

(19)



(11)

EP 1 316 087 B1

(12)

FASCICULE DE BREVET EUROPEEN

(45) Date de publication et mention de la délivrance du brevet:
02.01.2008 Bulletin 2008/01

(51) Int Cl.:
G10L 19/00^(2006.01) H04L 1/00^(2006.01)

(21) Numéro de dépôt: **01969857.0**

(86) Numéro de dépôt international:
PCT/FR2001/002747

(22) Date de dépôt: **05.09.2001**

(87) Numéro de publication internationale:
WO 2002/021515 (14.03.2002 Gazette 2002/11)

(54) **DISSIMULATION D'ERREURS DE TRANSMISSION DANS UN SIGNAL AUDIO**

ÜBERTRAGUNGSFEHLER-VERDECKUNG IN EINEM AUDIOSIGNAL

TRANSMISSION ERROR CONCEALMENT IN AN AUDIO SIGNAL

(84) Etats contractants désignés:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE TR**

(56) Documents cités:
**FR-A- 2 774 827 US-A- 5 717 822
US-A- 5 884 010**

(30) Priorité: **05.09.2000 FR 0011285**

(43) Date de publication de la demande:
04.06.2003 Bulletin 2003/23

(73) Titulaire: **FRANCE TELECOM
75015 Paris (FR)**

(72) Inventeurs:
• **KOVESI, Balazs**
F-22300 Lannion (FR)
• **MASSALOUX, Dominique**
F-22700 Perros-Guirec (FR)
• **DELEAM, David**
F-22700 Perros Guirec (FR)

- **COMBESCURE P ET AL: "A 16, 24, 32 KBIT/S WIDEBAND SPEECH CODEC BASED ON ATCELP" PHOENIX, AZ, MARCH 15 - 19, 1999, NEW YORK, NY: IEEE, US, 15 mars 1999 (1999-03-15), pages 5-8, XP000898251 ISBN: 0-7803-5042-1 cité dans la demande**
- **ERDOL N ET AL: "RECOVERY OF MISSING SPEECH PACKETS USING THR SHORT-TIME ENERGY AND ZERO-CROSSING MEASUREMENTS" IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, US, IEEE INC. NEW YORK, vol. 1, no. 3, 1 juillet 1993 (1993-07-01), pages 295-303, XP000388573 ISSN: 1063-6676**

(74) Mandataire: **Joly, Jean-Jacques et al**
Cabinet Beau de Loménie
158, rue de l'Université
75340 Paris Cédex 07 (FR)

EP 1 316 087 B1

Il est rappelé que: Dans un délai de neuf mois à compter de la date de publication de la mention de la délivrance du brevet européen, toute personne peut faire opposition au brevet européen délivré, auprès de l'Office européen des brevets. L'opposition doit être formée par écrit et motivée. Elle n'est réputée formée qu'après paiement de la taxe d'opposition. (Art. 99(1) Convention sur le brevet européen).

Description

1. DOMAINE TECHNIQUE

5 **[0001]** La présente invention concerne les techniques de dissimulation d'erreurs de transmission consécutives dans des systèmes de transmission utilisant n'importe quel type de codage numérique du signal de la parole et/ou du son.

[0002] On distingue classiquement deux grandes catégories de codeurs :

- 10 - les codeurs dits temporels, qui effectuent la compression des échantillons de signal numérisé échantillon par échantillon (cas des codeurs MIC ou MICDA [DAUMER][MAITRE] par exemple)
- et les codeurs paramétriques qui analysent des trames successives d'échantillons du signal à coder pour extraire, à chacune de ces trames, un certain nombre de paramètres qui sont ensuite codés et transmis (cas des vocodeurs [TREMAIN], des codeurs IMBE [HARDWICK], ou des codeurs par transformée [BRANDENBURG]).

15 **[0003]** Il existe des catégories intermédiaires qui complètent le codage des paramètres représentatifs des codeurs paramétriques par le codage d'une forme d'onde temporelle résiduelle. Pour simplifier, ces codeurs peuvent être rangés dans la catégorie des codeurs paramétriques.

[0004] Dans cette catégorie on trouve les codeurs prédictifs et notamment la famille des codeurs à analyse par synthèse tels le RPE-LTP ([HELLWIG]) ou les CELP ([ATAL]).

20 **[0005]** Pour tous ces codeurs, les valeurs codées sont ensuite transformées en un train binaire qui sera transmis sur un canal de transmission. Selon la qualité de ce canal et le type de transport, des perturbations peuvent affecter le signal transmis et produire des erreurs sur le train binaire reçu par le décodeur. Ces erreurs peuvent intervenir de manière isolée dans le train binaire mais se produisent très fréquemment par rafales. C'est alors un paquet de bits correspondant à une portion complète de signal qui est erroné ou non reçu. Ce type de problèmes se rencontre par exemple pour les

25 transmissions sur les réseaux mobiles. Il se rencontre aussi dans les transmissions sur les réseaux par paquets et en particulier sur les réseaux de type internet.

[0006] Lorsque le système de transmission ou les modules chargés de la réception permettent de détecter que les données reçues sont fortement erronées (par exemple sur les réseaux mobiles), ou qu'un bloc de données n'a pas été

30 reçu (cas de systèmes à transmission par paquets par exemple), des procédures de dissimulation des erreurs sont alors mises en oeuvre. Ces procédures permettent d'extrapoler au décodeur les échantillons du signal manquant à partir des signaux et données disponibles issus des trames précédant et éventuellement suivant les zones effacées.

[0007] De telles techniques ont été mises en oeuvre principalement dans le cas des codeurs paramétriques (techniques de récupération des trames effacées). Elles permettent de limiter fortement la dégradation subjective du signal perçue au décodeur en présence de trames effacées. La plupart des algorithmes développés reposent sur la technique utilisée

35 pour le codeur et le décodeur, et constituent en fait une extension du décodeur.

Un but général de l'invention est d'améliorer, pour tout système de compression de la parole et du son, la qualité subjective du signal de parole restitué au décodeur lorsque, à cause d'une mauvaise qualité du canal de transmission ou par suite de la perte ou non réception d'un paquet dans un système à transmission par paquets, un ensemble de données codées consécutives ont été perdues.

40 **[0008]** Elle propose à cet effet une technique permettant de dissimuler les erreurs de transmission successives (paquets d'erreur) quelle que soit la technique de codage utilisée, la technique proposée pouvant être utilisée par exemple dans le cas des codeurs temporels dont la structure se prête moins bien a priori à la dissimulation des paquets d'erreurs.

2. ETAT DE LA TECHNIQUE ANTERIEURE

45 **[0009]** La plupart des algorithmes de codage de type prédictif proposent des techniques de récupération de trames effacées ([GSM-FR], [REC G.723.1A], [SALAMI], [HONKANEN], [COX-2], [CHEN-2], [CHEN-3], [CHEN-4], [CHEN-5], [CHEN-6], [CHEN-7], [KROON-2], [WATKINS]). Le décodeur est informé de l'occurrence d'une trame effacée d'une manière ou d'une autre, par exemple dans le cas des systèmes radiomobiles par la transmission de l'information d'ef-

50 facement de trame provenant du décodeur canal. Les dispositifs de récupération de trames effacées ont pour objectif d'extrapoler les paramètres de la trame effacée à partir de la (ou des) dernières trames précédentes considérées comme valides. Certains paramètres manipulés ou codés par les codeurs prédictifs présentent une forte corrélation inter-trames (cas des paramètres de prédiction à court terme, encore dénommés "LPC" de "Linear Predictive Coding" (voir [RABINER]) qui représentent l'enveloppe spectrale, et des paramètres de prédiction à long terme pour les sons voisés, par exemple).

55 Du fait de cette corrélation, il est beaucoup plus avantageux de réutiliser les paramètres de la dernière trame valide pour synthétiser la trame effacée que d'utiliser des paramètres erronés ou aléatoires.

[0010] Pour l'algorithme de codage CELP (de "Code Excited Linear Prediction", se reporter à [RABINER]), les paramètres de la trame effacée sont classiquement obtenus de la manière suivante :

EP 1 316 087 B1

- le filtre LPC est obtenu à partir des paramètres LPC de la dernière trame valide soit par recopie des paramètres ou avec introduction d'un certain amortissement (cf. codeur G723.1 [REC G.723.1A]).
- on détecte le voisement pour déterminer le degré d'harmonicité du signal au niveau de la trame effacée ([SALAMI], cette détection se intervenant de la façon suivante :

■ dans le cas d'un signal non voisé :

un signal d'excitation est généré de manière aléatoire (tirage d'un mot de code et gain de l'excitation passée légèrement amorti [SALAMI], sélection aléatoire dans l'excitation passée [CHEN], usage des codes transmis éventuellement totalement erronés [HONKANEN],...)

■ dans le cas d'un signal voisé :

le délai LTP est généralement le délai calculé à la trame précédente, éventuellement avec une légère "gigue" ([SALAMI]), le gain LTP étant pris très voisin de 1 ou égal à 1. Le signal d'excitation est limité à la prédiction à long terme effectuée à partir de l'excitation passée.

[0011] Dans tous les exemples cités précédemment, les procédures de dissimulation des trames effacées sont fortement liées au décodeur et utilisent des modules de ce décodeur, comme le module de synthèse du signal. Ils utilisent aussi des signaux intermédiaires disponibles au sein de ce décodeur comme le signal d'excitation passé et mémorisé lors du traitement des trames valides précédant les trames effacées.

[0012] La plupart des méthodes utilisées pour dissimuler les erreurs produites par des paquets perdus lors du transport de données codées par des codeurs de type temporel font appel à des techniques de substitution de formes d'ondes telles celles présentées dans [GOODMAN], [ERDÖL], [AT&T]. Les méthodes de ce type reconstituent le signal en sélectionnant des portions du signal décodé avant la période perdue et ne font pas appel à des modèles de synthèse. Des techniques de lissage sont également mises en oeuvre pour éviter les artefacts produits par la concaténation des différents signaux.

[0013] Pour les codeurs par transformée, les techniques de reconstruction des trames effacées s'appuient également sur la structure de codage utilisée : les algorithmes, tels [PICTEL, MAHIEUX-2], visent à régénérer les coefficients transformés perdus à partir des valeurs prises par ces coefficients avant l'effacement.

[0014] La méthode décrite dans [PARIKH] peut s'appliquer à tout type de signaux; elle repose sur la construction d'un modèle sinusoïdal à partir du signal valide décodé précédant l'effacement, pour régénérer la partie du signal perdue.

[0015] Enfin, il existe une famille de techniques de dissimulation des trames effacées développées conjointement avec le codage canal. Ces méthodes, telle celle décrite dans [FINGSCHEIDT], se servent d'informations fournies par le décodeur canal, par exemple d'informations concernant le degré de fiabilité des paramètres reçus. Elles sont fondamentalement différentes de la présente invention qui ne présuppose pas l'existence d'un codeur canal.

[0016] Un art antérieur qui peut être considéré comme le plus proche de la présente invention est celui qui est décrit dans [COMBESCURE], qui proposait une méthode de dissimulation des trames effacées équivalente à celle utilisée dans les codeurs CELP pour un codeur par transformée. Les inconvénients de la méthode proposée étaient l'introduction de distorsions spectrales audibles (voix "synthétique", résonances parasites,...), dus notamment à l'usage de filtres de synthèse à long terme mal contrôlés (composante harmonique unique en sons voisés, génération du signal d'excitation limitée à l'usage de portions du signal résiduel passé). En outre, le contrôle d'énergie s'effectuait dans [COMBESCURE] au niveau du signal d'excitation, la cible énergétique de ce signal était maintenue constante pendant toute la durée de l'effacement, ce qui engendrait également des artefacts gênants. Les mêmes remarques s'appliquent au document US5884010.

3. PRESENTATION DE L'INVENTION

[0017] L'invention telle que définie dans les revendications 1, 17 et 18 permet quant à elle la dissimulation des trames effacées sans distorsion marquée à des taux d'erreurs plus élevés et/ou pour des intervalles effacés plus longs.

[0018] Elle propose notamment un procédé de dissimulation d'erreur de transmission dans un signal audio-numérique selon lequel on reçoit un signal décodé après transmission, on mémorise les échantillons décodés lorsque les données transmises sont valides, on estime au moins un opérateur de prédiction à court terme et au moins un opérateur de prédiction à long terme en fonction des échantillons valides mémorisés et on génère d'éventuels échantillons manquants ou erronés dans le signal décodé à l'aide des opérateurs ainsi estimés.

[0019] Selon un premier aspect particulièrement avantageux de l'invention, on contrôle l'énergie du signal de synthèse ainsi généré à l'aide d'un gain calculé et adapté échantillon par échantillon.

[0020] Ceci contribue en particulier à améliorer les performances de la technique sur des zones d'effacement d'une durée plus longue.

[0021] Notamment, le gain pour le contrôle du signal de synthèse est avantageusement calculé en fonction d'au moins un des paramètres suivantes : valeurs d'énergie préalablement mémorisées pour les échantillons correspondant à des données valides, période fondamentale pour les sons voisés, ou tout paramètre caractérisant le spectre de fréquences.

[0022] De façon avantageuse également, le gain appliqué au signal de synthèse décroît progressivement en fonction de la durée pendant laquelle les échantillons de synthèse sont générés.

[0023] De façon préférée également, on discrimine dans les données valides les sons stationnaires et les sons non stationnaires et on met en oeuvre des lois d'adaptation de ce gain (vitesse de décroissance, par exemple), différentes d'une part pour les échantillons générés à la suite de données valides correspondant à des sons stationnaires et d'autre part pour les échantillons générés à la suite de données valides correspondant à des sons non stationnaires.

[0024] Selon un autre aspect indépendant de l'invention, on met à jour en fonction des échantillons de synthèse générés le contenu des mémoires utilisées pour le traitement de décodage.

[0025] De cette façon, d'une part on limite l'éventuelle désynchronisation du codeur et du décodeur (voir paragraphe 5.1.4 ci-après), et on évite les brusques discontinuités entre la zone effacée reconstruite selon l'invention et les échantillons suivant cette zone.

[0026] Notamment, on met en oeuvre au moins partiellement sur les échantillons synthétisés un codage analogue à celui mis en oeuvre à l'émetteur suivi éventuellement d'une opération (éventuellement partielle) de décodage, les données obtenues servant à régénérer les mémoires du décodeur.

[0027] En particulier, cette opération de codage-décodage éventuellement partielle peut être avantageusement utilisée pour régénérer la première trame effacée car elle permet d'exploiter le contenu des mémoires du décodeur avant la coupure, lorsque ces mémoires contiennent des informations non fournies par les derniers échantillons valides décodés (par exemple dans le cas des codeurs par transformée à addition-recouvrement, voir paragraphe 5.2.2.1 point 10).

[0028] Selon un aspect encore différent de l'invention, on génère en entrée de l'opérateur de prédiction à court terme un signal d'excitation qui, en zone voisée, est la somme d'une composante harmonique et d'une composante faiblement harmonique ou non harmonique, et en zone voisée limitée à la composante non harmonique.

[0029] Notamment, la composante harmonique est avantageusement obtenue en mettant en oeuvre un filtrage au moyen de l'opérateur de prédiction à long terme appliqué sur un signal résiduel calculé en mettant en oeuvre un filtrage à court terme inverse sur les échantillons mémorisés.

[0030] L'autre composante peut être déterminée l'aide d'un opérateur de prédiction à long terme auquel on applique des perturbations (par exemple perturbation du gain, ou de la période), pseudo-aléatoires.

[0031] De façon particulièrement préférée, pour la génération d'un signal d'excitation voisé, la composante harmonique représente les basses fréquences du spectre, tandis que l'autre composante la partie haute fréquence.

[0032] Selon un autre aspect encore, l'opérateur de prédiction à long terme est déterminé à partir des échantillons de trames valides mémorisés, avec un nombre d'échantillons utilisés pour cette estimation variant entre une valeur minimale et une valeur égale à au moins deux fois la période fondamentale estimée pour le son voisé.

[0033] Par ailleurs, le signal résiduel est avantageusement modifié par des traitements de type non linéaire pour éliminer des pics d'amplitude.

[0034] Également, selon un autre aspect avantageux, on détecte l'activité vocale en estimant des paramètres de bruit lorsque le signal est considéré comme non actif, et on fait tendre des paramètres du signal synthétisé vers ceux du bruit estimé.

[0035] De façon préférentielle encore, on estime l'enveloppe spectrale du bruit des échantillons décodés valides et on génère un signal synthétisé évoluant vers un signal possédant la même enveloppe spectrale.

[0036] L'invention propose également un procédé de traitement de signaux de sons, caractérisé en ce qu'on met en oeuvre une discrimination entre la parole et les sons musicaux et lorsqu'on détecte des sons musicaux, on met en oeuvre un procédé du type précité sans estimation d'un opérateur de prédiction à long terme, le signal d'excitation étant limité à une composante non harmonique obtenue par exemple en générant un bruit blanc uniforme.

[0037] L'invention concerne en outre un dispositif de dissimulation d'erreur de transmission dans un signal audio-numérique qui reçoit en entrée un signal décodé que lui transmet un décodeur et qui génère des échantillons manquants ou erronés dans ce signal décodé, caractérisé en ce qu'il comporte des moyens de traitement aptes à mettre en oeuvre le procédé précité.

[0038] Elle concerne également un système de transmission comportant au moins un codeur, au moins un canal de transmission, un module apte à détecter que des données transmises ont été perdues ou sont fortement erronées, au moins un décodeur et un dispositif de dissimulation d'erreurs qui reçoit le signal décodé, caractérisé en ce que ce dispositif de dissimulation d'erreurs est un dispositif du type précité.

4. PRESENTATION DES FIGURES

[0039] D'autres caractéristiques et avantages de l'invention ressortiront encore de la description qui suit, laquelle est purement illustrative et non limitative et doit être lue en regard des dessins annexés sur lesquels :

- 5 - la figure 1 est un schéma synoptique illustrant un système de transmission conforme à un mode de réalisation possible de l'invention ;
- 10 - la figure 2 et la figure 3 sont des schémas synoptiques illustrant une mise en oeuvre conforme à un mode possible de l'invention ;
- les figures 4 à 6 illustrent schématiquement les fenêtres utilisées avec le procédé de dissimulation d'erreurs conforme à un mode de mise en oeuvre possible de l'invention ;
- 15 - les figures 7 et 8 sont des représentations schématiques illustrant un mode de mise en oeuvre possible de l'invention dans le cas de signaux musicaux.

5. DESCRIPTION D'UN OU PLUSIEURS MODES DE REALISATIONS POSSIBLE DE L'INVENTION

20 5.1 Principe d'un mode de réalisation possible

[0040] La figure 1 présente un dispositif de codage et décodage du signal audio numérique, comprenant un codeur 1, un canal de transmission 2, un module 3 permettant de détecter que des données transmises ont été perdues ou sont fortement erronées, un décodeur 4, et un module 5 de dissimulation des erreurs ou paquets perdus conforme à un mode de réalisation possible de l'invention.

[0041] On notera que ce module 5, outre l'indication de données effacées, reçoit le signal décodé en période valide et transmet au décodeur des signaux utilisés pour sa mise à jour.

[0042] Plus précisément, le traitement mis en oeuvre par le module 5 repose sur :

- 30 1. la mémorisation des échantillons décodés lorsque les données transmises sont valides (traitement 6);
- 2. durant un bloc de données effacées, la synthèse des échantillons correspondant aux données perdues (traitement 7) ;
- 3. lorsque la transmission est rétablie, le lissage entre les échantillons de synthèse produits pendant la période effacée et les échantillons décodés (traitement 8);
- 35 4. la mise à jour des mémoires du décodeur (traitement 9) (mise à jour qui s'effectue soit pendant la génération des échantillons effacés, soit au moment du rétablissement de la transmission).

5.1.1 En période valide

40 [0043] Après décodage des données valides, on met à jour la mémoire des échantillons décodés, contenant un nombre d'échantillons suffisant pour la régénération des éventuelles périodes effacées dans la suite. Typiquement, on mémorise de l'ordre de 20 à 40 ms de signal. On calcule également l'énergie des trames valides et on retient en mémoire les énergies correspondant aux dernières trames valides traitées (typiquement de l'ordre de 5 s).

45 5.1.2 Pendant un bloc de données effacées.

[0044] On effectue les opérations suivantes, illustrées par la figure 3 :

- 50 1. Estimation de l'enveloppe spectrale courante :

[0045] On calcule cette enveloppe spectrale sous la forme d'un filtre LPC [RABINER][KLEIJN]. L'analyse est effectuée par des méthodes classiques ([KLEIJN]) après fenêtrage des échantillons mémorisés en période valide. Notamment, on met en oeuvre une analyse LPC (étape 10) pour obtenir les paramètres d'un filtre A(z), dont l'inverse est utilisé pour le filtrage LPC (étape 11). Comme les coefficients ainsi calculés n'ont pas à être transmis, on peut utiliser pour cette analyse un ordre élevé, ce qui permet d'obtenir de bonnes performances sur les signaux musicaux.

2. Détection des sons voisés et calcul des paramètres LTP:

[0046] Une méthode de détection des sons voisés (traitement 12 de la figure 3 : détection V/NV, pour "voisé/non voisé") est utilisée sur les dernières données mémorisées. Par exemple on peut utiliser pour cela la corrélation normalisée ([KLEIJN]), ou le critère présenté dans l'exemple de réalisation qui suit.

[0047] Lorsque le signal est déclaré voisé, on calcule les paramètres permettant la génération d'un filtre de synthèse à long terme, encore dénommé filtre LTP ([KLEIJN]) (figure 3 : analyse LTP, on définit par B (Z) le filtre inverse LTP calculé). Un tel filtre est généralement représenté par une période correspondant à la période fondamentale et un gain. La précision de ce filtre peut être améliorée par l'usage de pitch fractionnaire ou d'une structure multi-coefficients [KROON].

[0048] Lorsque le signal est déclaré non voisé, une valeur particulière est attribuée au filtre de synthèse LTP (voir paragraphe 4).

[0049] Il est particulièrement intéressant dans cette estimation du filtre de synthèse LTP de restreindre la zone analysée à la fin de la période précédant l'effacement. La longueur de la fenêtre d'analyse varie entre une valeur minimale et une valeur liée à la période fondamentale du signal.

3. Calcul d'un signal résiduel :

[0050] On calcule un signal résiduel par filtrage inverse LPC (traitement 10) des derniers échantillons mémorisés. Ce signal est ensuite utilisé pour générer un signal d'excitation du filtre de synthèse LPC 11 (voir ci-dessous).

4. Synthèse des échantillons manquants :

[0051] La synthèse des échantillons de remplacement s'effectue en introduisant un signal d'excitation (calculé en 13 à partir du signal en sortie du filtre LPC inverse) dans le filtre de synthèse LPC 11 ($1/A(z)$) calculé en 1. Ce signal d'excitation est engendré de deux façons différentes suivant que le signal est voisé ou non voisé:

4.1 En zone voisée :

[0052] Le signal d'excitation est la somme de deux signaux, une composante fortement harmonique et l'autre moins harmonique ou pas du tout.

[0053] La composante fortement harmonique est obtenue par filtrage LTP (module de traitement 14) à l'aide des paramètres calculés en 2, du signal résiduel mentionné en 3.

[0054] La seconde composante peut être obtenue également par filtrage LTP mais rendu non périodique par des modifications aléatoires des paramètres, par génération d'un signal pseudo-aléatoire.

[0055] Il est particulièrement intéressant de limiter la bande passante de la première composante aux basses fréquences du spectre. De même il sera intéressant de limiter aux plus hautes fréquences la seconde composante.

4.2 En zone non voisée :

[0056] Lorsque le signal est non voisé, un signal d'excitation non harmonique est engendré. Il est intéressant d'utiliser une méthode de génération similaire à celle utilisée pour les sons voisés, avec des variations de paramètres (période, gain, signes) permettant de la rendre non harmonique.

4.3 Contrôle de l'amplitude du signal résiduel :

[0057] Lorsque le signal est non voisé, ou faiblement voisé, le signal résiduel utilisé pour la génération de l'excitation est traité pour éliminer les pics d'amplitude significativement au dessus de la moyenne.

5. Contrôle de l'énergie du signal de synthèse

[0058] L'énergie du signal de synthèse est contrôlée à l'aide d'un gain calculé et adapté échantillon par échantillon. Dans le cas où la période d'effacement est relativement longue, il est nécessaire de faire baisser progressivement l'énergie du signal de synthèse. La loi d'adaptation du gain est calculée en fonction de différents paramètres : valeurs d'énergies mémorisées avant l'effacement (voir en 1), période fondamentale, et stationnarité locale du signal au moment de la coupure.

[0059] Si le système comprend un module permettant la discrimination des sons stationnaires (comme la musique) et non stationnaires (comme la parole), des lois d'adaptation différentes peuvent aussi être utilisées.

[0060] Dans le cas de codeurs par transformée avec addition-recouvrement, la première moitié de la mémoire de la dernière trame correctement reçue contient des informations assez précises sur la première moitié de la première trame perdue (son poids dans l'addition-recouvrement est plus important que celui de la trame actuelle). Cette information peut être également utilisée pour le calcul du gain adaptatif.

5

6. Evolution de la procédure de synthèse au cours du temps :

[0061] Dans le cas de période d'effacement relativement longues, on peut également faire évoluer les paramètres de synthèse. Si le système est couplé à un dispositif de détection d'activité vocale avec estimation des paramètres du bruit (tel [REC-G.723.1A], [SALAMI-2], [BENYASSINE]), il est particulièrement intéressant de faire tendre les paramètres de génération du signal à reconstruire vers ceux du bruit estimé: en particulier au niveau de l'enveloppe spectrale (interpolation du filtre LPC avec celui du bruit estimé, les coefficients de l'interpolation évoluant au cours du temps jusqu'à obtention du filtre du bruit) et de l'énergie (niveau évoluant progressivement vers celui du bruit, par exemple par fenêtrage).

10

15 5.1.3 Au rétablissement de la transmission

[0062] Au rétablissement de la transmission, il est particulièrement important d'éviter les ruptures brutales entre la période effacée que l'on a reconstruite selon les techniques définies aux paragraphes précédents et les périodes qui suivent, au cours desquelles on dispose de toute l'information transmise pour décoder le signal. La présente invention effectue une pondération dans le domaine temporel avec interpolation entre les échantillons de remplacement précédent le rétablissement de la communication et les échantillons décodés valides suivant la période effacée. Cette opération est a priori indépendante du type du codeur employé.

20

[0063] Dans le cas de codeurs par transformée avec addition-recouvrement, cette opération est commune avec la mise à jour des mémoires décrite au paragraphe suivant (voir exemple de réalisation).

25

5.1.4 Mise à jour des mémoires du décodeur

[0064] Lorsque le décodage des échantillons valides reprend après une période effacée, il peut y avoir une dégradation lorsque le décodeur utilise des données normalement produites aux trames précédentes et mémorisées. Il est important de mettre à jour proprement ces mémoires pour éviter ces artefacts.

30

[0065] Ceci est particulièrement important pour les structures de codage utilisant des procédés récursifs, qui pour un échantillon ou une séquence d'échantillons, se servent d'informations obtenues après décodage des échantillons précédents. Ce sont par exemple des prédictions ([KLEIJN]) qui permettent d'extraire de la redondance du signal. Ces informations sont normalement disponibles à la fois au codeur, qui doit pour cela avoir effectué pour ces échantillons précédents une forme de décodage local, et au décodeur distant présent à la réception. Dès que le canal de transmission est perturbé et que le décodeur distant ne dispose plus des mêmes informations que le décodeur local présent à l'émission, il y a désynchronisation entre le codeur et le décodeur. Dans le cas de systèmes de codage fortement récursifs, cette désynchronisation peut provoquer des dégradations audibles qui peuvent perdurer longtemps voir même s'amplifier au cours du temps s'il existe des instabilités dans la structure. Dans ce cas, il est donc important de s'efforcer de resynchroniser le codeur et le décodeur, c'est à dire de faire une estimation des mémoires du décodeur la plus proche possible de celles du codeur. Cependant les techniques de resynchronisation dépendent de la structure de codage utilisée. On en présentera une dont le principe est général dans le présent brevet, mais dont la complexité est potentiellement importante.

35

40

[0066] Une méthode possible consiste à introduire dans le décodeur à la réception un module de codage du même type que celui présent à l'émission, permettant d'effectuer le codage-décodage des échantillons du signal produit par les techniques mentionnées au paragraphe précédent pendant les périodes effacées. De cette façon les mémoires nécessaires pour décoder les échantillons suivant sont complétées avec des données a priori proches (sous réserve d'une certaine stationnarité pendant la période effacée) de celles qui ont été perdues. Dans le cas où cette hypothèse de stationnarité ne serait pas respectée, après une longue période effacée par exemple, on ne dispose pas de toute façon d'informations suffisantes pour faire mieux.

45

50

[0067] En fait il n'est généralement pas nécessaire d'effectuer le codage complet de ces échantillons, on se limite aux modules nécessaires pour mettre à jour les mémoires.

[0068] Cette mise à jour peut s'effectuer au moment de la production des échantillons de remplacement, ce qui répartit la complexité sur toute la zone d'effacement, mais se cumule avec la procédure de synthèse décrite précédemment.

55

[0069] Lorsque la structure de codage le permet, on peut aussi limiter la procédure ci-dessus à une zone intermédiaire au début de la période de données valides succédant à une période effacée, la procédure de mise à jour se cumulant alors avec l'opération de décodage.

5.2. Description d'exemples de réalisation particuliers

[0070] Des exemples particuliers de mise en oeuvre possible sont donnés ci-après. Le cas des codeurs par transformée de type TDAC ou TCDM ([MAHIEUX]) est en particulier abordé.

5.2.1 Description du dispositif

[0071] Système de codage/décodage numérique par transformée de type TDAC.

[0072] Codeur en bande élargie (50-7000 Hz) à 24 kb/s ou 32 kb/s.

[0073] Trame de 20 ms (320 échantillons).

[0074] Fenêtres de 40 ms (640 échantillons) avec addition-recouvrements de 20 ms. Une trame binaire contient les paramètres codés obtenus par la transformation TDAC sur une fenêtre. Après le décodage de ces paramètres, en faisant la transformation inverse TDAC, on obtient une trame de sortie de 20 ms qui est la somme de la deuxième moitié de la fenêtre précédente et la première moitié de la fenêtre actuelle. Sur la figure 4, il a été marqué en gras les deux parties de fenêtres utilisées pour la reconstruction de la trame n (en temporel). Ainsi, une trame binaire perdue perturbe la reconstruction de deux trames consécutives (l'actuelle et la suivante, figure 5). Par contre, en faisant correctement le remplacement des paramètres perdus, on peut récupérer les parties de l'information provenant de la trame binaire précédente et suivante (figure 6), pour la reconstruction de ces deux trames.

5.2.2 Mise en oeuvre

[0075] Toutes les opérations décrites ci-dessous sont mises en oeuvre à la réception, conformément aux figures 1 et 2, soit au sein du module de dissimulation des trames effacées qui communique avec le décodeur, soit dans le décodeur lui même (mise à jour des mémoires du décodeur).

5.2.2.1 En période valide

[0076] Correspondant au paragraphe 5.1.2, on met à jour la mémoire des échantillons décodés. Cette mémoire est utilisée pour les analyses LPC et LTP du signal passé dans le cas d'un effacement d'une trame binaire. Dans l'exemple présenté ici, l'analyse LPC est faite sur une période de signal de 20 ms (320 échantillons). En général, l'analyse LTP nécessite plus d'échantillons à mémoriser. Dans notre exemple, pour pouvoir faire l'analyse LTP correctement, le nombre des échantillons mémorisés est égal à deux fois la valeur maximale du pitch. Par exemple, si la valeur maximale du pitch MaxPitch est fixée à 320 échantillons (50 Hz, 20 ms), les derniers 640 échantillons seront mémorisés (40 ms du signal). On calcule également l'énergie des trames valides et on les stocke dans un tampon circulaire de longueur de 5 s. Lorsqu'une trame effacée est détectée, on compare l'énergie de la dernière trame valide au maximum et au minimum de ce tampon circulaire pour connaître son énergie relative.

5.2.2.2 Pendant un bloc de données effacées

[0077] Lorsqu'une trame binaire est perdue, on distingue deux cas différents :

5.2.2.2.1 Première trame binaire perdue après une période valide

[0078] D'abord, on fait une analyse du signal mémorisé pour estimer les paramètres du modèle servant à synthétiser le signal régénéré. Ce modèle nous permet ensuite de synthétiser 40 ms de signal, ce qui correspond à la fenêtre de 40 ms perdue. En faisant la transformation TDAC suivie de la transformation inverse TDAC sur ce signal synthétisé (sans codage - décodage des paramètres), on obtient le signal de sortie de 20 ms. Grâce à ces opérations TDAC - TDAC inverse, on exploite l'information provenant de la fenêtre précédente correctement reçue (voir figure 6). En même temps, on met à jour les mémoires du décodeur. Ainsi, la trame binaire suivante, si elle est bien reçue, peut être décodée normalement, et les trames décodées seront automatiquement synchronisées (figure 6).

[0079] Les opérations à effectuer sont les suivantes :

1. Fenêtrage du signal mémorisé. Par exemple, on peut utiliser une fenêtre asymétrique de Hamming de 20 ms.
2. Calcul de la fonction d'autocorrélation sur le signal fenêtré.
3. Détermination des coefficients du filtre LPC. Pour cela, classiquement on utilise l'algorithme itératif de Levinson-Durbin. L'ordre d'analyse peut être élevé, surtout lorsque le codeur est utilisé pour coder des séquences de musique.

4. Détection de voisement et analyse à long terme du signal mémorisé pour la modélisation de l'éventuelle périodicité du signal (sons voisés). Dans la réalisation présentée, les inventeurs ont limité l'estimation de la période fondamentale T_p aux valeurs entières, et calculé une estimation du degré de voisement sous la forme du coefficient de corrélation $MaxCorr$ (voir ci-dessous) évalué à la période sélectionnée. Soit $T_m = \max(T, F_s/200)$, où F_s est la fréquence d'échantillonnage, donc $F_s/200$ échantillons correspondent à une durée de 5 ms. Pour mieux modéliser l'évolution du signal à la fin de la trame précédente, on calcule les coefficients de corrélation $Corr(T)$ correspondant à un retard T en n'utilisant que $2 \cdot T_m$ échantillons à la fin du signal mémorisé :

$$Corr(T) = \frac{2 \sum_{i=L_{mem}-2T_m+T}^{L_{mem}-1} m_i m_{i-T}}{\sum_{i=L_{mem}-2T_m}^{L_{mem}-1} m_i^2 + \sum_{i=L_{mem}-2T_m+T}^{L_{mem}-1-T} m_i^2}$$

où $m_0 \dots m_{L_{mem}-1}$ est la mémoire du signal décodé précédemment. De cette formule, on voit que la longueur de cette mémoire L_{mem} doit être au moins 2 fois la valeur maximale de la période fondamentale (encore appelée "pitch") $MaxPitch$.

On a également fixé la valeur minimale de la période fondamentale $MinPitch$ correspondant à une fréquence de 600 Hz (26 échantillons à $F_s = 16$ kHz).

On calcule $Corr(T)$ pour $T = 2, \dots, MaxPitch$. Si T' est le plus petit retard tel que $Corr(T') < 0$ (on élimine ainsi les corrélations à très court terme), alors on cherche $MaxCorr$, maximum de $Corr(T)$ pour $T' < T \leq MaxPitch$. Soit T_p la période correspondant à $MaxCorr$ ($Corr(T_p) = MaxCorr$). On cherche également $MaxCorrMP$, maximum de $Corr(T)$ pour $T' < T \leq 0.75 \cdot MinPitch$. Si $T_p < MinPitch$ ou $MaxCorrMP > 0.7 \cdot MaxCorr$ et si l'énergie de la dernière trame valide est relativement faible, on décide que la trame est non voisée, car en utilisant la prédiction LTP on risquerait d'obtenir une résonance dans les hautes fréquences très gênante. Le pitch choisi est $T_p = MaxPitch/2$, et le coefficient de corrélation $MaxCorr$ fixé à une valeur faible (0.25).

On considère également la trame comme non-voisée lorsque plus de 80% de son énergie se concentre dans les derniers $MinPitch$ échantillons. Il s'agit donc d'un démarrage de la parole, mais le nombre d'échantillons n'est pas suffisant pour estimer la période fondamentale éventuelle, il vaut mieux le traiter comme trame non voisée, et même diminuer plus rapidement l'énergie du signal synthétisé (pour signaler cela, on met $DiminFlag=1$).

Dans le cas où $MaxCorr > 0.6$, on vérifie que l'on n'a pas trouvé un multiple (4, 3 ou 2 fois) de la période fondamentale. Pour cela, on cherche le maximum local de la corrélation autour de $T_p/4$, $T_p/3$ et $T_p/2$. Notons T_1 la position de ce maximum, et $MaxCorrL = Corr(T_1)$. Si $T_1 > MinPitch$ et $MaxCorrL > 0.75 \cdot MaxCorr$, on choisit T_1 comme nouvelle période fondamentale.

Si T_p est inférieur à $MaxPitch/2$, on peut vérifier s'il s'agit vraiment d'une trame voisée en cherchant le maximum local de la corrélation autour de $2 \cdot T_p$ (TPP) et en vérifiant si $Corr(T_{pp}) > 0.4$. Si $Corr(T_{pp}) < 0.4$ et si l'énergie du signal diminue, on met $DiminFlag=1$ et on diminue la valeur de $MaxCorr$, sinon on cherche le maximum local suivant entre le T_p actuel et $MaxPitch$.

Un autre critère de voisement consiste à vérifier si au moins dans 2/3 des cas le signal retardé par la période fondamentale a le même signe que le signal non retardé.

On vérifie cela sur une longueur égale au maximum entre 5ms et $2 \cdot T_p$.

On vérifie également si l'énergie du signal a tendance à diminuer ou non. Si oui, on met $DiminFlag=1$ et on fait décroître la valeur de $MaxCorr$ en fonction de degré de diminution.

La décision de voisement tient compte également de l'énergie du signal: si l'énergie est forte, on augmente la valeur de $MaxCorr$, ainsi il est plus probable que la trame soit décidée voisée. Par contre, si l'énergie est très faible, on diminue la valeur de $MaxCorr$.

Finalement, on prend la décision de voisement en fonction de la valeur de $MaxCorr$: la trame est non voisée si et seulement si $MaxCorr < 0.4$. La période fondamentale T_p d'une trame non voisée est bornée, elle doit être inférieure ou égale à $MaxPitch/2$.

5. Calcul du signal résiduel par filtrage inverse LPC des derniers échantillons mémorisés. Ce signal résiduel est stocké dans la mémoire $ResMem$.

6. Egalisation de l'énergie du signal résiduel. Dans le cas d'un signal non voisé ou faiblement voisé ($MaxCorr <$

EP 1 316 087 B1

0.7), l'énergie du signal résiduel stocké dans ResMem peut changer brusquement d'une partie à l'autre. La répétition de cette excitation entraîne une perturbation périodique très désagréable dans le signal synthétisé. Pour éviter cela, on s'assure qu'aucun pic d'amplitude important ne se présente dans l'excitation d'une trame faiblement voisée. Comme l'excitation est construite à partir des derniers T_p échantillons du signal résiduel, on traite ce vecteur de T_p échantillons. La méthode utilisée dans notre exemple est la suivante :

- On calcule la moyenne MeanAmpl des valeurs absolues des derniers T_p échantillons du signal résiduel.
- Si le vecteur d'échantillons à traiter contient n passages à zéro, on le coupe en $n+1$ sous-vecteurs, le signe du signal dans chaque sous-vecteur étant donc invariant.
- On cherche l'amplitude maximale MaxAmplSv de chaque sous-vecteur. Si $\text{MaxAmplSv} > 1.5 * \text{MeanAmpl}$, on multiplie le sous-vecteur par $1.5 * \text{MeanAmpl} / \text{MaxAmplSv}$.

7. Préparation du signal d'excitation d'une longueur de 640 échantillons correspondant à la longueur de la fenêtre TDAC. On distingue 2 cas selon le voisement :

✘ Le signal d'excitation est la somme de deux signaux, une composante fortement harmonique limitée en bande aux basses fréquences du spectre excb et une autre moins harmonique limitée aux plus hautes fréquences exch.

La composante fortement harmonique est obtenue par filtrage LTP d'ordre 3 du signal résiduel :

$$\text{excb}(i) = 0.15 * \text{exc}(i - T_p - 1) + 0.7 * \text{exc}(i - T_p) + 0.15 * \text{exc}(i - T_p + 1)$$

Les coefficients [0.15, 0.7, 0.15] correspondent à un filtre FIR passe-bas de 3 dB d'atténuation à $F_s/4$. La seconde composante est obtenue également par un filtrage LTP rendu non périodique par la modification aléatoire de sa période fondamentale T_{ph} . T_{ph} est choisie comme la partie entière d'une valeur réelle aléatoire T_{pa} . La valeur initiale de T_{pa} est égale à T_p puis elle est modifiée échantillon par échantillon en l'additionnant une valeur aléatoire dans [-0.5, 0.5]. De plus, ce filtrage LTP est combiné avec un filtrage IIR passe haut :

$$\begin{aligned} \text{exch}(i) = & -0.0635 * (\text{exc}(i - T_{ph} - 1) + \text{exc}(i - T_{ph} + 1)) + \\ & 0.1182 * \text{exc}(i - T_{ph}) - 0.9926 * \text{exch}(i - 1) - \\ & 0.7679 * \text{exch}(i - 2) \end{aligned}$$

L'excitation voisée est alors la somme de ces 2 composante :

$$\text{Exc}(i) = \text{excb}(i) + \text{exch}(i)$$


- Dans le cas d'une trame non voisée, le signal d'excitation exc est obtenu également par filtrage LTP d'ordre 3 avec les coefficients [0.15, 0.7, 0.15] mais il est rendu non périodique par augmentation de la période fondamentale d'une valeur égale à 1 tous les 10 échantillons, et inversion du signe avec une probabilité de 0.2.

8. Synthèse des échantillons de remplacement en introduisant le signal d'excitation exc dans le filtre LPC calculé en 3.

9. Contrôle du niveau de l'énergie du signal de synthèse. L'énergie tend progressivement vers un niveau fixé par avance dès la première trame de remplacement synthétisée. Ce niveau peut être défini, par exemple, comme l'énergie de la trame de sortie la plus faible trouvée durant les 5 dernières secondes précédant l'effacement. Nous avons défini deux lois d'adaptation du gain qui sont choisies en fonction du drapeau DiminFlag calculé en 4. La vitesse de diminution de l'énergie dépend également de la période fondamentale. Il existe une troisième loi d'adaptation plus radicale qui est utilisée quand on détecte que le début du signal généré ne correspond pas bien au signal original, comme expliqué ultérieurement (voir point 11).

10. Transformation TDAC sur le signal synthétisé en 8, comme expliqué au début de ce chapitre. Les coefficients TDAC obtenus remplacent les coefficients TDAC perdus. Ensuite, en faisant la transformation inverse TDAC, on obtient la trame de sortie. Ces opérations ont trois buts :

- 5 ■ Dans le cas de la première fenêtre perdue, de cette façon on exploite l'information de la fenêtre précédente correctement reçue qui contient la moitié des données nécessaires pour reconstruire la première trame perturbée (figure 6).
- On met à jour la mémoire du décodeur pour le décodage de la trame suivante (synchronisation du codeur et du décodeur, voir paragraphe 5.1.4).
- 10 ■ On assure automatiquement la transition continue (sans rupture) du signal de sortie lorsque la première trame binaire correctement reçue arrive après une période effacée que l'on a reconstruite selon les techniques présentées ci-dessus (voir paragraphe 5.1.3).

11. La technique d'addition-recouvrement permet de vérifier si le signal voisé synthétisé correspond bien au signal d'origine ou non car pour la première moitié de la première trame perdue le poids de la mémoire de dernière fenêtre correctement reçue est plus important (figure 6). Donc en prenant la corrélation entre la première moitié de la première trame synthétisée et la première moitié de la trame obtenue après les opérations TDAC  TDAC inverse, on peut estimer la similitude entre la trame perdue et la trame de remplacement. Une corrélation faible (< 0.65) signale que le signal originel est assez différent de celui obtenu par la méthode de remplacement, et il vaut mieux diminuer l'énergie de ce dernier rapidement vers le niveau minimal.

5.2.2.2.2 Trames perdues suivant la première trame d'une zone effacée

25 **[0080]** Dans le paragraphe précédent, les points 1-6 concernent l'analyse du signal décodé précédant la première trame effacée et permettant la construction d'un modèle de synthèse (LPC et éventuellement LTP) de ce signal. Pour les trames effacées suivantes, on ne refait pas l'analyse, le remplacement du signal perdu est basé sur les paramètres (coefficients LPC, pitch, MaxCorr, ResMem) calculés lors de première trame effacée. On fait donc uniquement les opérations correspondant à la synthèse du signal et à la synchronisation du décodeur, avec les modifications suivantes par rapport à la première trame effacée :

- Dans la partie synthèse (points 7 et 8), on génère uniquement 320 nouveaux échantillons, car la fenêtre de la transformation TDAC couvre les derniers 320 échantillons générés lors de la trame effacée précédente et ces nouveaux 320 échantillons.
- 35 ■ Dans le cas où la période d'effacement serait relativement longue, il est important de faire évoluer les paramètres de synthèse vers les paramètres d'un bruit blanc ou vers ceux du bruit de fond (voir point 5 dans paragraphe 3.2.2.2). Comme le système présenté dans cet exemple ne comprend pas de VAD /CNG, nous avons, par exemple, la possibilité de faire une ou plusieurs des modifications suivantes :
- 40 ■ Interpolation progressive du filtre LPC avec un filtre plat pour rendre le signal synthétisé moins coloré.
- Augmentation progressive de la valeur du pitch.
- En mode voisé, on bascule en mode non-voisé après un certain temps (par exemple quand l'énergie minimale est atteinte).

45 **[0081] 5.3 Traitement spécifique pour les signaux musicaux.** Si le système comprend un module permettant la discrimination parole/musique, on peut alors, après sélection d'un mode de synthèse de musique mettre en oeuvre un traitement spécifique au signaux musicaux. Sur la figure 7, le module de synthèse de musique a été référencé par 15, celui de synthèse de parole par 16 et le commutateur parole/musique par 17.

50 **[0082]** Un tel traitement met par exemple en oeuvre pour le module de synthèse de musique les étapes suivantes, illustrées sur la figure 8 :

1. Estimation de l'enveloppe spectrale courante :

55 **[0083]** On calcule cette enveloppe spectrale sous la forme d'un filtre LPC [RABINER][KLEIJN]. L'analyse est effectuée par des méthodes classiques ([KLEIJN]). Après fenêtrage des échantillons mémorisés en période valide, on met en oeuvre une analyse LPC pour calculer un filtre LPC A(Z) (étape 19). On utilise pour cette analyse un ordre élevé (>100) afin d'obtenir de bonnes performances sur les signaux musicaux.

2. Synthèse des échantillons manquants :

5 [0084] La synthèse des échantillons de remplacement s'effectue en introduisant un signal d'excitation dans le filtre de synthèse LPC ($1/A(z)$) calculé à l'étape 19. Ce signal d'excitation - calculé dans une étape 20 - est un bruit blanc dont l'amplitude est choisie pour obtenir un signal ayant la même énergie que celle des derniers N échantillons mémorisés en période valide. Sur la figure 8, l'étape de filtrage est référencée par 21.

Exemple du contrôle de l'amplitude du signal résiduel :

10 [0085] Si l'excitation se présente comme un bruit blanc uniforme multiplié par un gain, on peut calculer ce gain G comme suit :

Estimation du gain du filtre LPC:

15 [0086] L'algorithme de Durbin donne l'énergie du signal résiduel. Connaissant également l'énergie du signal à modélisé on estime le gain G_{LPC} du filtre LPC comme le rapport de ces deux énergies.

Calcul de l'énergie cible :

20 [0087] On estime l'énergie cible égale à l'énergie des derniers N échantillons mémorisés en période valide (N est typiquement < la longueur du signal utilisé pour l'analyse LPC).

[0088] L'énergie du signal synthétisé est le produit de l'énergie du bruit blanc par G^2 et G_{LPC} . On choisit G pour que cette énergie soit égale à l'énergie cible.

25 3. Contrôle de l'énergie du signal de synthèse

[0089] Comme pour les signaux de parole, sauf que la vitesse de diminution de l'énergie du signal de synthèse est beaucoup plus lente, et qu'elle ne dépend pas de période fondamentale (inexistante) :

30 [0090] L'énergie du signal de synthèse est contrôlée à l'aide d'un gain calculé et adapté échantillon par échantillon. Dans le cas où la période d'effacement est relativement longue, il est nécessaire de faire baisser progressivement l'énergie du signal de synthèse. La loi d'adaptation du gain peut être calculée en fonction de différents paramètres comme les valeurs d'énergies mémorisées avant l'effacement, et stationnarité locale du signal au moment de la coupure.

35 6. Evolution de la procédure de synthèse au cours du temps :

Comme pour les signaux de parole :

40 [0091] Dans le cas de périodes d'effacement relativement longues, on peut également faire évoluer les paramètres de synthèse. Si le système est couplé à un dispositif de détection d'activité vocale ou de signaux musicaux avec estimation des paramètres du bruit (tel [REC-G.723.1A], [SALAMI-2], [BENYASSINE]), il sera particulièrement intéressant de faire tendre les paramètres de génération du signal à reconstruire vers ceux du bruit estimé: en particulier au niveau de l'enveloppe spectrale (interpolation du filtre LPC avec celui du bruit estimé, les coefficients de l'interpolation évoluant au cours du temps jusqu'à obtention du filtre du bruit) et de l'énergie (niveau évoluant progressivement vers celui du bruit, par exemple par fenêtrage).

45 **6. REMARQUE GENERALE**

50 [0092] Comme on l'aura compris, la technique qui vient d'être décrite présente l'avantage d'être utilisable avec tout type de codeur ; en particulier elle permet de remédier aux problèmes des paquets de bits perdus pour les codeurs temporels ou par transformée, sur des signaux de parole et musique avec de bonnes performances : en effet dans la présente technique, les seuls signaux mémorisés lors des périodes où les données transmises sont valides sont les échantillons issus du décodeur, information qui est disponible quelle que soit la structure de codage utilisée.

55 **7. RÉFÉRENCES BIBLIOGRAPHIQUES**

[0093]

[AT&T] AT&T (D.A. KapiLOW, R.V. Cox) « A high quality low-complexity algorithm for frame erasure concealment

EP 1 316 087 B1

(FEC) with G.711 » , Delayed Contribution D.249 (WP 3/16), ITU, may 1999.

[ATAL] B.S. Atal et M.R. Schroeder. "Predictive coding of speech signal and subjectives error criteria". IEEE Trans. on Acoustics, Speech and Signal Processing, 27 :247-254, juin 1979.

5 [BENYASSINE] A. Benyassine, E. Shlomot et H.Y. Su. "ITU-T recommendation G.729 Annex B : A silence compression scheme for use with G.729 optimized for V.70 digital simultaneous voice and data applications". IEEE Communication Magazine, septembre 97, PP. 56-63.

10 [BRANDENBURG] K. H. Brandenburg et M. Bossi. "Overview of MPEG audio : current and future standards for low-bit-rate audio coding". Journal of Audio Eng. Soc., Vol.45-1/2, janvier/février 1997, PP.4-21.

[CHEN] J. H. Chen, R. V. Cox, Y. C. Lin, N. Jayant et M. J. Melchner. "A low-delay CELP coder for the CCITT 16 kb/s speech coding standard". IEEE Journal on Selected Areas on Communications, Vol.10-5, juin 1992, PP.830-849.

15 [CHEN-2] J. H. Chen, C.R. Watkins. "Linear prediction coefficient generation during frame erasure or packet loss". Brevet US5574825, EP0673018.

[CHEN-3] J. H. Chen, C.R. Watkins. "Linear prediction coefficient generation during frame erasure or packet loss". Brevet 884010.

20 [CHEN-4] J. H. Chen, C.R. Watkins. "Frame erasure or packet loss compensation method". Brevet US5550543, EP0707308.

25 [CHEN-5] J. H. Chen. "Excitation signal synthesis during frame erasure or packet loss". Brevet US5615298, EP0673017.

[CHEN-6] J. H. Chen. "Computational complexity reduction during frame erasure of packet loss". Brevet US5717822.

30 [CHEN-7] J. H. Chen. "Computational complexity reduction during frame erasure or packet loss". Brevet US940212435, EP0673015.

[COX] R. V. Cox. "Three new speech coders from the ITU cover a range of applications". IEEE Communication Magazine, septembre 97, PP. 40-47.

35 [COX-2] R. V. Cox. "An improved frame erasure concealment method for ITU-T Rec. G728". Delayed contribution D.107 (WP 3/16), ITU-T, janvier 1998.

[COMBESURE] P.Combesure, J. Schnitzler, K. Fischer, R. Kirchherr, C. Lamblin, A. Le Guyader, D. Massaloux, C. Quinquis, J. Stegmann, P. Vary. "A 16,24,32 kbit/s Wideband Speech Codec Based on ATCELP". Proc. of ICASSP conference, 1998.

[DAUMER] W. R. Daumer, P. Mermelstein, X. Maître et I. Tokizawa. "Overview of the ADPCM coding algorithm". Proc. of GLOBECOM 1984, PP.23.1.1-23.1.4.

45 [ERDÖL]. N. Erdöl, C. Castelluccia, A. Zilouchian "Recovery of Missing Speech Packets Using the Short-Time Energy and Zero-Crossing Measurements" IEEE Trans. on Speech and Audio Processing, Vol.1-3, juillet 1993, PP. 295-303.

50 [FINGSCHEIDT] T. Fingscheidt, P. Vary, "Robust speech decoding: a universal approach to bit error concealment", Proc. of ICASSP conference, 1997, pp. 1667-1670.

[GOODMAN] D.J. Goodman, G.B. Lockhart, O.J. Wasem, W.C. Wong. "Waveform Substitution Techniques for Recovering Missing Speech Segments in Packet Voice Communications". IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-34, décembre 1986, PP. 1440-1448.

55 [GSM-FR] Recommendation GSM 06.11. "Substitution and muting of lost frames for full rate speech traffic channels". ETSI/TC SMG, ver. : 3.0.1. , février 1992.

[HARDWICK] J. C. Hardwick et J. S. Lim. "The application of the IMBE speech coder to mobile communications". Proc. of ICASSP conference, 1991, PP.249-252.

5 [HELLWIG] K. Hellwig, P. Vary, D. Massaloux, J. P. Petit, C. Galand et M. Rosso. "Speech codec for the European mobile radio system". GLOBECOM conference, 1989, PP. 1065-1069.

[HONKANEN] T. Honkanen, J. Vainio, P. Kapanen, P. Haavisto, R. Salami, C. Laflamme et J. P. Adoul. "GSM enhanced full rate speech codec ". Proc. of ICASSP conference, 1997, PP.771-774.

10 [KROON] P. Kroon, B.S. Atal. "On the use of pitch predictors with high temporal resolution". IEEE Trans. on Signal Processing, Vol.39-3, mars.1991, PP.733-735.

[KROON-2] P. Kroon. "Linear prediction coefficient generation during frame erasure or packet loss". Brevet US5450449, EP0673016.

15 [MAHIEUX] Y. Mahieux, J. P. Petit. "High quality audio transform coding at 64 kbit/s". IEEE Trans. on Com., Vol. 42-11, nov.1994, PP.3010-3019.

[MAHIEUX-2] Y. Mahieux, "Dissimulation erreurs de transmission", brevet 92 06720 déposé le 3 juin 1992.

20 [MAITRE] X. Maitre. "7 kHz audio coding within 64 kbit/s". IEEE Journal on Selected Areas on Communications, Vol.6-2, février 1988, PP.283-298.

[PARIKH] V.N. Parikh, J.H. Chen, G. Aguilar. "Frame Erasure Concealment Using Sinusoidal Analysis-Synthesis and Its Application to MDCT-Based Codecs". Proc. of ICASSP conference, 2000.

[PICTEL] PictureTel Corporation, "Detailed Description of the PTC (PictureTel Transform Coder), Contribution ITU-T, SG15/WP2/Q6, 8-9 Octobre 1996 Baltimore meeting, TD7

30 [RABINER] L.R. Rabiner, R.W. Schafer. "Digital processing of speech signals". Bell Laboratories inc., 1978.

[REC G.723.1A] ITU-T Annex A to recommendation G.723.1 "Silence compression scheme for dual rate speech coder for multimedia communications transmitting at 5.3 & 6.3 kbit/s"

35 [SALAMI] R. Salami, C. Laflamme, J. P. Adoul, A. Kataoka, S. Hayashi, T. Moriya, C. Lamblin, D. Massaloux, S. Proust, P. Kroon et Y. Shoham. "Design and description of CS-ACELP : a toll quality 8 kb/s speech coder". IEEE Trans. on Speech and Audio Processing, Vol.6-2, mars 1998, PP.116-130.

40 [SALAMI-2] R. Salami, C. Laflamme, J. P. Adoul. "ITU-T G.729 Annex A : Reduced complexity 8 kb/s CS-ACELP codec for digital simultaneous voice and data". IEEE Communication Magazine, septembre 97, PP. 56-63.

[TREMAIN] T. E. Tremain. "The government standard linear predictive coding algorithm : LPC 10". Speech technology, avril 1982, PP.40-49.

45 [WATKINS] C.R. Watkins, J.H. Chen. "Improving 16 kb/s G.728 LD-CELP Speech Coder for Frame Erasure Channels". Proc. of ICASSP conference, 1995, PP.241-244.

Revendications

50 1. Procédé de dissimulation d'erreur de transmission dans un signal audio-numérique dans lequel à la détection (3) d'échantillons manquants ou erronés dans un signal, on génère des échantillons de synthèse (5) à l'aide d'au moins un opérateur de prédiction à court terme et au moins pour les sons voisés un opérateur de prédiction à long terme estimé en fonction d'échantillons décodés d'un signal décodé passé, lesdits échantillons décodés étant mémorisés (6) précédemment lorsque les données transmises dudit signal passé sont valides, **caractérisé en ce qu'**on contrôle l'énergie du signal de synthèse ainsi généré à l'aide d'un gain calculé et adapté échantillon par échantillon selon une loi d'adaptation dépendant d'au moins un paramètre desdits échantillons décodés mémorisés.

2. Procédé selon la revendication 1, **caractérisé en ce que** le gain pour le contrôle du signal de synthèse est calculé en fonction d'au moins un des paramètres suivants : valeurs d'énergie préalablement mémorisées pour les échantillons correspondant à des données valides, période fondamentale pour les sons voisés, ou tout paramètre caractérisant le spectre de fréquences.
- 5
3. Procédé selon l'une des revendications précédentes, **caractérisé en ce que** le gain appliqué au signal de synthèse décroît progressivement en fonction de la durée pendant laquelle les échantillons de synthèse sont générés.
- 10
4. Procédé selon l'une des revendications précédentes, **caractérisé en ce qu'**on discrimine dans les données valides les sons stationnaires et les sons non stationnaires et on met en oeuvre des lois d'adaptation du gain permettant de contrôler le signal de synthèse différentes d'une part pour les échantillons générés à la suite de données valides correspondant à des sons stationnaires et d'autre part pour les échantillons générés à la suite de données valides correspondants à des sons non stationnaires.
- 15
5. Procédé selon l'une des revendications précédentes, **caractérisé en ce qu'**on met à jour en fonction des échantillons de synthèse générés le contenu de mémoires utilisées pour le traitement de décodage.
- 20
6. Procédé selon la revendication 5, **caractérisé en ce qu'**on met en oeuvre au moins partiellement sur les échantillons synthétisés un codage analogue à celui mis en oeuvre à l'émetteur suivi éventuellement d'une opération de décodage au moins partielle, les données obtenues servant à régénérer les mémoires du décodeur.
- 25
7. Procédé selon la revendication 6, **caractérisé en ce qu'**on régénère la première trame effacée au moyen de cette opération de codage-décodage, en exploitant le contenu des mémoires du décodeur avant la coupure, lorsque lesdites mémoires contiennent des informations exploitables dans cette opération.
- 30
8. Procédé selon l'une des revendications précédentes, **caractérisé en ce qu'**on génère en entrée de l'opérateur de prédiction à court terme un signal d'excitation qui, en zone voisée, est la somme d'une composante harmonique et d'une composante faiblement harmonique ou non harmonique, et en zone non voisée, limitée à une composante non harmonique.
- 35
9. Procédé selon la revendication 8, **caractérisé en ce que** la composante harmonique est obtenue en mettant en oeuvre un filtrage au moyen de l'opérateur de prédiction à long terme appliqué sur un signal résiduel calculé en mettant en oeuvre un filtrage à court terme inverse sur les échantillons mémorisés.
- 40
10. Procédé selon la revendication 9, **caractérisé en ce que** l'autre composante est déterminée à l'aide d'un opérateur de prédiction à long terme auquel on applique des perturbations pseudo-aléatoires.
- 45
11. Procédé selon l'une des revendications 8 à 10, **caractérisé en ce que** pour la génération d'un signal d'excitation voisé, la composante harmonique est limitée aux basses fréquences du spectre, tandis que l'autre composante est limitée aux hautes fréquences.
- 50
12. Procédé selon l'une des revendications précédentes, **caractérisé en ce que** l'opérateur de prédiction à long terme est déterminé à partir des échantillons de trames valides mémorisés, avec un nombre d'échantillons utilisés pour cette estimation variant entre une valeur minimale et une valeur égale à au moins deux fois la période fondamentale estimée pour le son voisé.
- 55
13. Procédé selon l'une des revendications précédentes, **caractérisé en ce que** le signal résiduel est traité de manière non linéaire pour éliminer des pics d'amplitude.
14. Procédé selon l'une des revendications précédentes, **caractérisé en ce qu'**on détecte l'activité vocale en estimant des paramètres de bruit et **en ce qu'**on fait tendre des paramètres du signal synthétisé vers ceux du bruit estimé.
15. Procédé selon la revendication 14, **caractérisé en ce qu'**on estime l'enveloppe spectrale du bruit des échantillons décodés valides et on génère un signal synthétisé évoluant vers un signal possédant la même enveloppe spectrale.
16. Procédé de traitement de signaux de sons, **caractérisé en ce qu'**on met en oeuvre une discrimination entre les sons voisés et les sons musicaux et lorsqu'on détecte des sons musicaux, on met en oeuvre un procédé selon l'une des revendications précédentes sans estimation d'un opérateur de prédiction à long terme.

17. Dispositif de dissimulation d'erreur de transmission dans un signal audio-numérique qui reçoit en entrée un signal décodé que lui transmet un décodeur et qui génère des échantillons manquants ou erronés dans ce signal décodé, **caractérisé en ce qu'il** comporte des moyens de traitement aptes à mettre en oeuvre le procédé selon l'une des revendications précédentes.

5
18. Système de transmission comportant au moins un codeur, au moins un canal de transmission, un module apte à détecter que des données transmises ont été perdues ou sont fortement erronées, au moins un décodeur et un dispositif de dissimulation d'erreurs qui reçoit le signal décodé, **caractérisé en ce que** ce dispositif de dissimulation d'erreurs est un dispositif selon la revendication 17.

10
Claims

15
1. Method of concealing transmission error in a digital audio signal in which upon detecting (3) samples that are missing or erroneous in a signal, synthesis samples (5) are generated with the aid of at least one short-term prediction operator and at least for the voiced sounds a long-term prediction operator estimated as a function of decoded samples of a past decoded signal, said decoded samples being stored (6) previously when the transmitted data of said past signal are valid, **characterized in that** the energy of the synthesis signal thus generated is controlled with the aid of a gain that is calculated and adapted sample by sample according to an adaptation law dependent on at least one parameter of said stored decoded samples.

20
2. Method according to Claim 1, **characterized in that** the gain for controlling the synthesis signal is calculated as a function of at least one of the following parameters: energy values previously stored for the samples corresponding to valid data, pitch period for the voiced sounds, or any parameter characterizing the frequency spectrum.

25
3. Method according to one of the preceding claims, **characterized in that** the gain applied to the synthesis signal decreases progressively as a function of the duration for which the synthesis samples are generated.

30
4. Method according to one of the preceding claims, **characterized in that** the steady sounds and the non-steady sounds are discriminated in the valid data, and gain adaptation laws are implemented for controlling the synthesis signal that differ on the one hand for the samples generated following valid data corresponding to steady sounds and on the other hand for the samplers generated following valid data corresponding to non-steady sounds.

35
5. Method according to one of the preceding claims, **characterized in that** the content of memories that are used for the decoding processing is updated as a function of the synthesis samples generated.

40
6. Method according to Claim 5, **characterized in that** a coding analogous to that implemented at the transmitter is implemented at least partially on the synthesized samples, optionally followed by an at least partial decoding operation, the data obtained serving to regenerate the memories of the decoder.

45
7. Method according to Claim 6, **characterized in that** the first erased frame is regenerated by means of this coding-decoding operation, utilizing the content of the memories of the decoder prior to the interruption, when said memories contain information that can be utilized in this operation.

50
8. Method according to one of the preceding claims, **characterized in that** an excitation signal is generated as input to the short-term prediction operator, which signal, in a voiced zone, is the sum of a harmonic component and a weakly harmonic or non-harmonic component, and in a non-voiced zone, limited to a non-harmonic component.

55
9. Method according to Claim 8, **characterized in that** the harmonic component is obtained by implementing a filtering by means of the long-term prediction operator applied to a residual signal calculated by implementing an inverse short-term filtering on the stored samples.

10
10. Method according to Claim 9, **characterized in that** the other component is determined with the aid of a long-term prediction operator to which pseudo-ransom disturbances are applied.

11. Method according to one of Claims 8 to 10, **characterized in that** in order to generate a voiced excitation signal, the harmonic component is limited to the low frequencies of the spectrum, while the other component is limited to the high frequencies.

12. Method according to one of the preceding claims, **characterized in that** the long-term prediction operator is determined from the stored valid frame samples, with a number of samples used for this estimation varying between a minimum value and a value that is equal to at least twice the pitch period estimated for the voiced sound.
- 5 13. Method according to one of the preceding claims, **characterized in that** the residual signal is processed in a non-linear manner in order to eliminate amplitude peaks.
- 10 14. Method according to one of the preceding claims, **characterized in that** voice activity is detected by estimating noise parameters and **in that** parameters of the synthesized signal are made to tend towards those of the estimated noise.
- 15 15. Method according to Claim 14, **characterized in that** the spectral envelope of the noise of the valid decoded samples is estimated and a synthesized signal that evolves towards a signal possessing the same spectral envelope is generated.
- 20 16. Method of processing sound signals, **characterized in that** a discrimination is implemented between the voiced sounds and the musical sounds, and when ;musical sounds are detected, a method is implemented according to one of the preceding claims without estimating a long-term prediction operator.
- 25 17. Device for concealing transmission error in a digital audio signal which receives as input a decoded signal transmitted to it by a decoder and which generates samples that are missing or erroneous in this decoded signal, **characterized in that** it comprises processing means suitable for implementing the method according to one of the preceding claims.
- 30 18. Transmission system comprising at least one coder, at least one transmission channel, a module suitable for detecting that transmitted data have been lost or are highly erroneous, at least one decoder and a device for concealing errors which receives the decoded signal, **characterized in that** this device for concealing errors is a device according to Claim 17.

30 Patentansprüche

- 35 1. Verfahren zur Verdeckung eines Übertragungsfehlers in einem digitalen Audiosignal, bei dem bei Erfassung (3) von fehlenden oder fehlerhaften Proben in einem Signal Syntheseproben (5) mit Hilfe mindestens eines Kurzzeitvorhersage-Operators und mindestens für die stimmhaften Töne eines Langzeitvorhersage-Operators, geschätzt in Abhängigkeit von decodierten Proben eines vergangenen decodierten Signals, erzeugt, wobei die decodierten Proben vorher gespeichert werden (6), wenn die übertragenen Daten des vergangenen Signals gültig sind, **dadurch gekennzeichnet, dass** die Energie des so erzeugten Synthesesignals mit Hilfe einer berechneten und angepassten Verstärkung Probe für Probe gemäß einem Anpassungsgesetz gesteuert wird, das von mindestens einem Parameter der decodierten gespeicherten Proben abhängt.
- 40 2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** die Verstärkung für die Steuerung des Synthesesignals in Abhängigkeit von mindestens einem der folgende Parameter berechnet wird: vorher gespeicherte Energiewerte für die gültigen Daten entsprechenden Proben, Grundperiode für die stimmhaften Töne, oder jeder das Frequenzspektrum kennzeichnende Parameter.
- 45 3. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die an das Synthesesignal angewendete Verstärkung progressiv in Abhängigkeit von der Dauer abnimmt, während der die Syntheseproben erzeugt werden.
- 50 4. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** in den gültigen Daten die stationären Töne und die nicht-stationären Töne unterschieden werden und dass Anpassungsgesetze der Verstärkung, die es ermöglichen, das Synthesesignal zu steuern, angewendet werden, die einerseits für die Proben, die nach stationären Tönen entsprechenden gültigen Daten erzeugt werden, und andererseits für die Proben, die nach nicht-stationären Tönen entsprechenden gültigen Daten erzeugt werden, unterschiedlich sind.
- 55 5. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** in Abhängigkeit von den erzeugten Syntheseproben der Inhalt von Speichern aktualisiert wird, die für die Decodierungsverarbeitung verwendet werden.

- 5
6. Verfahren nach Anspruch 5, **dadurch gekennzeichnet, dass** mindestens teilweise an die synthetisierten Proben ein Codierung analog zu derjenigen angewendet wird, die an den Emitter angewendet wird, gefolgt ggf. von einem mindestens teilweisen Decodierungsvorgang, wobei die erhaltenen Daten dazu dienen, die Speicher des Decodierers zu regenerieren.
- 10
7. Verfahren nach Anspruch 6, **dadurch gekennzeichnet, dass** der erste gelöscht Rahmen mittels dieses Codier-Decodier-Vorgangs regeneriert wird, indem der Inhalt der Speicher des Decodierers vor dem Abschalten ausgewertet wird, wenn die Speicher in diesem Vorgang auswertbare Informationen enthalten.
- 15
8. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** am Eingang des Kurzzeitvorhersage-Operators ein Anregungssignal erzeugt wird, das in der stimmhaften Zone die Summe einer harmonischen Komponente und einer gering harmonischen oder nicht-harmonischen Komponente ist und in der nicht-stimmhaften Zone auf eine nicht-harmonische Komponente beschränkt ist.
- 20
9. Verfahren nach Anspruch 8, **dadurch gekennzeichnet, dass** die harmonische Komponente durch Anwenden einer Filterung mittels des Langzeitvorhersage-Operators erhalten wird, die an ein Restsignal angewendet wird, das unter Anwendung einer umgekehrten Kurzzeit-Filterung an die gespeicherten Proben berechnet wird.
- 25
10. Verfahren nach Anspruch 9, **dadurch gekennzeichnet, dass** die andere Komponente mit Hilfe eines Langzeitvorhersage-Operators bestimmt wird, an den pseudo-zufällige Störungen angelegt werden.
- 30
11. Verfahren nach einem der Ansprüche 8 bis 10, **dadurch gekennzeichnet, dass** zur Erzeugung eines stimmhaften Anregungssignals die harmonische Komponente auf die niederen Frequenzen des Spektrums begrenzt ist, während die andere Komponente auf die hohen Frequenzen begrenzt ist.
- 35
12. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** der Langzeitvorhersage-Operator ausgehend von den gültigen gespeicherten Rahmenproben bestimmt wird, mit einer Anzahl von für diese Schätzung verwendeten Proben, die zwischen einem minimalen Wert und einem Wert gleich mindestens der doppelten Grundperiode variiert, die für den stimmhaften Ton geschätzt wird.
- 40
13. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** das Restsignal nichtlinear verarbeitet wird, um Amplitudenspitzen zu unterdrücken.
- 45
14. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Stimmaktivität erfasst wird, indem Rauschparameter geschätzt werden und indem Parameter des synthetisierten Signals zu denjenigen des geschätzten Rauschens ausgeweitet werden.
- 50
15. Verfahren nach Anspruch 14, **dadurch gekennzeichnet, dass** die Spektralhülle des Rauschens der gültigen decodierten Proben geschätzt wird, und dass ein synthetisiertes Signal erzeugt wird, das sich zu einem Signal entwickelt, das die gleiche Spektralhülle besitzt.
- 55
16. Verfahren zur Verarbeitung von Tonsignalen, **dadurch gekennzeichnet, dass** eine Unterscheidung zwischen den stimmhaften Tönen und den musikalischen Tönen angewendet wird, und dass, wenn musikalische Töne erfasst werden, ein Verfahren nach einem der vorhergehenden Ansprüche ohne Schätzung eines Langzeitvorhersage-Operators angewendet wird.
17. Vorrichtung zur Übertragungsfehler-Verdeckung in einem digitalen Audiosignal, die am Eingang ein decodiertes Signal empfängt, das ihr ein Dekodierer überträgt, und die fehlende oder fehlerhaft Proben in diesen decodierten Signal erzeugt, **dadurch gekennzeichnet, dass** sie Verarbeitungsmittel aufweist, die in der Lage sind, das Verfahren nach einem der vorhergehenden Ansprüche anzuwenden.
18. Übertragungssystem, das mindestens einen Codierer, mindestens einen Übertragungskanal, ein Modul, das erfassen kann, ob übertragene Daten verloren wurden oder stark fehlerhaft sind, mindestens einen Dekodierer und eine Fehlerverdeckungs Vorrichtung aufweist, die das decodierte Signal empfängt, **dadurch gekennzeichnet, dass** diese Fehlerverdeckungs Vorrichtung eine Vorrichtung gemäß Anspruch 17 ist.

FIG. 1

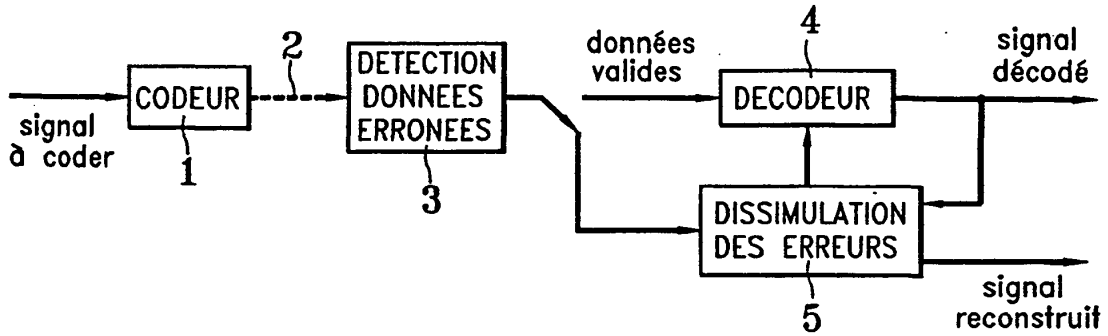


FIG. 2

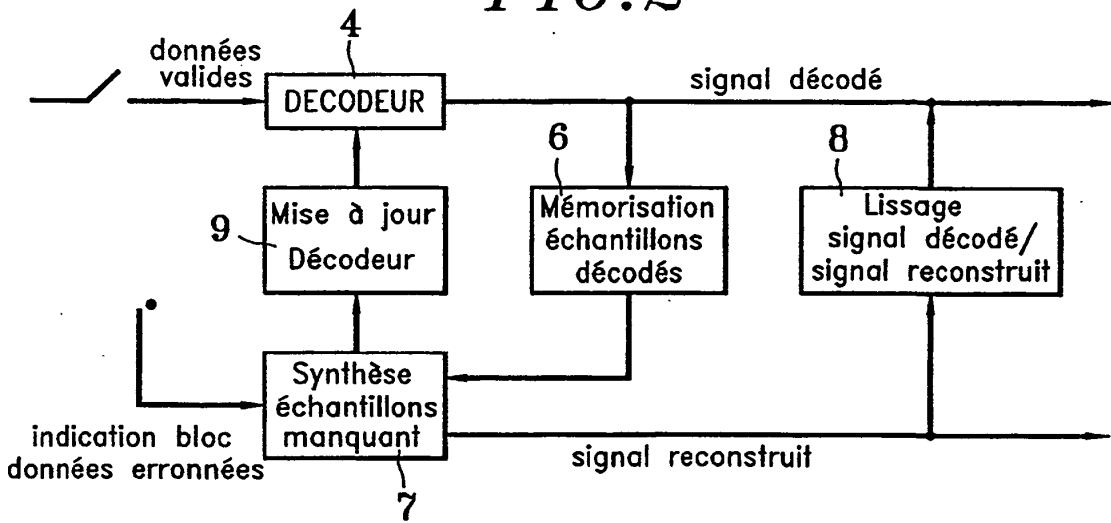


FIG. 3

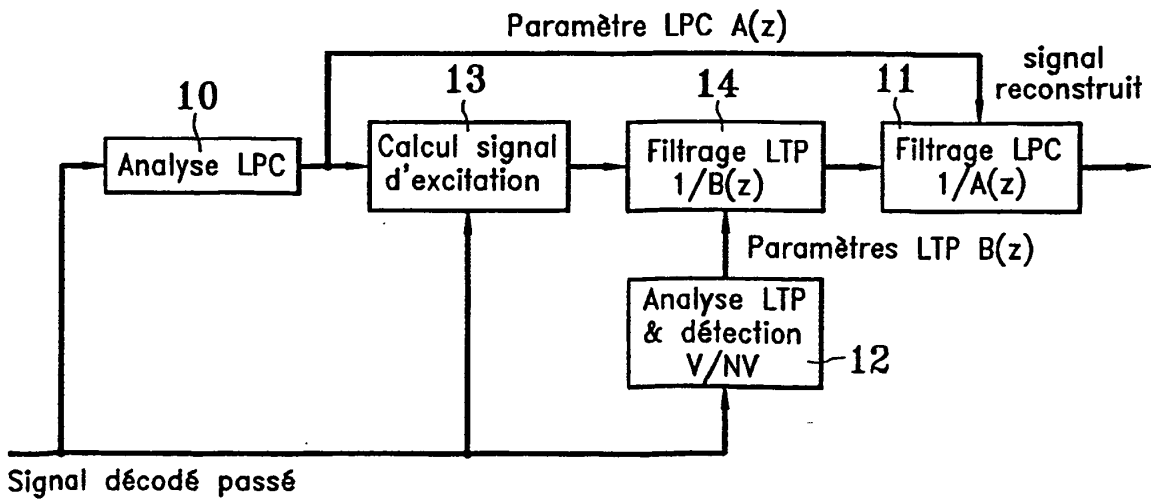


FIG. 4

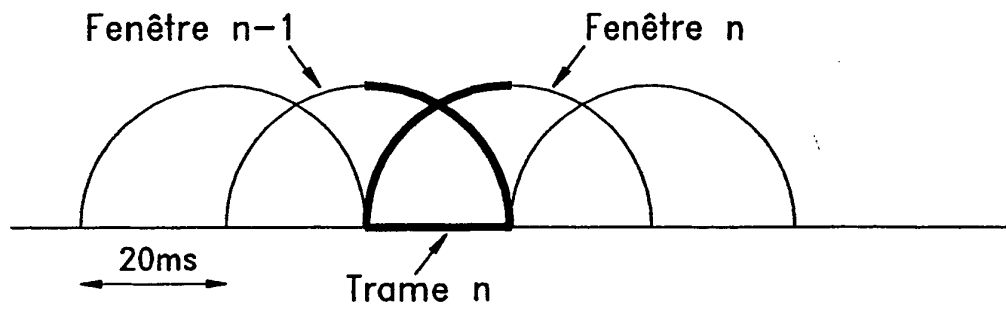


FIG. 5

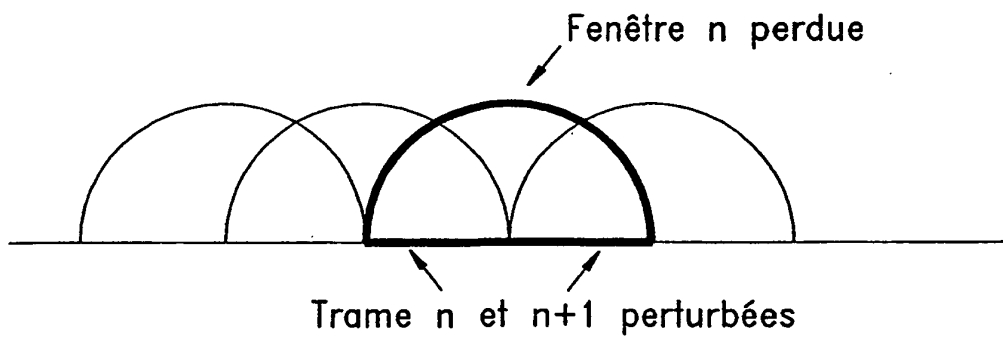


FIG. 6

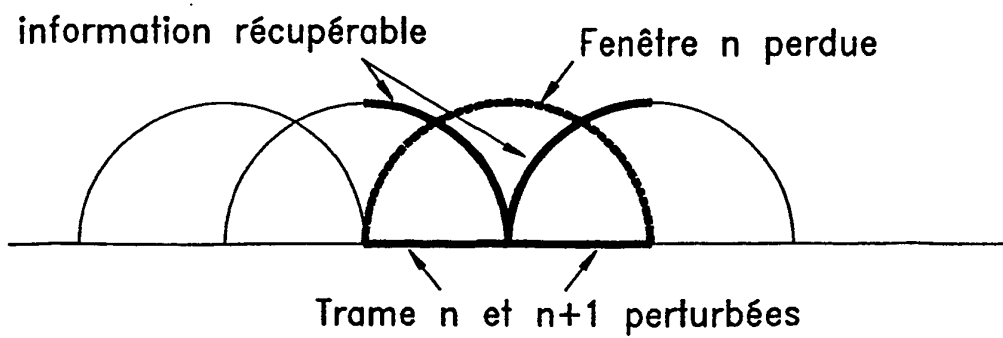


FIG. 7

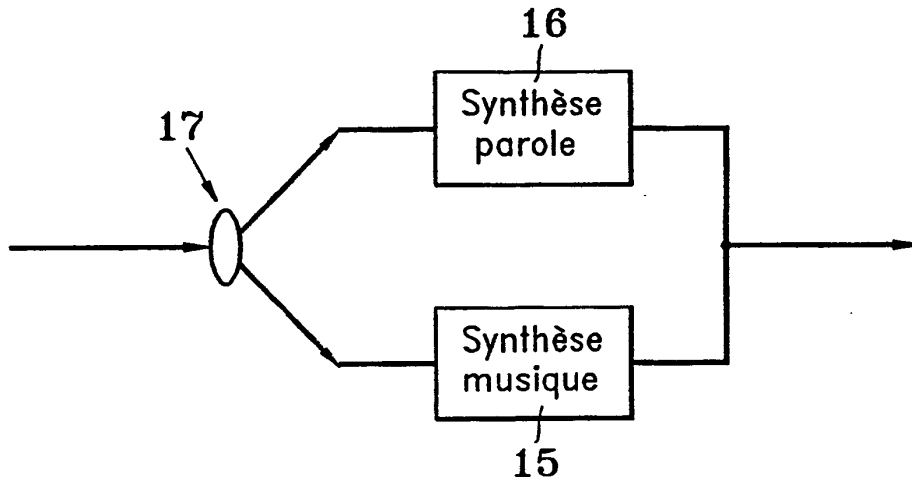
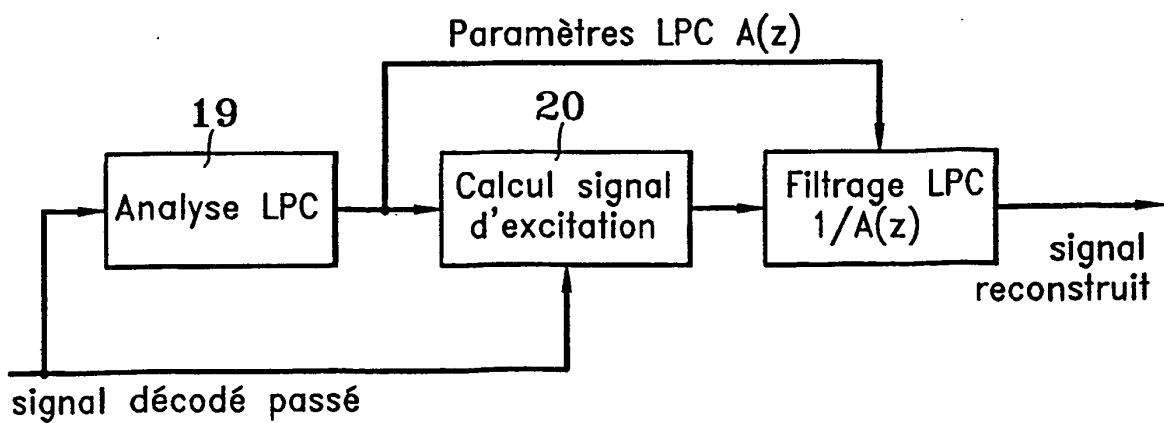


FIG. 8



RÉFÉRENCES CITÉES DANS LA DESCRIPTION

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

Documents brevets cités dans la description

- US 5884010 A [0016]
- US 5574825 A [0093]
- EP 0673018 A [0093]
- WO 884010 A [0093]
- US 5550543 A [0093]
- EP 0707308 A [0093]
- US 5615298 A [0093]
- EP 0673017 A [0093]
- US 5717822 A [0093]
- US 940212435 B [0093]
- EP 0673015 A [0093]
- US 5450449 A [0093]
- EP 0673016 A [0093]
- WO 9206720 A [0093]

Littérature non-brevet citée dans la description

- **B.S. ATAL ; M.R. SCHROEDER.** Predictive coding of speech signal and subjective error criteria. *IEEE Trans. on Acoustics, Speech and Signal Processing*, Juin 1979, vol. 27, 247-254 [0093]
- **A. BENYASSINE ; E. SHLOMOT ; H.Y. SU.** ITU-T recommendation G.729 Annex B : A silence compression scheme for use with G.729 optimized for V.70 digital simultaneous voice and data applications. *IEEE Communication Magazine*, Septembre 1997, 56-63 [0093]
- **K. H. BRANDENBURG ; M. BOSSI.** Overview of MPEG audio : current and future standards for low-bit-rate audio coding. *Journal of Audio Eng. Soc.*, Janvier 1997, vol. 45-1/2, 4-21 [0093]
- **J. H. CHEN ; R. V. COX ; Y. C. LIN ; N. JAYANT ; M. J. MELCHNER.** A low-delay CELP coder for the CCITT 16 kb/s speech coding standard. *IEEE Journal on Selected Areas on Communications*, Juin 1992, vol. 10-5, 830-849 [0093]
- **J. H. CHEN ; C.R. WATKINS.** Linear prediction coefficient generation during frame erasure or packet loss [0093] [0093]
- **J. H. CHEN ; C.R. WATKINS.** Frame erasure or packet loss compensation method [0093]
- **J. H. CHEN.** Excitation signal synthesis during frame erasure or packet loss [0093]
- **J. H. CHEN.** Computational complexity reduction during frame erasure of packet loss [0093]
- **J. H. CHEN.** Computational complexity reduction during frame erasure or packet loss [0093]
- **R. V. COX.** Three new speech coders from the ITU cover a range of applications. *IEEE Communication Magazine*, Septembre 1997, 40-47 [0093]
- **R. V. COX.** An improved frame erasure concealment method for ITU-T Rec. G728. *Delayed contribution D.107 (WP 3/16), ITU-T*, Janvier 1998 [0093]
- **P.COMBESCURE ; J. SCHNITZLER ; K. FICHER ; R. KIRCHHERR ; C. LAMBLIN ; A. LE GUYADER ; D. MASSALOUX ; C. QUINQUIS ; J. STEGMANN ; P. VARY.** A 16,24,32 kbit/s Wideband Speech Codec Based on ATCELP. *Proc. of ICASSP conference*, 1998 [0093]
- **W. R. DAUMER ; P. MERMELSTEIN ; X. MAÎTRE ; I. TOKIZAWA.** Overview of the ADPCM coding algorithm. *Proc. of GLOBECOM*, 1984, 23.1.1-23.1.4 [0093]
- **N. ERDÖL ; C. CASTELLUCCIA ; A. ZILOUCHIAN.** Recovery of Missing Speech Packets Using the Short-Time Energy and Zero-Crossing Measurements. *IEEE Trans. on Speech and Audio Processing*, Juillet 1993, vol. 1-3, 295-303 [0093]
- **T. FINGSCHEIDT ; P. VARY.** Robust speech decoding: a universal approach to bit error concealment. *Proc. of ICASSP conference*, 1997, 1667-1670 [0093]
- **D.J. GOODMAN ; G.B. LOCKHART ; O.J. WASEM ; W.C. WONG.** Waveform Substitution Techniques for Recovering Missing Speech Segments in Packet Voice Communications. *IEEE Trans. on Acoustics, Speech and Signal Processing*, Décembre 1986, vol. ASSP-34, 1440-1448 [0093]
- Substitution and muting of lost frames for full rate speech traffic channels. *ETSI/TC SMG*, Février 1992 [0093]
- **J. C. HARDWICK ; J. S. LIM.** The application of the IMBE speech coder to mobile communications. *Proc. of ICASSP conference*, 1991, 249-252 [0093]
- **K. HELLWIG ; P. VARY ; D. MASSALOUX ; J. P. PETIT ; C. GALAND ; M. ROSSO.** Speech codec for the European mobile radio system. *GLOBECOM conference*, 1989, 1065-1069 [0093]

- **T. HONKANEN ; J. VAINIO ; P. KAPANEN ; P. HAAVISTO ; R. SALAMI ; C. LAFLAMME ; J. P. ADOUL.** GSM enhanced full rate speech codec. *Proc. of ICASSP conference*, 1997, 771-774 [0093]
- **P. KROON ; B.S. ATAL.** On the use of pitch predictors with high temporal resolution. *IEEE Trans. on Signal Processing*, Mars 1991, vol. 39-3, 733-735 [0093]
- **P. KROON.** Linear prediction coefficient generation during frame erasure or packet loss [0093]
- **Y. MAHIEUX ; J. P. PETIT.** High quality audio transform coding at 64 kbit/s. *IEEE Trans. on Com.*, Novembre 1994, vol. 42-11, 3010-3019 [0093]
- **Y. MAHIEUX.** Dissimulation erreurs de transmission [0093]
- **X. MAITRE.** 7 kHz audio coding within 64 kbit/s. *IEEE Journal on Selected Areas on Communications*, Février 1988, vol. 6-2, 283-298 [0093]
- **V.N. PARIKH ; J.H. CHEN ; G. AGUILAR.** Frame Erasure Concealment Using Sinusoidal Analysis-Synthesis and Its Application to MDCT-Based Coders. *Proc. of ICASSP conference*, 2000 [0093]
- Detailed Description of the PTC (PictureTel Transform Coder. *Contribution ITU-T, SG15/WP2/Q6*, 08 Octobre 1996 [0093]
- **L.R. RABINER ; R.W. SCHAFER.** Digital processing of speech signals. Bell Laboratories inc, 1978 [0093]
- **R. SALAMI ; C. LAFLAMME ; J. P. ADOUL ; A. KATAOKA ; S. HAYASHI ; T. MORIYA ; C. LAMBLIN ; D. MASSALOUX ; S. PROUST ; P. KROON.** Design and description of CS-ACELP : a toll quality 8 kb/s speech coder. *IEEE Trans. on Speech and Audio Processing*, Mars 1998, vol. 6-2, 116-130 [0093]
- **R. SALAMI ; C. LAFLAMME ; J. P. ADOUL.** ITU-T G.729 Annex A: Reduced complexity 8 kb/s CS-ACELP codec for digital simultaneous voice and data. *IEEE Communication Magazine*, Septembre 1997, 56-63 [0093]
- **T. E. TREMAIN.** The government standard linear predictive coding algorithm : LPC 10. *Speech technology*, Avril 1982, 40-49 [0093]
- **C.R. WATKINS ; J.H. CHEN.** Improving 16 kb/s G.728 LD-CELP Speech Coder for Frame Erasure Channels. *Proc. of ICASSP conference*, 1995, 241-244 [0093]