 par la sous-unité de transmission 210.

L'unité de contrôle 202 reçoit ainsi les ordres émanant du locuteur et, en conséquence, le terminal affiche alors une page vide correspondant au contenu du SMS.

Le locuteur émet ensuite les données « Où es-tu ? ».

Par le même procédé que précédemment décrit pour « envoi SMS Alice », le
5 terminal affiche alors une page dans laquelle est écrite « Où es-tu ? ».

Le locuteur émet ensuite les données « Envoyer ».

Par le même procédé que précédemment décrit pour « envoi SMS Alice », l'ordre d'envoyer du locuteur est transmis à l'unité de contrôle 202.

L'unité de contrôle reçoit cet ordre émanant du locuteur et, en conséquence, le
10 terminal affiche alors une demande de confirmation de l'ordre « Voulez-vous envoyer ce message ? ».

Le locuteur émet ensuite les données « Oui ».

Par le même procédé que précédemment décrit pour « envoi SMS Alice », la confirmation orale du locuteur à l'unité de contrôle 202.

15 Le téléphone mobile envoie alors le message à Alice.

Pour comparaison, dans le cas d'un téléphone mobile comprenant un terminal muni d'un clavier, le locuteur clique d'abord sur Alice, puis envoyer SMS. Le terminal affiche alors une page vide correspondant au contenu du SMS.

En supposant que le SMS est saisi sur un clavier de type azerty, le locuteur appuie
20 successivement sur les touches « o », « ù », espace, « e », « s », « - », « t », « u » et « ? ». Le terminal affiche alors une page dans laquelle est écrite « Où es-tu ? ».

Le locuteur clique alors sur l'option envoyer sur le terminal. Le terminal affiche alors une demande de confirmation de l'ordre « Voulez-vous envoyer ce message ? Oui / Non ». Le locuteur clique sur la réponse « Oui ».

25 Le terminal envoie alors le message à Alice

Ainsi, avec le terminal muni du dispositif selon l'invention, le locuteur a prononcé quatre phrases alors qu'avec le terminal muni d'un clavier, le locuteur a treize interactions distinctes avec le clavier (quatre clics et neuf appuis sur une touche alpha-numérique).

Il apparaît alors que dicter en utilisant un téléphone mobile 200 muni du dispositif
30 12 d'acquisition est plus rapide que d'utiliser un téléphone mobile muni d'un clavier.

En outre, les mouvements de la langue sont plus aisément répétables que les manipulations d'un clavier par les doigts, notamment pour éviter les erreurs de saisie. Il suffit, en effet, de prononcer correctement le SMS pour que celui-ci apparaisse sur l'écran.

35 De plus, du fait que la dictée peut se faire sans production de son, la dictée ne produit pas de fatigue lors d'un exercice prolongé au contraire de la communication à voix

haute. La communication à voix haute sollicite les cordes vocales et peut fatiguer le locuteur.

En variante, le téléphone mobile est muni du dispositif 12 d'acquisition selon l'invention et d'un clavier. L'envoi du SMS est alors effectué de manière mixte. Par exemple, seul le contenu textuel du SMS est dicté, le reste des opérations étant effectué par utilisation du clavier.

Le procédé proposé s'applique également pour toute situation où un locuteur commande un appareil électronique 200.

Ainsi, selon un autre mode de réalisation, l'appareil électronique 200 est une console de jeux vidéo. Dans ce cas, au lieu de taper la combinaison de touches requises par la situation de jeu comme dans l'état de la technique, l'utilisateur de la console de jeux vidéo se contente de donner un ordre comme « course », « saut », « arrêt » ou « tir ».

Cela permet d'augmenter la cadence des jeux vidéo puisqu'une manipulation par la langue est plus rapide qu'une manipulation par la main.

En outre, alors que le nombre limitée de touches pour une console selon l'état de la technique impose des combinaisons complexes pour pouvoir multiplier les possibilités d'actions, la console proposée permet potentiellement d'obtenir plus aisément de multiples possibilités d'action puisque le nombre de mots à disposition du locuteur est beaucoup plus grand que le nombre de touches des consoles selon l'état de la technique.

Selon une variante, l'appareil électronique 200 est une interface homme-machine.

Par exemple, l'interface homme-machine fait partie d'un casque pour le domaine militaire. Là encore, le contrôle par les mots prononcés par le locuteur est plus rapide qu'une manipulation de commande avec des mains.

Cela permet de réduire les temps de réaction, ce qui est crucial, notamment dans les situations de tir.

Selon des variantes, l'appareil électronique 200 est un robot, un ordinateur ou un fauteuil pour handicapé.

Selon une autre application, l'appareil 200 permet de suppléer l'usage des mains pour un handicap. En ce sens, l'appareil 200 est une prothèse pour un handicapé privé de l'usage de ses mains.

En variante, au lieu d'une mise en œuvre de l'étape de détermination avec des ordres de priorité ainsi que précédemment décrit, il est fait appel à une logique floue pour comparer des profils de sons appris avec des profils incidents pour établir des choix probabilistes.

Parallèlement, les signatures des souffles correspondants à des sons appris sont comparées aux souffles incidents mesurés par le capteur 102 de souffle différentiel.

Cela permet une reconnaissance instantanée, c'est-à-dire au rythme de la prononciation de la phrase. Cela impose un traitement informatique dont le temps de mise en œuvre est suffisamment rapide pour respecter cette contrainte de temps.

Selon encore une autre variante, des processeurs neuronaux sont utilisés.

5 La langue est le plus petit dénominateur fonctionnel commun de communication entre les êtres humains conscients: handicapés : sourds, muets, aveugles, paraplégiques et valides.

10 L'optimisation de la reconnaissance des signes induit la création d'un nouveau langage basé sur la recherche de l'optimisation des distances morphologiques entre morphèmes, eu égard aux diversités des langues parlées, la recherche de la transmission d'ordres combinés et simultanés aux acteurs.

REVENDICATIONS

1.- Appareil électronique (200) comprenant :

- au moins un actionneur (204),
- une unité de contrôle (202) contrôlant l'au moins un actionneur (204),
- un dispositif (12) d'acquisition de données d'un locuteur comprenant :
 - une unité (40) d'illumination d'au moins une partie de la langue du locuteur et d'au moins une partie supérieure d'une dent inférieure, et
 - une unité (42) d'imagerie adaptée à la prise d'image de la partie de langue du locuteur et de la partie supérieure de la dent inférieure illuminées par l'unité (40) d'illumination, et
 - une unité de traitement (17) configurée pour déterminer des données à communiquer (DC) à partir des données issues de l'unité (42) d'imagerie et transmettre les données à communiquer (DC) à l'unité de contrôle (202),

l'appareil électronique (200) étant caractérisé en ce que :

- le dispositif (12) d'acquisition de données comprend, en outre, un capteur (102) de souffle différentiel adapté à mesurer le gradient de température et/ou de pression entre l'air expiré par le nez du locuteur et l'air expiré par la bouche du locuteur, et
- l'unité de traitement (17) est propre à déterminer des données à communiquer (DC) également à partir des données issues du capteur (102) de pression différentielle.

2.- Appareil selon la revendication 1, dans lequel le dispositif (12) d'acquisition de données comprend, en outre, un organe (44) de maintien sur au moins une dent du locuteur.

3.- Appareil selon la revendication 1 ou 2, dans lequel le dispositif (12) d'acquisition de données comprend, en outre, un microphone (104).

4.- Appareil selon l'une quelconque des revendications 1 à 3, dans lequel l'unité de traitement (17) est propre à déterminer des données à communiquer également à partir des données issues du microphone (104).

5.-Appareil selon l'une quelconque des revendications 1 à 4, dans lequel l'appareil électronique (200) est choisi dans le groupe constitué de :

- une console de jeux vidéo,
- un téléphone mobile,
- une interface homme-machine,
- un ordinateur,
- un robot, et
- un fauteuil pour handicapé.

6.- Procédé de communication entre un locuteur et un appareil électronique (200) selon l'une quelconque des revendications 1 à 5, le procédé étant caractérisé en ce que le procédé comprend les étapes de :

- émission de données par le locuteur
- 5 - acquisition d'au moins une image d'une partie de la langue du locuteur simultanément à l'étape d'émission,
- mesure du gradient de température et/ou de pression entre l'air expiré par le nez du locuteur et l'air expiré par la bouche du locuteur simultanément à l'étape d'émission,
- détermination de données à communiquer (DC) à partir des données acquises et
- 10 mesurées,
- commande de l'appareil électronique (200) à partir des données déterminées (DC).

7.- Procédé selon la revendication 6, dans lequel l'étape de détermination est une étape de reconnaissance vocale comprenant :

- 15 - une première étape de reconnaissance vocale sur la base des données acquises seulement, la première étape de reconnaissance vocale permettant d'obtenir des premières données de reconnaissance et un premier taux de probabilité d'erreur entre les premières données et les données émises par le locuteur,
- une étape de comparaison du premier taux de probabilité d'erreur à une valeur
- 20 seuil,
- lorsque le premier taux de probabilité est inférieur ou égal à la valeur seuil, les données à communiquer sont les premières données à communiquer,
- lorsque le premier taux de probabilité est supérieur à une valeur seuil, une deuxième étape de reconnaissance vocale sur la base des données acquises et
- 25 mesurées seulement, la deuxième étape de reconnaissance vocale permettant d'obtenir des deuxièmes données, les données à communiquer étant les deuxièmes données.

8.- Procédé selon la revendication 6, dans lequel le procédé comprend également une étape de :

- enregistrement de données de son par le microphone (104),
- 30 l'étape de détermination de données à communiquer prenant également en compte les données enregistrées.

9.- Procédé selon la revendication 8, dans lequel l'étape de détermination comprend :

- 35 - une première étape de reconnaissance vocale sur la base des données enregistrées seulement, la première étape de reconnaissance vocale permettant d'obtenir des premières données de reconnaissance (D1) et un premier taux (T1) de probabilité

d'erreur entre les premières données de reconnaissance (D1) et les données émises par le locuteur,

- une étape de comparaison du premier taux (T1) de probabilité d'erreur à une première valeur seuil (V1) telle que lorsque le premier taux (T1) de probabilité d'erreur est inférieur ou égal à la première valeur seuil (V1), les données à communiquer (DC) sont les premières données de reconnaissance (D1),

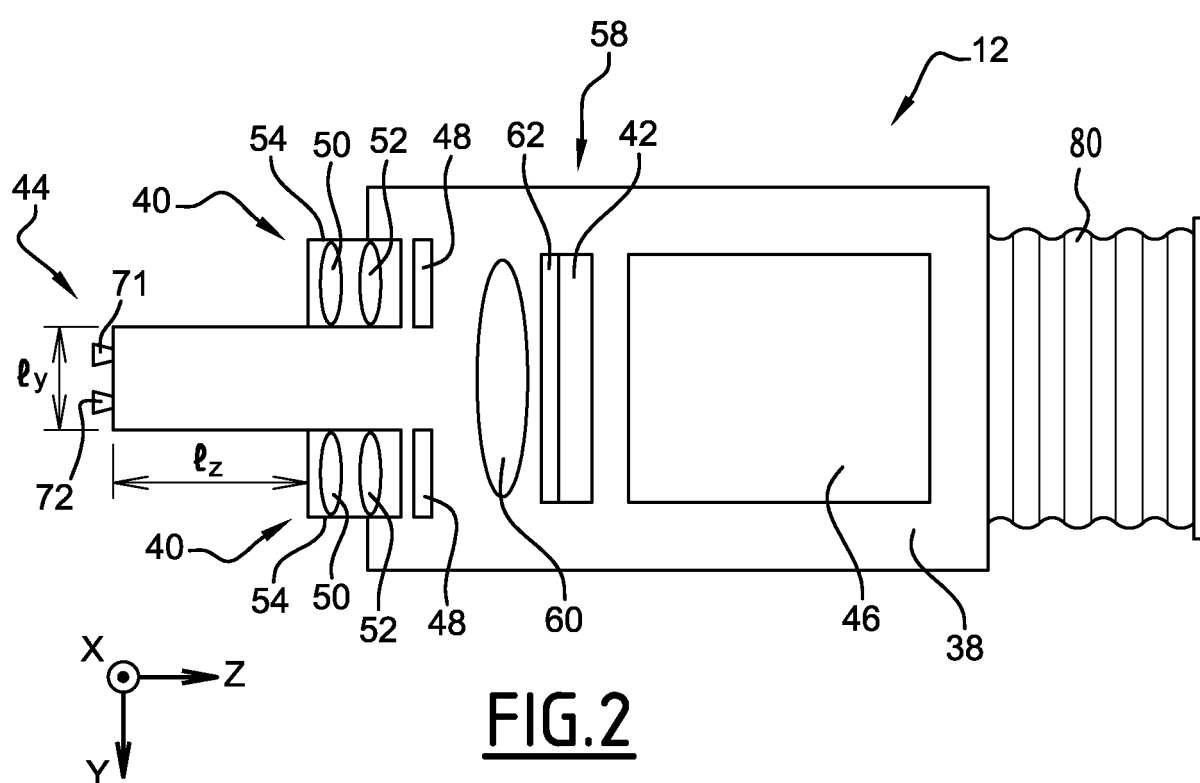
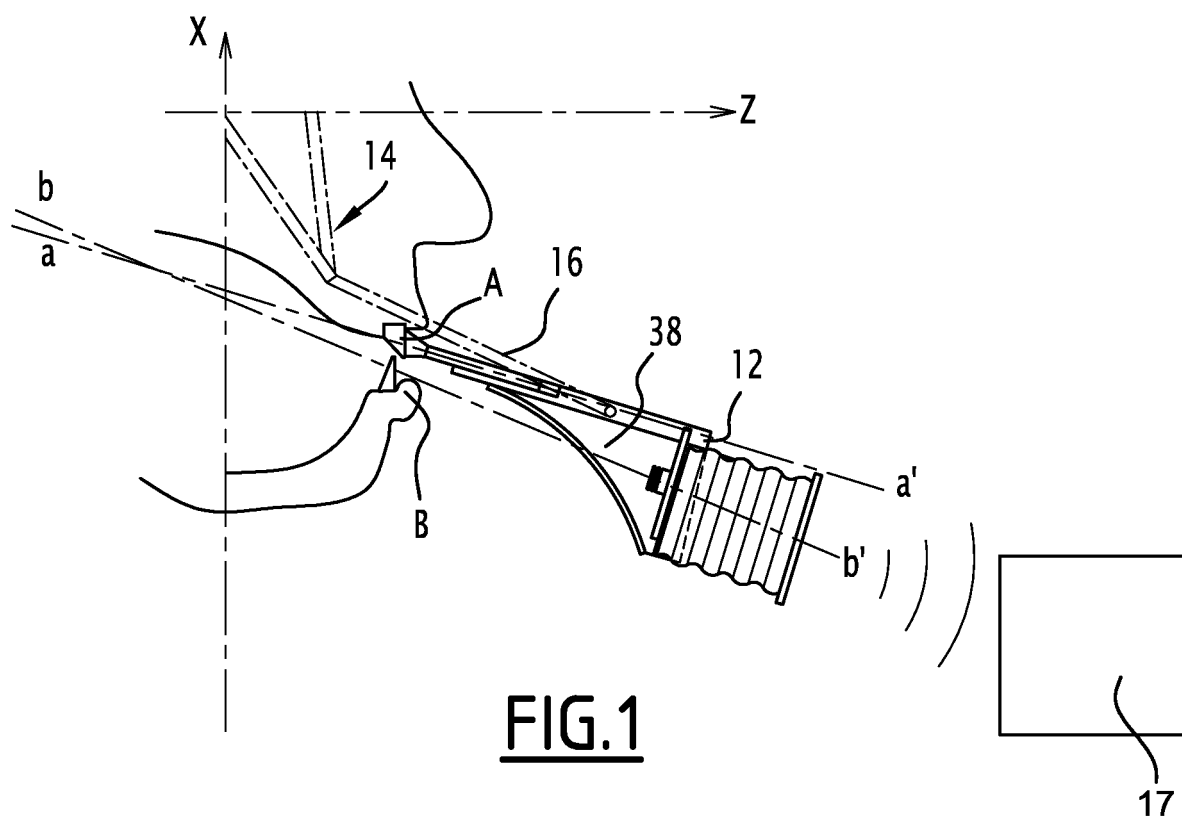
- une deuxième étape de reconnaissance vocale sur la base des données acquises et enregistrées lorsque le premier taux (T1) de probabilité d'erreur est supérieur à la première valeur seuil (V1), la deuxième étape de reconnaissance vocale permettant d'obtenir des deuxième données de reconnaissance (D2) et un deuxième taux (T2) de probabilité d'erreur entre les deuxième données de reconnaissance (D2) et les données émises par le locuteur,

- une étape de comparaison du deuxième taux (T2) de probabilité d'erreur à une deuxième valeur seuil (V2) telle que lorsque le deuxième taux (T2) de probabilité est inférieur ou égal à la deuxième valeur seuil (V2), les données à communiquer (DC) sont les deuxième données de reconnaissance (D2),

- une troisième étape de reconnaissance vocale sur la base des données acquises, mesurées et enregistrées lorsque le deuxième taux (T2) de probabilité est supérieur à une deuxième valeur seuil (V2), la troisième étape de reconnaissance vocale permettant d'obtenir des troisièmes données de reconnaissance (D3), les données à communiquer étant les troisièmes données de reconnaissance (D3).

10.- Procédé selon la revendication 9, dans lequel le premier taux (T1) de probabilité d'erreur et le deuxième taux (T2) de probabilité d'erreur sont égaux.

1/3



2/3

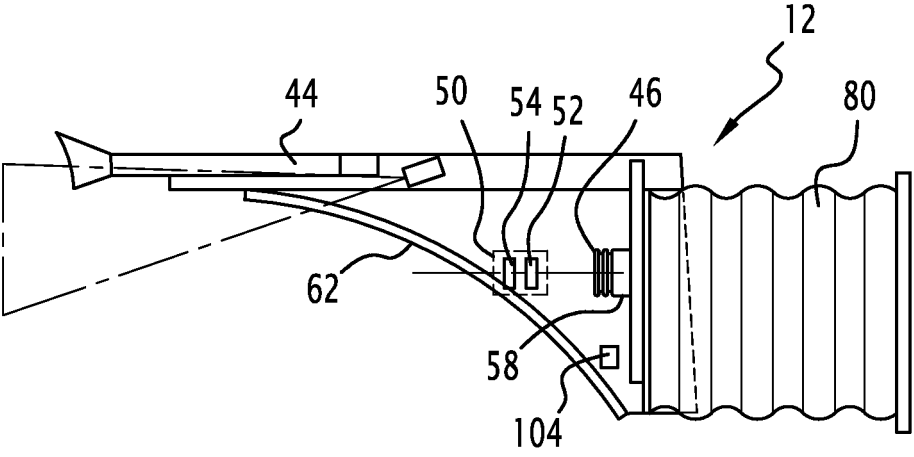


FIG. 3

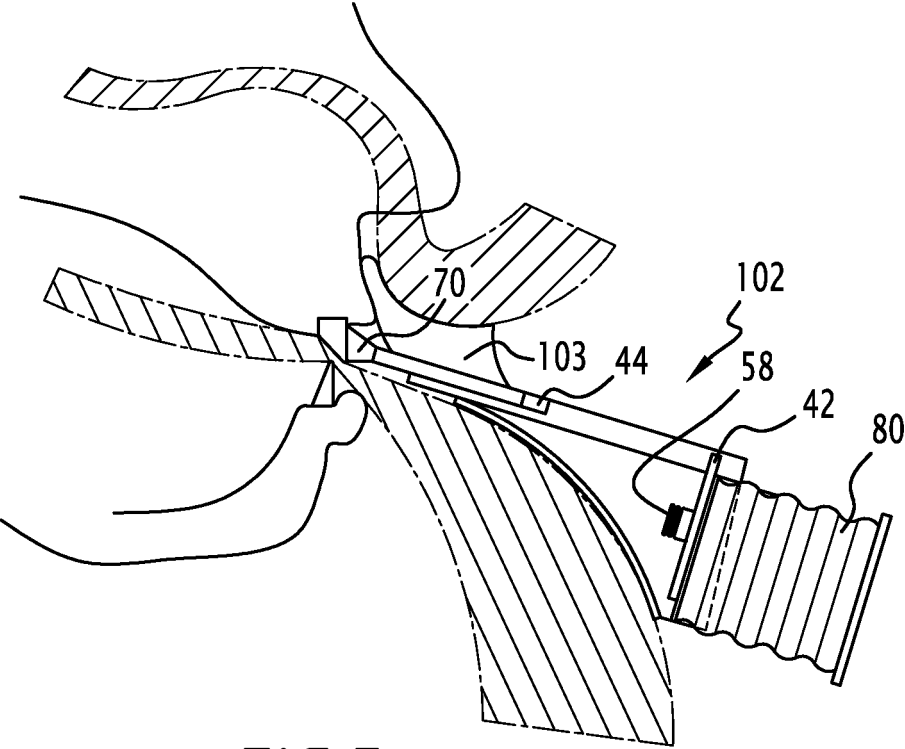
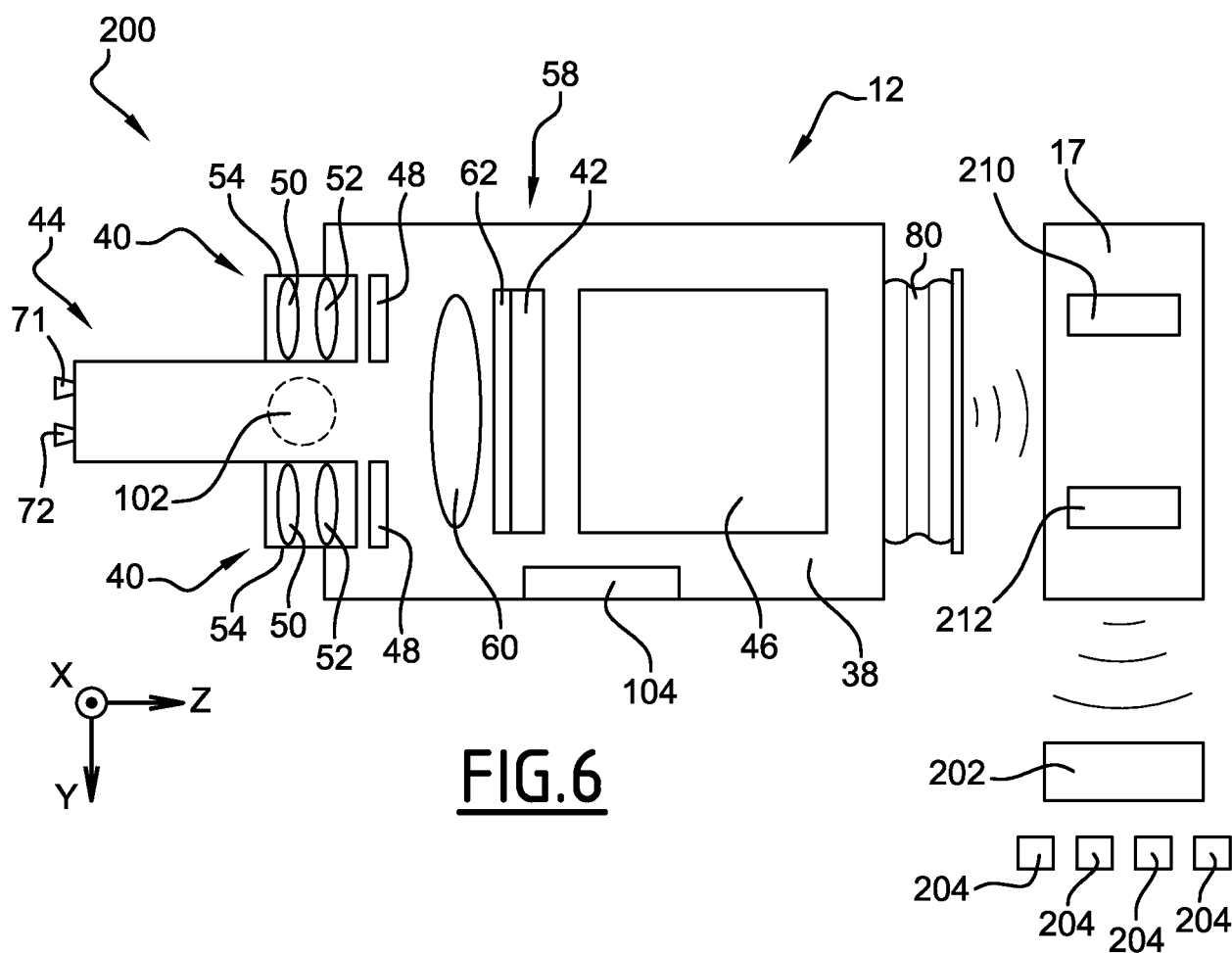
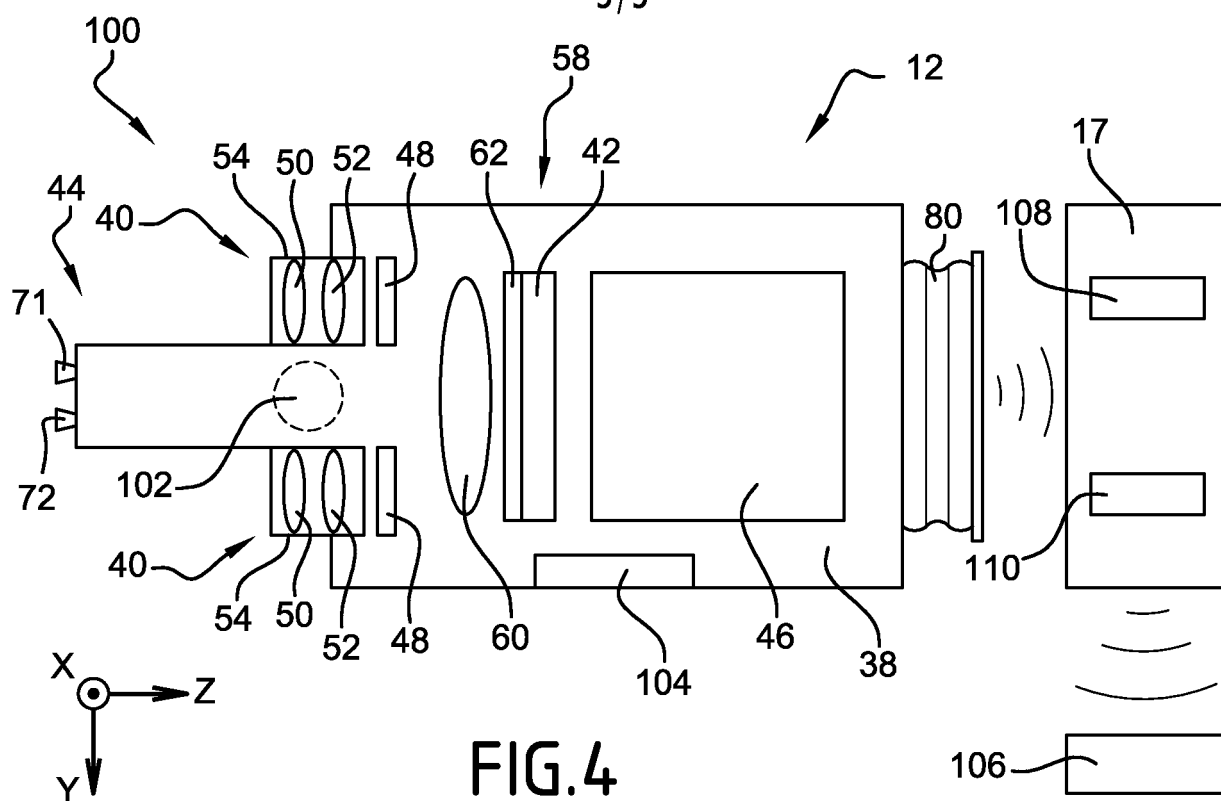


FIG. 5

3/3





RAPPORT DE RECHERCHE PRÉLIMINAIRE

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

N° d'enregistrement
national

FA 777869
FR 1262860

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,D	US 2010/036657 A1 (MORISAKI MITSUNORI [JP] ET AL) 11 février 2010 (2010-02-11) * figures 1,18,29 * * alinéa [0140] - alinéa [0151] * * alinéa [0070] * * alinéa [0214] - alinéa [0227] * -----	1-10	G10L15/24 H04N7/18 G01K13/00 G06T7/00 G01L23/00 G01K13/02
A	US 2002/194005 A1 (LAHR ROY J [US]) 19 décembre 2002 (2002-12-19) * alinéa [0091] - alinéa [0096] * * figures 6,7 * * alinéa [0023] - alinéa [0027] * -----	1-6	DOMAINES TECHNIQUES RECHERCHÉS (IPC) A61B G10L
A	US 2002/087103 A1 (FLETCHER SAMUEL G [US] ET AL) 4 juillet 2002 (2002-07-04) * alinéa [0055] - alinéa [0058]; figure 1 *	1,6	
A	US 2012/259554 A1 (CHEN RUXIN [US] ET AL) 11 octobre 2012 (2012-10-11) * figures 3A-3B,6 * * alinéas [0022], [0023], [0042] * -----	2-5	
A	US 2004/243416 A1 (GARDOS THOMAS R [US]) 2 décembre 2004 (2004-12-02) * figures 1-3 * * abrégé * * alinéa [0024] - alinéa [0030] * -----	1,3,6	
A	US 2009/305203 A1 (OKUMURA MACHI [JP] ET AL) 10 décembre 2009 (2009-12-10) * revendications 1,17 * -----	1,6	
Date d'achèvement de la recherche		Examineur	
28 août 2013		De la Hera, Germán	
CATÉGORIE DES DOCUMENTS CITÉS		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	
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			

ANNEXE AU RAPPORT DE RECHERCHE PRÉLIMINAIRE
RELATIF A LA DEMANDE DE BREVET FRANÇAIS NO. FR 1262860 FA 777869

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 **28-08-2013**

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 2010036657 A1	11-02-2010	US 2010036657 A1 WO 2008062782 A1	11-02-2010 29-05-2008

US 2002194005 A1	19-12-2002	US 2002194005 A1 WO 02077972 A1	19-12-2002 03-10-2002

US 2002087103 A1	04-07-2002	AU 775160 B2 AU 9286001 A CA 2422811 A1 CN 1635851 A EP 1320850 A2 JP 2004509677 A US 2002087103 A1 WO 0225635 A2	22-07-2004 02-04-2002 28-03-2002 06-07-2005 25-06-2003 02-04-2004 04-07-2002 28-03-2002

US 2012259554 A1	11-10-2012	US 2012259554 A1 WO 2012138450 A1	11-10-2012 11-10-2012

US 2004243416 A1	02-12-2004	AUCUN	

US 2009305203 A1	10-12-2009	EP 1947643 A1 JP 5120826 B2 JP 2007122004 A KR 20080059180 A US 2009305203 A1 WO 2007037356 A1	23-07-2008 16-01-2013 17-05-2007 26-06-2008 10-12-2009 05-04-2007
