

12

DEMANDE DE BREVET D'INVENTION

A1

②② **Date de dépôt : 26.08.94.**

③③ **Priorité** : 26.08.93 JP 21126993.

43 Date de la mise à disposition du public de la demande : 03.03.95 Bulletin 95/09.

(56) Liste des documents cités dans le rapport de recherche préliminaire : Ce dernier n'a pas été établi à la date de publication de la demande.

⑥ Références à d'autres documents nationaux apparentés :

(71) Demandeur(s) : *NEC Corporation* — JP.

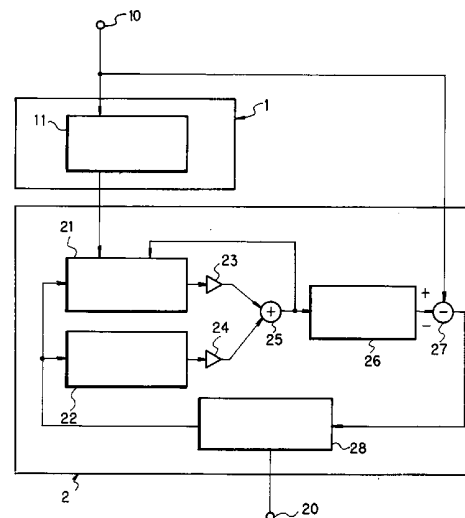
(72) Inventeur(s) : Masahiro Serizawa.

(73) Titulaire(s) :

(74) Mandataire : Société de Protection des Inventions.

(54) Système de codage de hauteur de son de parole.

(57) Une pluralité de voies de transition de périodes de hauteur de son sont extraites au moyen d'un suivi de hauteur de son (11) sur une séquence et une voie d'un gain de prédiction moyen minimum sur la séquence est choisie parmi les voies extraites. Une sélection de hauteur de son préliminaire suivante peut être exécutée lors d'un traitement de sous-séquence afin de sélectionner une pluralité de candidats au voisinage de la hauteur de son de la voie de transition sélectionnée pour chaque sous-séquence en utilisant le produit interne du signal de parole d'entrée et de chaque vecteur de code (21, 22). Pour finir, une période de hauteur de son présentant une distorsion de forme d'onde minimum est choisie (28) pour chaque sous-séquence.



SYSTEME DE CODAGE DE HAUTEUR DE SON DE PAROLE

La présente invention concerne un système de codage de hauteur de son de parole permettant de réaliser un codage haute qualité d'un signal de parole selon un débit binaire faible, plus particulièrement de 4 kilobits/seconde ou moins.

- 5 Un système de codage de parole de l'art antérieur code un signal de parole sur la base de données de paramètres caractéristiques obtenues pour chaque séquence (laquelle présente une longueur de 40 millisecondes par exemple) du signal de parole et de données de paramètres caractéristiques
- 10 obtenues pour chacune des sous-séquence (laquelle présente une longueur de 8 millisecondes par exemple) qui constituent des divisions supplémentaires de la séquence. Le système comprend deux sources d'excitation, c'est-à-dire un livre de codes adaptatifs produit en répétant un signal d'excitation précédent
- 15 selon une période de hauteur de son et un livre de codes de source d'excitation constitué par un signal produit précédemment et il produit un signal d'excitation synthétisé en faisant passer le signal d'excitation au travers d'un filtre de synthèse de prédiction linéaire. Le filtre de synthèse est
- 20 construit en utilisant un jeu de coefficients de filtre (par exemple un jeu de coefficients de filtre de prédiction linéaire) obtenu par l'intermédiaire d'une analyse d'une parole d'entrée de séquence courante qui doit être quantifiée. En tant que système de codage de ce type, un système de codage par prédiction
- 25 linéaire excité par code ou système CELP est bien connu, lequel est décrit par exemple dans le traité de M. Schroeder et B. Atal intitulé "Code-Excited Linear Prediction : High Quality Speech at Very Low Bit Rates", IEEE Proc., ICASSP-85, pages 937-940, 1985.
- 30 Selon un autre système de l'art antérieur, le codage de hauteur de son est réalisé selon une faible quantité d'opérations au moyen d'une sélection préliminaire de hauteur de son. En tant

que systèmes de ce type, on connaît un système de recherche à deux étages (décrit dans la demande de brevet du Japon n° Heisei 4-305135) qui comprend des étapes de sélection préliminaire de hauteur de son dans une boucle ouverte en utilisant des coefficients d'auto-corrélation d'un signal résiduel et de sélection finale de hauteur de son à partir de candidats sélectionnés en utilisant une distorsion en boucle fermée, un système de recherche à deux étages (décrit dans la demande de brevet du Japon n° Heisei 4-270398) qui comprend des étapes de sélection préliminaire de hauteur de son dans une boucle ouverte en utilisant des coefficients d'auto-corrélation d'un signal d'entrée et de sélection de hauteur de son finale à partir de retards proches des candidats sélectionnés en utilisant une distorsion en boucle fermée et un système de recherche à trois étages (décrit dans TECHNICAL REPORT OF IEICE. SP92-133, 1993-02, Para. 5.1.2) qui comprend des étapes de sélection de hauteur de son préliminaire dans une boucle ouverte en utilisant des coefficients d'auto-corrélation d'un signal résiduel, de sélection préliminaire de hauteur de son suivante dans une boucle fermée moyennant seulement le produit interne d'un signal d'entrée et de chaque vecteur de code et de sélection finale de hauteur de son à partir de candidats sélectionnés en utilisant une distorsion en boucle fermée.

Dans les systèmes de l'art antérieur mentionnés ci-avant cependant, la sélection préliminaire de hauteur de son est réalisée lors de chaque traitement de sous-trame. Par conséquent, si le nombre de candidats lors de la sélection finale de hauteur de son est excessivement réduit, une hauteur de son présentant une distorsion de forme d'onde localement faible peut être sélectionnée, ce qui augmente la détérioration de la qualité de parole de la parole codée. Afin d'éviter ce problème, un certain nombre de candidats sont nécessaires, ce qui rend difficile de réduire la quantité d'opérations mises en jeu.

Un objet de la présente invention consiste par conséquent à proposer un système de codage de hauteur de son de parole

permettant un codage de hauteur de son moyennant une faible quantité d'opérations par comparaison avec l'art antérieur.

Selon un aspect de la présente invention, on propose un système de codage de hauteur de son de parole permettant de
5 coder un signal de parole en utilisant des paramètres caractéristiques obtenus pour chaque séquence du signal de parole et des paramètres caractéristiques obtenus pour chacune des sous-séquences qui constituent des divisions supplémentaires de la séquence et de synthétiser un signal de
10 parole au moyen d'un filtre de synthèse de prédiction linéaire, dans lequel des signaux de source d'excitation d'un livre de codes adaptatifs obtenu en répétant un signal d'excitation précédent selon une période de hauteur de son et d'un livre de codes d'excitation constitué par un signal produit préliminairement
15 sont appliqués, comprenant : un moyen de suivi de hauteur de son pour extraire une période de hauteur de son pour chaque unité plus longue que la sous-séquence ; et un moyen de sélection finale de période de hauteur de son pour sélectionner en final une période de hauteur de son présentant une distorsion de forme
20 d'onde minimum, obtenue par l'intermédiaire dudit filtre de synthèse de prédiction linéaire, pour chacune des sous-séquences, parmi des périodes de hauteur de son au voisinage de la période de hauteur de son extraite dans ledit moyen de suivi de hauteur de son.

25 Selon un autre aspect de la présente invention, on propose un système de codage de hauteur de son de parole permettant de coder un signal de parole en utilisant des paramètres caractéristiques obtenus pour chaque séquence du signal de parole et des paramètres caractéristiques obtenus pour chacune
30 des sous-séquences qui constituent des divisions supplémentaires de la séquence et permettant de synthétiser un signal de parole au moyen d'un filtre de synthèse de prédiction linéaire dans lequel des signaux de source d'excitation d'un livre de codes adaptatifs obtenu en répétant un signal d'excitation
35 précédent selon une période de hauteur de son et d'un livre de

codes d'excitation constitué par un signal produit
préliminairement sont appliqués, comprenant : un moyen de suivi
de hauteur de son pour extraire une période de hauteur de son
pour chaque unité plus longue que la sous-séquence ; un moyen de
5 sélection préliminaire de période de hauteur de son pour
extraire, pour chacune des sous-séquences, des candidats de
période de hauteur de son par rapport à une période de hauteur de
son au voisinage de la période de hauteur de son extraite dans
ledit moyen de section de suivi de hauteur de son ; et un moyen
10 de sélection finale de période de hauteur de son pour
sélectionner une période de hauteur de son présentant une
distorsion de forme d'onde minimum parmi les candidats de
période de hauteur de son extraits dans ledit moyen de sélection
de période préliminaire de hauteur de son par l'intermédiaire
15 dudit filtre de synthèse de prédiction linéaire.

La présente invention utilise le fait que la période de
hauteur de son d'un signal de parole n'est pas modifiée de façon
soudaine. Une pluralité de voies de transition de période de
hauteur de son sont extraites au moyen d'un suivi de hauteur de
20 son sur une séquence et une voie d'un gain de prédiction moyen
minimum sur la séquence est choisie parmi les voies extraites.
Selon un autre aspect selon lequel une sélection de hauteur de
son préliminaire suivante est exécutée lors d'un traitement de
sous-séquence, une pluralité de candidats sont sélectionnés au
25 voisinage de la hauteur de son de la voie de transition
sélectionnée pour chaque sous-séquence en utilisant le produit
interne du signal de parole d'entrée et de chaque vecteur de code.
Pour finir, une période de hauteur de son présentant une
distorsion de forme d'onde minimum est choisie pour chaque
30 sous-séquence. De la manière présentée ci-avant, les candidats
de hauteur de son sont réduits à un unique candidat lors du suivi
de hauteur de son afin de réduire notablement la quantité
d'opérations.

En outre, puisque le suivi de hauteur de son est réalisé, il
35 est possible d'obtenir une réduction des bits de transmission de

période de hauteur de son en exprimant la période de hauteur de son à l'aide de la différence entre la période de hauteur de son pour la sous-séquence et la période de hauteur de son pour la sous-séquence précédente.

5 Comme représenté, moyennant le système de codage de hauteur de son de parole selon la présente invention, il est possible d'obtenir un codage de hauteur de son haute qualité moyennant une très faible quantité d'opérations nécessaires par comparaison avec le système de l'art antérieur de telle sorte que
10 la sélection d'une hauteur de son minimum d'une distorsion de forme d'onde locale est empêchée. Il est également possible d'obtenir un codage de hauteur de son moyennant une quantité davantage faible de bits de transmission.

D'autres objets et caractéristiques de la présente
15 invention apparaîtront à la lumière de la description détaillée qui suit que l'on lira en relation avec les dessins annexés parmi lesquels :

la figure 1 est un schéma fonctionnel qui représente un premier mode de réalisation de la présente invention ; et

20 la figure 2 est un schéma fonctionnel qui représente un second mode de réalisation de la présente invention.

Des modes de réalisation de la présente invention sont maintenant décrits par report aux dessins.

La figure 1 est un schéma fonctionnel qui représente un
25 premier mode de réalisation de la présente invention.

Un signal de parole entré sur une borne d'entrée 10 est appliqué à une section de suivi de hauteur de son 11 contenue dans un processeur de séquence 1 en vue du suivi de hauteur de son dans chaque séquence et la voie de suivi de hauteur de son
30 résultante est appliquée à un processeur de sous-séquence 2. Selon un procédé de suivi de hauteur de son, moyennant une séquence prédéterminée (qui présente une longueur de 40 millisecondes par exemple) et moyennant des sous-séquences (qui présentent une longueur de 8 millisecondes par exemple) en
35 tant que divisions de la séquence, une voie de suivi de hauteur de

son présentant une distorsion de forme d'onde minimum ou un gain de prédiction de hauteur de son moyen maximum est choisie parmi B^N combinaisons de voies de suivi de hauteur de son où B est le nombre de bits du codage de hauteur de son dans chaque sous-séquence et N est le nombre de sous-séquences dans la séquence. Puisque ce procédé nécessite une quantité énorme d'opérations, la quantité d'opérations peut être extrêmement réduite en adoptant un procédé dans lequel la passe est déterminée en sélectionnant successivement des hauteurs de son à partir de l'une quelconque des sous-séquences.

Puis dans le processeur de sous-séquence 2 une section de livre de codes adaptatifs 21 produit des candidats de hauteur de son (par exemple environ cinq candidats de hauteur de son repérés à l'aide d'index) au voisinage de la hauteur de son correspondant à chaque sous-séquence de la voie de suivi de hauteur de son obtenue dans le processeur de séquence 1. Puis une section d'évaluation de distorsion minimum 28 sélectionne en tant que distorsion de forme d'onde minimum l'une des combinaisons des vecteurs correspondant aux candidats de hauteur de son pris parmi des vecteurs de code adaptatif accumulés dans la section de livre de codes adaptatifs 21 et pris parmi des vecteurs de code d'excitation accumulés dans une section de livre de codes d'excitation 22 et applique l'index de la combinaison sélectionnée sur une borne de sortie 20. La distorsion de forme d'onde est calculée en utilisant une différence obtenue depuis un soustracteur 27 qui calcule la différence entre le signal de parole d'entrée et un signal de parole synthétisé obtenu en faisant passer un signal d'excitation obtenu dans un additionneur 25 par l'intermédiaire du réglage d'amplitude et de l'addition de sorties de multiplieurs 23 et 24 qui multiplient les vecteurs de code adaptatif et d'excitation de chaque combinaison au travers d'un filtre de synthèse 26.

La figure 2 est un schéma fonctionnel qui représente un second mode de réalisation de la présente invention.

Ce mode de réalisation est le même que le premier mode de réalisation qui précède à ceci près que le processeur de sous-séquence inclut en outre une section de sélection préliminaire de hauteur de son 29. Une section de sélection préliminaire de
5 hauteur de son 11 exécute en outre la sélection préliminaire de hauteur de son par rapport à chaque sous-séquence au voisinage de la voie de suivi de hauteur de son obtenue dans la section de suivi de hauteur de son 11. Pour la sélection préliminaire de hauteur de son, aucun des procédés de l'art antérieur relevés
10 jusqu'ici n'est efficace.

Comme il a été décrit dans ce qui précède, selon la présente invention, il est possible de réduire la quantité d'opérations lors du codage de hauteur de son par comparaison avec les procédés de l'art antérieur.

REVENDECATIONS

1. Système de codage de hauteur de son de parole permettant de coder un signal de parole en utilisant des paramètres caractéristiques obtenus pour chaque séquence du signal de parole et des paramètres caractéristiques obtenus pour
5 chacune des sous-séquences qui constituent des divisions supplémentaires de la séquence et de synthétiser un signal de parole au moyen d'un filtre de synthèse de prédiction linéaire (26), dans lequel des signaux de source d'excitation d'un livre de codes adaptatifs (21) obtenu en répétant un signal d'excitation
10 précédent selon une période de hauteur de son et d'un livre de codes d'excitation (22) constitué par un signal produit préliminairement sont appliqués,
caractérisé en ce qu'il comprend :
un moyen de suivi de hauteur de son (11) pour extraire une
15 période de hauteur de son pour chaque unité plus longue que la sous-séquence ; et
un moyen de sélection finale de période de hauteur de son pour sélectionner en final une période de hauteur de son présentant une distorsion de forme d'onde minimum, obtenue par
20 l'intermédiaire dudit filtre de synthèse de prédiction linéaire, pour chacune des sous-séquences, parmi des périodes de hauteur de son au voisinage de la période de hauteur de son extraite dans ledit moyen de suivi de hauteur de son (11).
2. Système de codage de hauteur de son de parole
25 permettant de coder un signal de parole en utilisant des paramètres caractéristiques obtenus pour chaque séquence du signal de parole et des paramètres caractéristiques obtenus pour chacune des sous-séquences qui constituent des divisions supplémentaires de la séquence et permettant de synthétiser un
30 signal de parole au moyen d'un filtre de synthèse de prédiction linéaire (26) dans lequel des signaux de source d'excitation d'un livre de codes adaptatifs (21) obtenu en répétant un signal

d'excitation précédent selon une période de hauteur de son et d'un livre de codes d'excitation (22) constitué par un signal produit préliminairement sont appliqués,

caractérisé en ce qu'il comprend :

5 un moyen de suivi de hauteur de son (11) pour extraire une période de hauteur de son pour chaque unité plus longue que la sous-séquence ;

un moyen de sélection préliminaire de période de hauteur de son (29) pour extraire, pour chacune des sous-séquences, des
10 candidats de période de hauteur de son par rapport à une période de hauteur de son au voisinage de la période de hauteur de son extraite dans ledit moyen de section de suivi de hauteur de son (11) ; et

un moyen de sélection finale de période de hauteur de son
15 pour sélectionner une période de hauteur de son présentant une distorsion de forme d'onde minimum parmi les candidats de période de hauteur de son extraits dans ledit moyen de sélection de période préliminaire de hauteur de son (29) par l'intermédiaire dudit filtre de synthèse de prédiction linéaire
20 (26).

3. Système de codage de hauteur de son de parole permettant de coder un signal de parole en utilisant des paramètres caractéristiques obtenus pour chaque séquence du signal de parole et des paramètres caractéristiques obtenus pour
25 chacune des sous-séquences qui constituent des divisions supplémentaires de la séquence et permettant de synthétiser un signal de parole au moyen d'un filtre de synthèse de prédiction linéaire (26) dans lequel des signaux de source d'excitation d'un livre de codes adaptatifs (21) obtenu en répétant un signal
30 d'excitation précédent pour une période de hauteur de son et d'un livre de codes d'excitation (22) constitué par un signal produit préliminairement sont appliqués,

caractérisé en ce qu'il comprend :

un processeur de séquence (1) pour suivre la hauteur de
35 son en réalisant, à l'aide de la séquence du signal de parole et

des sous-séquences qui constituent des divisions de la séquence, pour sélectionner une voie de suivi de hauteur de son présentant une distorsion de forme d'onde minimum ou un gain de prédiction de hauteur de son moyen maximum parmi B^N combinaisons de
5 voies de suivi de hauteur de son où B est le nombre de bits du codage de hauteur de son dans chaque sous-séquence et N est le nombre de sous-séquences de la séquence ;

un dispositif de production de candidats de hauteur de son pour produire un nombre prédéterminé de candidats de hauteur de son au voisinage de la hauteur de son correspondant à chaque
10 sous-séquence de la voie de suivi de hauteur de son obtenue dans ledit processeur de séquence (1) ;

un calculateur de distorsion de forme d'onde pour calculer une distorsion de forme d'onde en utilisant une différence entre
15 le signal de parole d'entrée et le signal de parole synthétisé sur la base dudit livre de codes adaptatifs (21) et desdits vecteurs de code d'excitation selon chaque combinaison par l'intermédiaire dudit filtre de synthèse (26) ; et

un évaluateur de distorsion minimum (28) pour
20 sélectionner en tant que distorsion de forme d'onde minimum l'une des combinaisons de vecteurs correspondant aux candidats de hauteur de son parmi des vecteurs de code adaptatif accumulés dans ledit livre de codes adaptatifs (21) et parmi les vecteurs de code d'excitation accumulés dans ledit livre de
25 codes d'excitation (22) et pour appliquer la combinaison sélectionnée sur une borne de sortie.

4. Système de codage de hauteur de son de parole permettant de coder un signal de parole selon la revendication 3, caractérisé en ce qu'il comprend en outre un sélecteur
30 préliminaire de hauteur de son (29) pour exécuter une sélection préliminaire de hauteur de son par rapport à chaque sous-séquence au voisinage de la voie de suivi de hauteur de son obtenue dans ledit moyen de suivi de hauteur de son (11).

5. Système de codage de hauteur de son de parole permettant de coder un signal de parole selon la revendication 3, caractérisé en ce que ledit processeur de séquence (1) détermine une voie en sélectionnant successivement des hauteurs de son à
- 5 partir de l'une quelconque des sous-séquences.

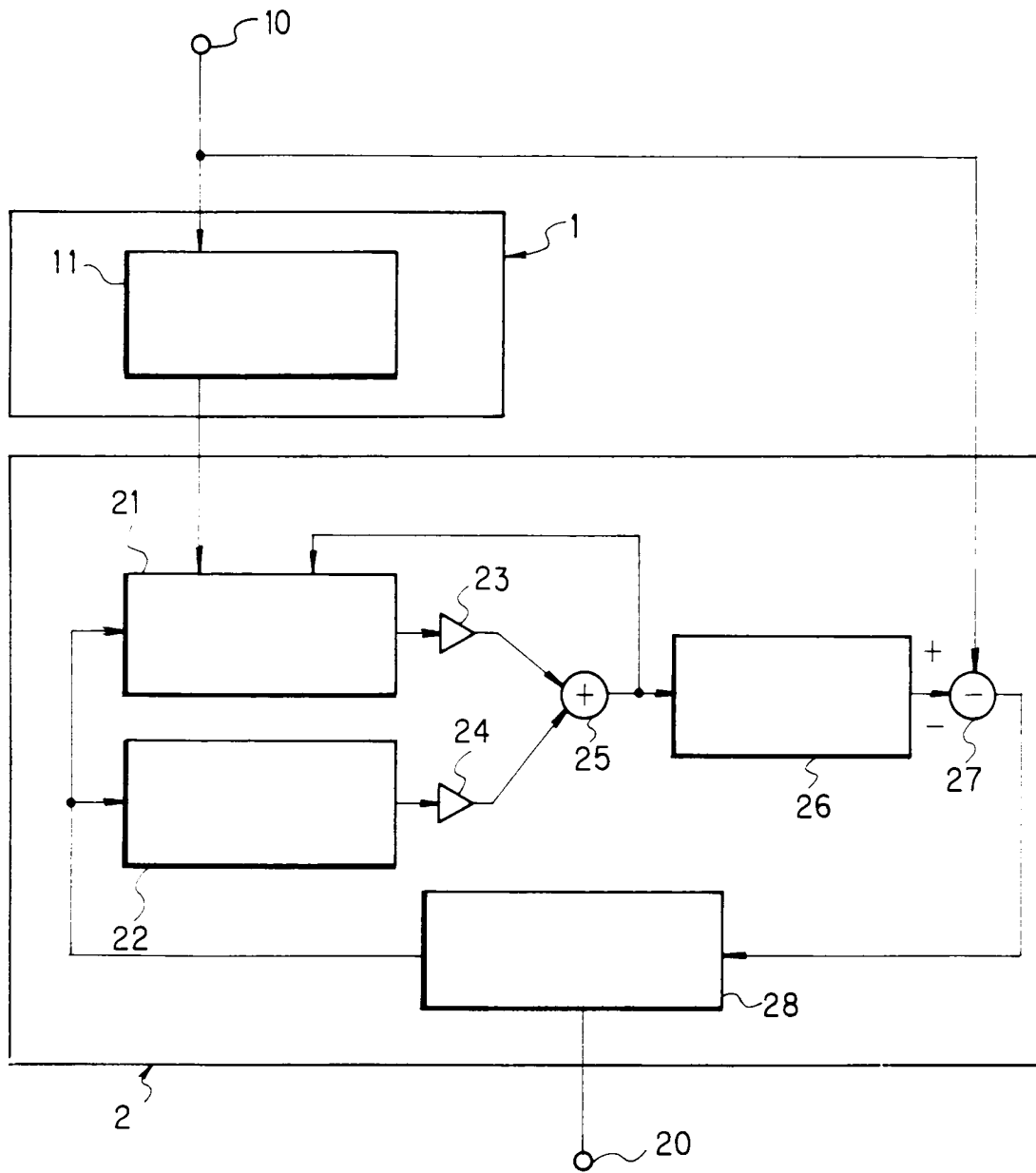
FIG. 1

FIG. 2

