

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-508597  
(P2004-508597A)

(43) 公表日 平成16年3月18日(2004.3.18)

(51) Int. Cl. <sup>7</sup>	F I	テーマコード (参考)
G 1 0 L 13/00	G 1 0 L 3/00	5 D 0 4 5
G 1 0 L 19/00	G 1 0 L 3/00	

審査請求 未請求 予備審査請求 有 (全 75 頁)

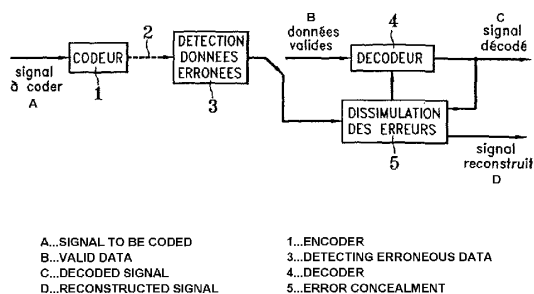
(21) 出願番号	特願2002-525647 (P2002-525647)	(71) 出願人	591034154 フランス テレコム FRANCE TELECOM フランス国、75015 パリ、プラス・ ダルレ、6
(86) (22) 出願日	平成13年9月5日 (2001.9.5)	(74) 代理人	100080447 弁理士 太田 恵一
(85) 翻訳文提出日	平成15年3月4日 (2003.3.4)	(72) 発明者	コヴェジ, バラズ フランス共和国、エフ-22300 ラニ ヨン, リュ コルレイ 12
(86) 国際出願番号	PCT/FR2001/002747	(72) 発明者	マッサルー, ドミニク フランス共和国、エフ-22700 ペロ ス-ギレク, リュ デュ プレ ドゥ サ ン-モール, 53
(87) 国際公開番号	W02002/021515		
(87) 国際公開日	平成14年3月14日 (2002.3.14)		
(31) 優先権主張番号	00/11285		
(32) 優先日	平成12年9月5日 (2000.9.5)		
(33) 優先権主張国	フランス (FR)		

最終頁に続く

(54) 【発明の名称】 オーディオ信号における伝送エラーの抑止シミュレーション

(57) 【要約】

伝送後に復号化された信号を受信し、伝送されたデータが健全な場合には、復号化されたサンプルを記憶し、短期予測演算子を少なくとも一つと、一つの長期予測演算子とを、サンプルに応じて算定し、そして、復号化された信号において、欠損しているかエラーを含みうるサンプルを、そのようにして算定した演算子によって行うものであり、オーディオ・デジタル信号の中の伝送エラーを抑止シミュレーションする、そのようにして生成された合成信号のエネルギー制御を、サンプルごとに計算され適合理化されたゲインを用いて制御するという特徴とする、オーディオ・デジタル信号の中の伝送エラーの抑止シミュレーションする方法。



## 【特許請求の範囲】

## 【請求項 1】

伝送後に復号化された信号を受信し、伝送されたデータが健全な場合には、復号化されたサンプルを記憶し、短期予測演算子を少なくとも一つと、少なくとも有声音の長期予測演算子一つとを、記憶された健全なサンプルに応じて算定し、そして、復号化された信号において、欠損しているかエラーを含みうるサンプルを、そのようにして算定した演算子によって生成する、オーディオ・デジタル信号における伝送エラーを抑止シミュレーションする方法であり、その特徴は、そのようにして生成された合成信号のエネルギー制御を、サンプルごとに計算され、そして適合化されたゲインを用いて制御するという特徴とする、オーディオ・デジタル信号における伝送エラーの抑止シミュレーション方法。

10

## 【請求項 2】

合成信号を制御するためのゲインは、健全なデータに対応するサンプルのために前もって記憶されたエネルギーの値、有声音のための基本波周期、あるいは周波数のスペクトルを特徴づけるあらゆるパラメータというような、パラメータの少なくとも一つに応じて計算することを特徴とする、請求項 1 に記載の方法。

## 【請求項 3】

合成信号に適用されるゲインは、合成サンプルが生成される持続時間に応じて、徐々に減少していくことを特徴とする、請求項 1 あるいは請求項 2 に記載の方法。

## 【請求項 4】

健全なデータにおいて、定常性音と非定常性音とを区別し、そして、異なる合成信号の制御を可能にするゲイン適合化法則を、一方では定常性音に対応する健全なデータの後に生成されるサンプルのために、他方では非定常性音に対応する健全なデータの後に生成されるサンプルのために用いることを特徴とする、請求項 1 から請求項 3 のいずれか一つに記載の方法。

20

## 【請求項 5】

復号化処理のために用いられるメモリの内容を、生成される合成サンプルに応じて更新することを特徴とする、請求項 1 から請求項 4 のいずれか一つに記載の方法。

## 【請求項 6】

少なくとも部分的には行われる復号化の作業に続いて、送信器で活用されるかもしれないものと類似の符号化を合成されたサンプルに少なくとも部分的に用い、その際には、そこで得られるデータが復号器のメモリを再生するのに役立つことを特徴とする、請求項 5 に記載の方法。

30

## 【請求項 7】

その符号化 - 復号化作業により、消失された最初のフレームを再生するのに際し、切断の前に復号器のメモリにある情報がその作業に利用可能な場合には、そのメモリの内容を利用することを特徴とする、請求項 6 に記載の方法。

## 【請求項 8】

短期予測演算子の入力において生成される励振信号は、有声音の区域においては、高調波成分の強い一つの成分と高調波成分の弱いまたは非高調波成分的なもう一つの成分との和であり、非有声音の区域においては、非高調波成分的な成分に限定される信号であることを特徴とする、請求項 1 から請求項 7 のいずれか一つに記載の方法。

40

## 【請求項 9】

高調波成分は、記憶されたサンプルに短期逆フィルタリングを用いて計算された残留信号に長期予測演算子を適用することによるフィルタリングを用いることにより得られることを特徴とする、請求項 8 に記載の方法。

## 【請求項 10】

もう一つの成分は、長期予測演算子に疑似ランダムな擾乱を加えることにより、決定されることを特徴とする、請求項 9 に記載の方法。

## 【請求項 11】

有声音の励振信号を生成するについては、高調波成分はスペクトルの周波数の低いものに限

50

定し、その一方で、もう一つの成分は高い周波数に限定することを特徴とする、請求項 8 から請求項 10 のいずれか一つに記載の方法。

【請求項 12】

長期予測演算子の決定は、記憶された健全なフレームのサンプルに基づいて行われ、この算定のために使用するサンプルの個数は、最小値に始まって、その有声音のために算定された基本波周期の少なくとも二倍に等しい値に至るまでの間を変化する個数であることを特徴とする、請求項 1 から請求項 11 のいずれか一つに記載の方法。

【請求項 13】

残留信号は非線形的に処理され、それにより、振幅のピークを除去することを特徴とする、請求項 1 から請求項 12 のいずれか一つに記載の方法。

10

【請求項 14】

ノイズのパラメータを算定して発声活性を検出すること、そして、合成された信号のパラメータを算定されたノイズのパラメータのものに近づけることを特徴とする、請求項 1 から請求項 13 のいずれか一つに記載の方法。

【請求項 15】

復号化された健全なサンプルのノイズのスペクトル包絡線を算定し、同じスペクトル包絡線を有する信号に向かって進展する合成された信号を生成することを特徴とする、請求項 14 に記載の方法。

【請求項 16】

有声音と楽音との間の区別を実施し、楽音が検出される場合には、長期予測演算子を算定することなく請求項 1 から 15 のいずれか一つの方法を用いることを特徴とする、音響信号処理方法。

20

【請求項 17】

復号器から伝送された復号化された信号を入力で受信し、この復号化された信号の中に欠けているサンプルまたはエラーのあるサンプルを生成する、デジタル・オーディオ信号の中の伝送エラーを抑止シミュレーションする方法で、請求項 1 から請求項 16 にのいずれか一つに記載の方法を用いるのに適した処理手段からなることを特徴とする、伝送エラー抑止シミュレーション装置。

【請求項 18】

少なくとも一つの符号器と、少なくとも一つの伝送路と、伝送されてきたデータが失われた、あるいはエラーの多いことを検出するに適したモジュールと、少なくとも一つの復号器と、その復号化された信号を受信するエラー抑止シミュレーション装置とからなる伝送システムで、そのエラー抑止シミュレーション装置が請求項 17 に記載の装置であることを特徴とする、伝送システム。

30

【発明の詳細な説明】

【0001】

1. 技術分野

本発明は、言葉及び/または音の信号のあらゆるタイプのデジタル符号化方法を用いる伝送システムにおいて、続発する伝送エラーを抑止シミュレーションする技術に関するものである。

40

【0002】

従来、符号化器には、大きく分けて、次の二つのカテゴリーのものがあった。

- ・いわゆる時間的といわれる符号化器で、サンプルごとにデジタル信号のサンプルの圧縮を行うもの（例えば符号化器 M I C または M I C D A [ D A U M E R ] [ M A I T R E ] の場合）。

- ・そしてパラメータ式の符号化器で、符号化すべき信号のサンプルの連続するフレームを分析し、それにより、それらのフレームのそれぞれで、ある一定数のパラメータを抽出し、つぎにその抽出したパラメータを符号化して伝送するというもの（音声合成機 [ T R E M A I N ]、I M B E 符号化器 [ H A R D W I C K ]、または変換値を用いる符号化器 [ B R A N D E N B U R G ] ) の場合）。

50

## 【 0 0 0 3 】

残留時間の波形を符号化することによるパラメータ式の符号化器を表すパラメータの符号化を補完する中間的カテゴリが存在する。単純にするために、これらの符号化器をパラメータ式の符号化器に含めてもよい。

## 【 0 0 0 4 】

このカテゴリに含まれるものとしては、予測符号化器があり、また、例えば、RPE-LTP ( [ H E L L W I G ] ) または CELP ( [ A T A L ] ) のような合成による分析式符号化器に分類されるものが幾つかある。

## 【 0 0 0 5 】

これらの符号化器のすべてについて、符号化される数値は、つぎに二進法列に変換され、それを伝送路にのせて伝送することになる。この伝送路の質と搬送のタイプによって、いくらかの擾乱が伝送される信号に影響を与え、復号器が受信する二進列でいくつかのエラーを発生させることになりかねない。これらのようなエラーが二進列に割り込むのは孤立した形になる可能性があるが、非常に多くの場合、一斉に発生する。そのような場合にこそ、信号の一つの部分に丸ごと対応する一パケット分のビットが、エラーを含んだり、あるいは受信されなかったりすることになるのである。この種の問題が発生するのは、例えば携帯電話のネットワークで伝送を行う場合である。この問題は、パケットによるネットワーク、特にインターネット・タイプのネットワークで伝送を行う場合にも発生する。

10

## 【 0 0 0 6 】

伝送システム、または、受信担当モジュールにより、(例えば携帯電話のネットワークのように)受信したデータにエラーが多いことや、あるいは(例えばパケット通信による伝送システムの場合のように)データの一つのまとまりが受信されなかったことを検知できる場合には、エラーの抑止シミュレーション方法を活用することになる。これらの方法を用いることにより、先行のフレームから発信される入手可能な信号とデータを基にして、そして場合によっては、消失された区域に基づいて、欠けている信号のサンプルを復号器に外挿することができるようになる。

20

## 【 0 0 0 7 】

このような技術は主に(消失されたフレームの回収技術として)パラメータ式の符号化器の場合に活用されていた。そのような技術により、消失されたフレームが存在する場合に復号器で感知される信号の主観的な劣化を大きく制限することができる。開発されたアルゴリズムの大部分は符号化器及び復号器に用いられる技術に基づくものであり、また、実際に復号器の延長となるものである。

30

本発明の全体的な目的は、言葉と音を圧縮するあらゆるシステムの、復号器で再生される言葉の信号の主観的な質を改善することである。そのような改善が必要となるのは、伝送路の質が悪く、または、パケット通信システムで一つのパケットが失われたり受信されなかったりするなど、連続した符号化済みデータ全体が失われたような場合である。

## 【 0 0 0 8 】

そのために本発明が提案する技術は、符号化技術を用いるかに係わらず、連続する伝送エラー(エラーのパケット)を抑止シミュレーションすることのできるものであり、提案されているその技術は、例えば、エラーのパケットの抑止シミュレーションに必ずしも適しているとは言い切れない構造の時間的符号化器の場合に使用可能なものである。

40

## 【 0 0 0 9 】

## 2. 従来技術の水準

予測式の符号化アルゴリズムの大部分は消失されたフレームの回復技術を提案するものである([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])。例えば伝送路復号器から来るフレームの消失の情報を伝送することによる無線携帯システムの場合に、何らかの形で、伝送路の符号器からの消失されたフレームが一つ発生しているという情報が、その復号器に与えられている。消失されたフレ

50

ムを回復する装置は、健全と認められている先行のフレームのうちの後の方のものにある、一つの（または複数の）ものに基づいて、消失されたフレームのパラメータを外挿することを目的とするものである。予測式の符号化器によって演算子を加えられ、または符号化された幾つかのパラメータには、フレーム間の強い相関関係がある（例えば、依然として「Linear Predictive Coding」線形予測式符号化「LPC」と呼ばれて（〔RABINER〕参照）スペクトル包絡線を示す短期予測のパラメータの場合と、有声音については長期予測のパラメータの場合）。この相関関係からして、エラーがあったり、あるいは乱雑であったりするパラメータを使うよりも、健全な最後のフレームのパラメータを再利用して消失されたフレームを合成する方がずっと好適になる。

【0010】

10

（「Code Excited Linear Prediction」の略である）CELP符号化アルゴリズム（〔RABINER〕参照）については、消失されたフレームのパラメータは従来、次のようにして得られてきた：

- ・LPCフィルタは、パラメータを再複製するか、あるいはある程度の減衰を導入することにより、健全なフレームの最後のもののLPCパラメータから得られる（符号化器G723.1〔REG.723.1A〕参照）。

- ・有声性を検出し、それにより、（〔SALAMI〕のような）消失されたフレームのこの信号の高調波成分度を決定する。この検出は、次のように行われる。

- ・非有声信号の場合：

励振信号はランダムな方法によって発生させられる（僅かに減衰した通過した励振の符号とゲインの単語を抽出し〔SALAMI〕、通過した励振においてランダムな選択を行い〔CHEN〕、完全にエラーになりうる符号が伝送されたものを用いる〔HONKANEN〕、...）

20

- ・有声信号の場合：

LTP遅延時間は一般的には先行するフレームで計算された遅延時間であり、場合によっては、それに軽い「ジグ」が加わっていることもあり〔SALAMI〕、LTPゲインはほぼ1、または1に等しく取る。励振信号は通過した励振に基づいて行われた長期予測のみに限定する。

【0011】

前述した例の全てにおいて、消失されたフレームを抑止シミュレーションする方法は復号器に強く関連付けられており、また、この復号器のモジュールを信号を合成するモジュールとして用いている。それらに用いられるものには、中間信号もあり、該信号は、通過した励振信号として、この復号器の内部で使用可能であり、消失されたフレームに先行する健全なフレームを処理する際に記憶される。

30

【0012】

時間によるタイプの符号化器で符号化されたデータを搬送する際に失われたパケットから生成されたエラーを抑止シミュレーションするために用いられる方法の大部分は、〔GODMAN〕、〔ERDOL〕、〔AT&T〕に示されているような波形置換技術を用いるものである。このタイプの方法で、信号の復元を行う際には、失われた周期の前で復号化された信号の幾つかの部分を選択し、合成モデルは用いない。平滑化技術も活用されるが、それは、様々に異なる信号の連鎖が生成した人工物を回避するためである。

40

【0013】

変換値を用いる符号化器については、消失されたフレームを再構成する技術は、そこで用いられた符号化構造にも適用される。〔PICTEL、MAHIEUX-2〕のようなアルゴリズムは、消失前に有していた数値に基づいて、失われた変換された係数を再生することを目指すものである。

【0014】

〔PARIKH〕において述べられた方法は、あらゆるタイプの信号に応用可能である。その方法の基礎となっているのは、消失に先立って復号化された健全な信号に基づいた正弦曲線のモデルを構成することであり、それにより、その信号の失われた部分を再生する

50

のである。

【0015】

結局のところ、消失されたフレームの抑止シミュレーション技術には、一つの「族」があるが、それらの技術の開発は伝送路の符号化に付随して行われてきた。〔F I N G S C H E I D T〕で述べられているような、これらの方法は、伝送路の復号器が供給する情報を用いるものであり、例えば、受信したパラメータの信頼度に関する情報である。それらの方法は、本発明とは根本的に異なるものであり、本発明は伝送路の符号化器が存在することを前提とはしていない。

【0016】

本発明に最も近いものと考えうる従来技術は、〔C O M B E S C U R E〕に記載されているものであり、そこで提案されている消失されたフレームの抑止シミュレーション方法は、変換による符号化器のためにC E L P符号化器で用いられているものと同等のものである。そのようにして提案された方法の不都合な点は、(「合成」音声、寄生共振等の)スペクトル音響歪みの導入であり、それは特に、(有声音における高調波成分が唯一であること、励振信号の生成が通過した残留信号を部分的に使う場合に限られているなど)制御不良の長期合成フィルタを用いることが原因となっている。さらに、エネルギー制御は、〔C O M B E S C U R E〕では励振信号において実施されており、この信号のエネルギー標的は消失が続いている間は、ずっと一定に維持されているため、邪魔な人工物が生じることにもなっている。

10

【0017】

3. 本発明の説明

本発明は、それ自体に関しては、消失されたフレームの抑止シミュレーションを、さらに高い値のエラーに対しても、そして/または、消失された間隔がもっと長くても格別の音響歪み無しにそれを行うことを可能にする。

20

【0018】

本発明では特に、伝送後に復号化された信号を受信し、伝送されたデータが健全な場合には、復号化されたサンプルを記憶し、短期予測演算子を少なくとも一つと、長期予測演算子を少なくとも一つとを、記憶された健全なサンプルに応じて算定し、そして、復号化された信号において、欠損しているかエラーを含みうるサンプルを、そのようにして算定した演算子によって、生成する、オーディオ・デジタル信号における伝送エラーを抑止シミュレーションする方法を提案する。

30

【0019】

本発明が特に好適な第一の様相によれば、そのようにして生成された合成信号のエネルギー制御を、サンプルごとに計算され、そして適合化されたゲインを用いて制御する。

【0020】

このことが特に有益なのは、その技術の性能を、消失がされる区域で発揮するにつき、さらに長い期間にわたって、改善することにおいてである。

【0021】

特に、合成信号の制御をするためのゲインは、健全なデータに対応するサンプルのために前もって記憶されたエネルギーの値、有声音のための基本波周期、あるいは周波数のスペクトルを特徴づけるあらゆるパラメータというような、パラメータの少なくとも一つに応じて計算するのが好適である。

40

【0022】

また、好適な面としては、合成信号に適用されるゲインは、合成サンプルが生成される持続時間に応じて、徐々に減少していく。

【0023】

また、より好ましい面として、健全なデータにおいて、定常性音と非定常性音とを区別し、そして、異なる法則を可能にするこのゲインの適合化法則の活用を、一方では、定常性音に対応する健全なデータの後に生成されるサンプルのためと、他方では、非定常性音に対応する健全なデータの後に生成されるサンプルのために用いる。

50

## 【0024】

本発明の他の独自の様相によれば、復号化処理のために用いられるメモリの内容を、生成される合成サンプルに応じて更新する。

## 【0025】

この方法によれば、一方では、符号化器と復号器が脱同期化してしまいかねないという可能性を制限し（後述するパラグラフ5.1.4参照）、そして、本発明により再構成した消失された区域と、その区域に続くサンプルとの間で急な不連続が生じるということ避けられる。

## 【0026】

特に、（場合によっては部分的でしかない）復号化の作業に続いて、送信器で活用されるものと類似の符号化を合成されたサンプルに少なくとも部分的に用い、そこで得られるデータが復号器のメモリを再生するのに役立つ。

## 【0027】

とりわけ、この、場合によっては部分的にしか行われな（符号化 - 復号化）作業は、消失された最初のフレームを再生するのに用いるのが好適なのであるが、その理由は、このようなメモリの中にある情報が復号化された健全なサンプルの後の方のものものによって供給されていない場合に、切断の前に復号器のメモリの内容を利用することができるからである（例えば、加算 - 被覆による変換値を用いる符号化器の場合、パラグラフ5.2.2.2.1の10参照）。

## 【0028】

本発明の異なるもう一つの様相によれば、短期予測演算子の入力において生成される励振信号は、有声の区域では、高調波成分と、高調波成分の弱いまたは非高調波成分との和であり、限定された有声の区域では、非高調波成分に限定されているということである。

## 【0029】

特に、高調波成分は、記憶されたサンプルに短期逆フィルタリングを用いることにより計算した残留信号に、長期予測演算子を適用することによるフィルタリングを用いることにより好適に得られる。

## 【0030】

もう一つの成分を決定するについては、長期予測演算子に疑似ランダム（例えばゲイン、または周期の擾乱のような）擾乱を加えることにより、決定される。

## 【0031】

特に好適な方法としては、有声の励振信号を生成するについて、高調波成分はそのスペクトルの低い周波数の方を表すようにしているのに対し、もう一方の成分は高い周波数の部分を表す。

## 【0032】

さらに他のもう一つの様相によれば、長期予測演算子の決定は、記憶された健全なフレームのサンプルに基づいて行われ、この算定のために使用するサンプルの数は、最小値に始まって、その有声音に算定された基本波周期の少なくとも二倍に等しい値に至るまでの間を変化する数である。

## 【0033】

また、残留信号の修正は好適には、非線形的に処理され、それにより、振幅のピークを除去する。

## 【0034】

また、もう一つの好適な様相によれば、その信号が非活性のものであると考えられる場合には、ノイズのパラメータを算定して発声活性を検出すること、そして合成された信号のパラメータを算定されたノイズのパラメータのものに近づける。

## 【0035】

さらに好適な方法としては、復号化された健全なサンプルのノイズのスペクトル包絡線を算定し、同じスペクトル包絡線を有する信号に向かって展開する合成された信号を生成する。

10

20

30

40

50

## 【0036】

本発明が更に提案するのは、言葉と楽音との間の区別を実施し、楽音が検出された場合には、長期予測演算子を算定することなく、前述したタイプの方法を実施し、その励振信号は、例えば一様なホワイトノイズを生成して得られる非高調波成分に限定されることを特徴とする、音声信号の処理方法である。

## 【0037】

本発明はさらに、デジタル・オーディオ信号における伝送エラーを抑止シミュレーションする装置に関するものであり、復号器から装置に伝送された復号化された信号を装置の入力で受信し、この復号化された信号において、欠損しているサンプル、またはエラーのあるサンプルを生成する装置なのであり、前述の方法を用いるのに適した装置の処理手段であるということの特徴とする。

10

## 【0038】

本発明はまた、伝送システムに関するものでもあり、少なくとも一つの符号化器と、少なくとも一つの伝送路と、伝送されたデータが失われてしまった、あるいはエラーの多いことを検出するのに適したモジュールと、少なくとも一つの復号器と、その復号化された信号を受信するエラー抑止シミュレーション装置とからなる伝送システムであり、そのエラー抑止シミュレーション装置が前述したタイプの装置であることを特徴とする。

## 【0039】

## 4. 図の説明

本発明の他の特徴と利点は以下の説明を読むことで、さらに明らかになっていくものであり、ただし、以下の説明はあくまで例示のためのものであり、非制限的なものであり、また、添付図面も参照しつつ、説明を読まなければならない。

20

- ・図1は、本発明で可能な実施態様に従った伝送システムを示す一覧図。
- ・図2と図3は、本発明で可能な実施態様に従った活用法を示す一覧図。
- ・図4から6は、本発明で可能な活用方法に従ったエラーの抑止シミュレーション方法で用いられるウィンドウの概略図。
- ・図7及び8は、音楽信号の場合に使用可能な本発明による活用方法を示す概略図。

## 【0040】

## 5. 本発明で可能な一つまたは複数の実施態様の説明

## 5.1 一つの実施可能な態様の原理

図1はデジタルオーディオ信号を符号化し復号化する装置を示すものであり、それを構成するものは、符号化器1、伝送路2、伝送されたデータが失われたか、もしくはエラーが多いということを検出できるモジュール3と復号器4と、エラーもしくは失われたパケットを本発明に従った実施態様の一つに沿った形で抑止シミュレーションするモジュール5とである。

30

## 【0041】

念のために申し添えると、このモジュール5は、消失されたデータを表示する他に、健全な周期において復号化された信号を受信し、それを更新するために用いられる信号を復号器に伝送するものである。

## 【0042】

さらに詳しくは、モジュール5で実施される処理の基礎となるのは以下のものである。

40

1. 復号化されたサンプルは、伝送されたデータが健全な場合記憶される(処理6)；
2. 消失されたデータの一区画を通して、失われたデータに対応するサンプルを合成する(処理7)；
3. 伝送が修復される際に、消失された周期内に生成された合成サンプルと復号化されたサンプルとの間の平滑化(処理8)。
4. 復号器のメモリの更新(処理9)(更新は、消失されたサンプルの生成中、あるいは伝送修復の時点で行われる)。

## 【0043】

## 5.1.1 健全な周期内で

50



健全なデータを復号化した後、復号化されたサンプルのメモリを更新するのであるが、該メモリには、後になって消失しうる周期ができて、それを再生するに十分な個数のサンプルが含まれている。典型的には、20から40マイクロ秒程度の信号を記憶する。また、健全なフレームのエネルギーを計算して処理された（典型的には5 s程度のもの）に対応するエネルギーをメモリに保存する。

#### 【0044】

5.1.2. 消失されたデータのブロック内で

図3に示された、以下のような作業を行う。

##### 1. 現在のスペクトル包絡線の算定

このスペクトル包絡線の計算は、具体的にはLPCフィルタ〔RABINER〕〔KLEIJN〕の形で行う。分析方法は、従来の方法で（〔KLEIJN〕）健全な周期内で記憶したサンプルをウィンドウ化した後で行う。特に、LPC分析を実施するのは（手順10）のは、フィルタA(z)のパラメータを得るためであり、その逆はLPCフィルタリングを実施するのに用いられる（手順11）。このようにして計算された係数は伝送する必要はないため、この分析の実施については高度な制御命令を用いることができ、その結果、音楽信号については高い性能が得られることになる。

#### 【0045】

##### 2. 有声音の検出及びLTPパラメータの計算

有声音の検出方法（図3の処理12：V/NV、つまり「有聲/非有聲」検出）を記憶されたデータの最後の幾つかに用いられる。そのために使用可能なのは、例えば正規化された相関関係（〔KLEIJN〕）、あるいは以下の実施例の中に示される基準である。

#### 【0046】

その信号が有聲であると表される場合には、なおもLTPフィルタ（〔KLEIJN〕）と呼ばれる長期の合成フィルタを生成できるパラメータを計算する（図3：LTP分析、B(z)により規定するのは計算されたLTP逆フィルタ）。そのようなフィルタは一般的には、基本波周期に対応する周期とゲインとで表される。このフィルタの精度は、分数ピッチまたは多係数構造を用いて改善することが可能である〔KROON〕。

#### 【0047】

その信号が非有聲のものと表われる場合には、LTP合成フィルタに特殊な値を割り当てる（パラグラフ4参照）。

このLTP合成フィルタの算定において特に有益なのは、前の周期が終わるところで分析される区域を限定することである。分析ウィンドウの長さは、最小値から始まって、その信号の基本波周期に関連する値に至るまでの間で変化する。

#### 【0048】

##### 3. 残留信号の計算

残留信号の計算は記憶されたサンプルの後の方のものものにLPC逆フィルタリング（処理10）を実施することにより行われる。つぎに、この信号を用いてLPC合成フィルタ11の励振信号を発生させる（以下を参照）。

#### 【0049】

##### 4. 欠損サンプルの合成

代替サンプルの合成は、（LPC逆フィルタの出力で、その信号に基づき13で計算した）励振信号を、1で計算したLPC合成フィルタ11（1/A(z)）の中に導入することで行う。この励振信号を生成する方法には二つあり、その信号が有聲のものかそうでないかによって異なる。

#### 【0050】

##### 4.1 有聲区域において

励振信号は、二つの信号を、一つの高調波成分の強い成分と一つの高調波成分の弱い、または全くない成分とを合計したものである。

#### 【0051】

高調波成分の強い成分は、2で計算されたパラメータを用いて、3で述べた残留信号に（

処理 14 のモジュールの) LTP フィルタリングによって得られる。

【0052】

第二の成分もまた、LTP フィルタリングによって得られるが、パラメータに乱数的修正を加え、疑似乱数信号を生成することにより非周期的なものになる。

【0053】

第一の成分の通過周波帯をスペクトルの周波数の低いものに限定することは特に有益である。同様に、第二の成分をさらに高い周波数に限定することも有益なものとなる。

【0054】

4.2 非有声区域において

その信号が非有声である場合、非高調波成分的な励振信号が生成される。有声音について用いられるのと同様の生成方法を(周期、ゲイン、徴候などの)パラメータを変化させて用いることにより、非高調波成分的な方法にすることが有益である

【0055】

4.3 残留信号の振幅制御

その信号が非有声である場合、あるいは、有声の度合いが弱い場合、励振の生成に用いられる残留信号を処理することにより、平均を有意に越える振幅のピークを除去する。

【0056】

5. 合成信号のエネルギー制御

合成信号のエネルギーを計算されたゲインによって制御し、そしてサンプルごとに適合化させる。消失の周期が比較的長い場合には、合成信号のエネルギーを徐々に下げることが必要になる。ゲインの適合化法則の計算は、消失される前に記憶されたエネルギーの値(1参照)、基本波周期、そして切断時の信号の局所的定常性などの、様々なパラメータに応じて行われる。

【0057】

そのシステムに、(音楽のような)定常的音と(言葉のような)非定常的音とを区別できるモジュールが含まれている場合には、様々に異なった適合化法則を用いることもまた可能である。

【0058】

加算 - 被覆によって変換値を用いる符号化器の場合には、正確に受信した最後のフレームのメモリの先の方のものには、失われた最初のフレームの先の方のものについてのかなり精度の高い情報が含まれている(加算 - 被覆におけるその重みは実際のフレームのものよりもさらに大きい)。この情報もまた適合化ゲインの計算に用いることが可能である。

【0059】

6. 合成の手順を時間の経過とともに辿る:

消失の周期が比較的長い場合には、合成のパラメータを展開することもできる。システムが([REC - G. 723.1A]、[SALAMI - 2]、[BENYASSINE])のように)ノイズのパラメータを検出する装置と結合されている場合、特に有益となるのが、再構成すべき信号を生成するパラメータを算定されたノイズのパラメータに近づけることである。それを特に、(LPCフィルタを算定されたノイズのそれと内挿し、その内挿の係数は、時の経過とともに、そのノイズのフィルタが得られるまで進展することになる)スペクトル包絡線のレベルで行い、そして(例えば、ウィンドウ化により、ノイズのものに向かって徐々に進展していくレベルである)エネルギーのレベルでも行う。

【0060】

5.1.3. 伝送の修復

伝送を修復させるに際して特に重要なのは、前記各パラグラフにおいて規定した技術により再構成した消失された周期と、その後続く周期、つまり、その信号を復号化するために伝送された情報の一切を自由に入手できる周期との間に突然、破綻が生じるということはないようにするということである。本発明は、時間の領域で加重を行うものであり、それは、通信の修復に先行する代替サンプルと消失された周期の後の健全な復号化されたサンプルとの間で内挿を行うことによる加重である。この作業は、どのようなタイプの符号

化器を用いるかに係わらないものであることは自明である。

【0061】

加算 - 被覆によって変換値を用いる符号化器の場合には、この作業は以下のパラグラフで述べられるメモリを更新するのと共通の作業である（実施例参照）。

【0062】

5.1.4. 復号器のメモリの更新

消失された周期の後に健全なサンプルの復号化を再開する場合、その前の記憶されたフレームで通常通り生成されたデータをその復号器が用いると劣化が生じる可能性がある。重要なのは、これらのメモリの更新を適切に行い、これらの人工物を回避することである。

【0063】

これは、一つのサンプルまたは一連のサンプルについて、先行するサンプルを復号化した後に得られる情報を利用する回帰的方法を用いる符号化構造にとって、特に重要である。これらは例えば、その信号の冗長性を抽出することのできる予測（[KLEIJN]）である。これらの情報は、通常、符号化器でも復号器でも同時に使用可能であり、符号化器は、そのために先行するサンプルに、一つの形式の局所的復号化を既に行っていないければならず、そして、復号器は受信時に遠くにあるものである。伝送路が擾乱を受け、遠隔復号器が送信に際し存在する局所的復号器と同じ情報をもはや用いられなくなるとすぐに、符号化器と復号器の間で脱同期化が生じる。回帰性の強い符号化システムの場合、この脱同期化によって、聞き取れる程の劣化が生じる恐れがあり、構造内部に不安定なものがある場合にはそれが長く続き、さらには時間の経過とともに増幅しかねない。よって、この場合に重要となるのは、符号化器と復号器との間で再同期化を行うように努力すること、つまり、復号器のメモリを符号化器のメモリにできるだけ近く算定するということである。しかしながら、再同期化技術は、そこで用いられる符号化構造に左右される。そのうちの一つを後述にて示すが、その原理は本特許願において一般的なものであるものの、その複雑さは潜在的に大きい。

【0064】

考えられる方法の一つは、要するに、受信に際しての復号器に、送信に際し存在するものと同じタイプの符号化モジュールを導入することであり、それにより、前述のパラグラフで述べた技術により生成された信号のサンプルの符号化 - 復号化を、消失された周期内で行えるようにすることである。この方法により、後に続くサンプルを復号化するのに必要なメモリを、（消失された周期内で一定の定常性がある場合は別として）失われてしまったデータと間違いなく近いデータで補完することになる。この定常性の仮説が、例えば消失された周期が長く続いた後で重要と思われない場合、事態を改善するに足るだけの情報は、得られないことになる。

【0065】

実際には、一般的にはこれらのサンプルの完全な符号化を行う必要はなく、メモリの更新に必要なモジュールに限定して行われるものである。

【0066】

この更新は、代替サンプルの生成時に行うことが可能であり、そのことにより、複雑さを消失区域全体にわたって分散させることになるが、前述の合成方法により併合されることになる。

符号化構造によりそれが可能ならば、消失された周期に続く健全なデータ周期の始めの中間区域に限定することでその方法を用いてもよい。その場合、更新方法は、復号化作業と併合されることになる。

【0067】

5.2. 特殊な実施例の説明

考えられる実施例の特殊なものを以下に示す。TDACまたはTCDM（[MAHIEUX]）タイプの変換値を用いた符号化器の場合を特に取り上げている。

【0068】

5.2.1 装置の説明

10

20

30

40

50

T D A C タイプの変換値を用いたデジタル符号化 - 復号化システム。

2 4 k b / s ないし 3 2 k b / s の拡大帯域 ( 5 0 - 7 0 0 0 H z ) の符号化器。

2 0 m s のフレーム ( 3 2 0 個のサンプル ) 。

2 0 m s の加算 - 被覆による 4 0 m s ( 6 4 0 個のサンプル ) のウィンドウ。一つの二進法フレームに符号化されたパラメータがあり、それは、一つのウィンドウで T D A C 変換によって得られたパラメータである。これらのパラメータを復号化した後、T D A C 逆変換を行い、2 0 m s の出力フレームを得るが、そのフレームは、先行するウィンドウの後半と現行のウィンドウの先の方のものとの和である。図 4 では、( 時間に関する ) フレーム n の再構成用に用いられるウィンドウの二つの部分が太字で示してある。このようにして、失われた二進法フレームが、連続する二つのフレーム ( 現行のものとその後に続くもの、図 5 ) の再構成を擾乱する。逆に、失われたパラメータの代替を正確に行うことにより、それら二つのフレームを再構成するための、( 図 6 の ) 二進法フレームからの情報の二つの部分、先行する部分とその後に続く部分を回復することができる。

【 0 0 6 9 】

### 5 . 2 . 2 実施

以下に述べる作業の全てを、図 1 及び図 2 に従って受信の際に実施するが、それは、復号器と交信する、消失されたフレームを抑止シミュレーションするモジュールの内部において実施したり、あるいは、その復号器そのものの内部において実施したりする ( 復号器のメモリの更新 ) 。

【 0 0 7 0 】

#### 5 . 2 . 2 . 1 健全な周期内

パラグラフ 5 . 1 . 2 に対応して、復号化されたサンプルのメモリを更新する。このメモリは二進法フレームが消失した場合の通過した信号の L P C 及び L T P 分析を行うために用いられるものである。ここに示された例においては、L P C 分析は、2 0 m s ( 3 2 0 個のサンプル ) の信号の周期で行われる。一般的には、L T P 分析には、記憶すべきサンプルがさらに多く必要となる。この例においては、L T P 分析を正確に行うことができるように、記憶されたサンプルの個数はピッチの最大値の二倍に等しい数である。例えば、ピッチの最大値  $M a x P i t c h$  を 3 2 0 個のサンプル ( 5 0 H z , 2 0 m s ) に定めると、後ろから数えて 6 4 0 個のサンプルが記憶されることになる ( その信号の 4 0 m s ) 。健全なフレームのエネルギーの計算も行い、それら健全なフレームを長さ 5 s の円形のバッファに保存する。消失されたフレームが検出されると、その最後の健全なフレームのエネルギーをこの円形緩衝器の最大値と最小値に比較し、それにより、その相対エネルギーを認識する。

【 0 0 7 1 】

#### 5 , 2 . 2 . 2 消失されたデータの一区画間

二進法フレームが失われる場合には、二つの異なるケースを区別する：

【 0 0 7 2 】

5 , 2 . 2 . 2 . 1 健全な一つの周期の後に失われた第一の二進法フレーム まず、記憶された信号の分析を行い、それにより、再生された信号を合成するのに役立つモデルのパラメータを算定する。このモデルにより、我々は、つぎに 4 0 m s の信号を合成することができるのであり、そのことは、失われた 4 0 m s のウィンドウに対応している。T D A C 変換を行った後に、( パラメータの符号化 - 復号化はせずに ) この合成された信号に T D A C 逆変換を行って、2 0 m s の出力信号を得る。このように T D A C - 逆 T D A C の作業を行うことにより、正確に受信された先行するウィンドウからの情報を利用することができる ( 図 6 参照 ) 。同時に、復号器のメモリの更新を行う。そのようにして、後に続く二進法フレームは、それが確かに受信される場合には、正常に復号化することができる ( 図 6 ) 。

行うべき作業は次の通りである。

【 0 0 7 3 】

1 . 記憶された信号のウィンドウ化。例えば、2 0 m s のハミングの非対称ウィンドウを

10

20

30

40

50

用いることができる。

【0074】

2. ウィンドウ化信号について自動相関関係の関数の計算

【0075】

3. LPCフィルタの係数の決定。そのためには、従来レビンソン - ダービンの反復アルゴリズムが用いられてきた。特に符号化器を用いて音楽シーケンスの符号化を行う場合に、分析の等級を上げることができる。

【0076】

4. 有声性を検出してその信号（有声音）に周期性があれば、それをモデル化するために記憶した信号の長期分析を行う。ここで示した実施例において、本発明者等は基本波周期  $T_p$  の算定を整数値に限定し、有声性の程度の算定を、具体的には、選択された周期で評価されたマックスコル相関係数（下記参照）の形で、計算した。Fs がサンプリングの頻度であるとする  $T_m = \max(T, Fs / 200)$  であれば、Fs / 200 個のサンプルが持続時間 5 ms に対応することになる。先行するフレームの終わりの信号の展開をさらによくモデル化するために、記憶された信号の終わりで  $2 * T_m$  個のサンプルのみを用いて、遅延 T に対応する相関係数  $Corr(T)$  の係数を計算する。

【0077】

【数1】

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

20

【0078】

但し、 $m_0 \cdots m_{Lmem-1}$  は先に復号化した信号のメモリである。この式から、このメモリ  $Lmem$  の長さは（また「ピッチ」と呼ばれる）基本波周期  $MaxPitch$  の最大値の少なくとも二倍でなければならないことがわかる。

600 Hz の周波数に対応する基本波周期  $MinPitch$  の最小値もまた定められた（ $Fs = 16kHz$  で 26 個のサンプル）。 30

【0079】

$T = 2, \dots, MaxPitch$  について  $Corr(T)$  を計算する。（非常に短期の相関係数は除外するとして） $T'$  が  $Corr(T') < 0$  のような最小の遅延である場合には、 $T' < T \leq MaxPitch$  の最大値  $MaxCorr$  を求める。すなわち  $T_p$  が  $MaxCorr$  に対応する周期（ $Corr(T_p) = MaxCorr$ ）。また、 $T' < T \leq 0.75 * MinPitch$  について  $Corr(T)$  の最大値、 $MaxCorrMp$  も求める。 $T_p < MinPitch$  または  $MaxCorrMp > 0.7 * MaxCorr$  の場合、そして、最後の健全なフレームのエネルギーが比較的弱い場合には、そのフレームは非有声であるという決定を下すことになるが、その理由は、LTP 予測を用いると、非常にやっかいな高周波の中に共振が得られるという危険を冒しかねないからである。選択されたピッチは  $T_p = MaxPitch / 2$  であり、そして相関係数  $MaxCorr$  は小さな値（0.25）に定められている。 40

【0080】

そのエネルギーの 80% を越えるものが終わりの方の  $MinPitch$  サンプルの中に集中している場合には、そのフレームもまた非有声であるものとして考える。それゆえに、言葉の開始ということなのであるが、サンプルの数は基本波周期でありうるものを算定するに足りるだけのものではなく、それを非有声であるものとして処理した方がよく、合成された信号のエネルギーをもっと早く減らした方がいいとさえいえる（それを知らせるため、 $DiminFlag = 1$  とする）。

50

## 【0081】

MaxCorr > 0.6 の場合には、基本波周期の倍数 (4 倍、3 倍または 2 倍) が見つからなかったということを確認する。そのために、 $T_p / 4$ 、 $T_p / 3$  そして  $T_p / 2$  の周辺の相関関係の局所的な最大値を求める。念のため、 $T_1$  はこの最大値の位置であり、 $MaxCorrL = Corr(T_1)$  である。 $T_1 > MinPitch$  で  $MaxCorrL > 0.75 * MaxCorr$  である場合には、 $T_1$  を新しい基本波周期として選ぶ。

## 【0082】

$T_p$  が  $MaxPitch / 2$  よりも小さい場合は、それが本当に有声のフレームなのかどうかを、 $2 * T_p$  ( $TPP$ ) の前後の相関関係の局所的な最大値を求め、そして  $Corr(T_{pp}) > 0.4$  であることを確かめて、検証してもよい。 $Corr(T_{pp}) < 0.4$  である場合、そして信号のエネルギーが減少する場合には、 $DiminFlag = 1$  とし、 $MaxCorr$  の値を減らし、さもなければ、それに続く局所的な最大値を実際の  $T_p$  と  $MaxPitch$  との間で求める。

## 【0083】

有声性のもう一つの基準は、つまりは、少なくとも  $2 / 3$  の場合に、基本波周期の分だけ遅延した信号が遅延のない信号と同じ特徴をもっているかどうかを検証することである。

## 【0084】

その検証を  $5ms$  と  $2 * T_p$  との間の最大値に等しい長さについて行う。

## 【0085】

信号のエネルギーに減少傾向があるかどうかを検証する。もしあるなら、 $DiminFlag = 1$  とし、 $MaxCorr$  の値を減少の割合に応じて下げる。

## 【0086】

有声性の判定には、信号のエネルギーも考慮に入れる。そのエネルギーが強い場合には、 $MaxCorr$  の値を増大させ、そのため、そのフレームが有声であると判定される可能性が高まることになる。逆に、そのエネルギーが非常に弱ければ、 $MaxCorr$  の値を減らす。

## 【0087】

結局のところ、有声性の判定は  $MaxCorr$  の値に応じて行う。 $MaxCorr < 0.4$  であれば、ただそれだけのことで、そのフレームは有声のものではない。非有声であるフレームの基本波周期  $T_p$  は制限され、それは  $MaxPitch / 2$  以下でなければならない。

## 【0088】

5. 記憶されたサンプルの後の方のものものを LPC 逆フィルタリングすることにより残留信号の計算を行う。この残留信号はメモリ ResMem に保存される。

## 【0089】

6. 残留信号のエネルギーの平均化。非有声であるか、または有声性が弱い信号の場合 ( $MaxCorr < 0.7$ ) である場合、ResMem に保存された残留信号のエネルギーは、ある部分から他の部分へと突然変化することがある。この励振の反復により、合成信号において非常に不愉快な周期的擾乱が引き起こされることになる。それを避けるためには、有声性の弱いフレームの励振において大きな振幅のピークは一切ないようにすることを確実にする。励振は残留信号の後の方の  $T_p$  個のサンプルに基づいて構成されるため、 $T_p$  個のサンプルのこのベクトルを処理する。我々の例において用いられる方法は次のようなものである。

- ・残留信号の後の方のもの  $T_p$  個のサンプルの絶対値の平均 MeanAmpl を計算する。

- ・処理対象のサンプルのベクトルにゼロの  $n$  個の通過がある場合には、それを  $n + 1$  個のサブ・ベクトルに切り、サブ・ベクトルそれぞれの信号の兆候が変化しないようにする。

- ・サブ・ベクトルそれぞれの最大振幅 MaxAmplSv を求める。 $MaxAmplSv > 1.5 * MeanAmpl$  である場合には、サブ・ベクトルに  $1.5 * MeanAmpl / MaxAmplSv$  を掛ける。

## 【0090】

7. TDACウィンドウの長さに対応する640個の長さの励振信号の準備。有声性に  
 応じて2つのケースを区別する。

・励振信号は、スペクトル  $e x c b$  の周波数の低いものに帯域が限定された高調波成分の  
 強い成分と、さらに周波数の高い  $e x c h$  に限定された高調波成分のより弱いもう一つの  
 成分との、二つの信号の和である。

高調波成分の強い成分は、残留信号の等級3のLTPフィルタリングを行うことにより得  
 られる。

$$e x c b ( i ) = 0 . 1 5 * e x c ( i - T p - 1 ) + 0 . 7 * e x c ( i - T p ) + 0 . 1 5 * e x c ( i - T p + 1 )$$

10

## 【0091】

係数〔0.15、0.7、0.15〕は  $F s / 4$  で3dBの減衰の低域フィルタリングFIR  
 に対応している。

第二の成分もまたLTPフィルタリングを行うことにより得られるのであるが、それは基  
 本波周期  $T p h$  の乱数的修正により周期性をなくしたものである。 $T p h$  は乱数実数値  $T p a$   
 の整数部分として選ばれる。

$T p a$  の初期の値は  $T p$  に等しく、つぎに、〔-0.5、0.5〕の乱数値を加算して、  
 サンプルごとに修正される。さらに、このLTPフィルタリングは高域フィルタリングIIR  
 と組み合わせられる。

$$e x c h ( i ) = - 0 . 0 6 3 5 * ( e x c ( i - T p h - 1 ) + e x c ( i - T p h + 1 ) ) + 0 . 1 1 8 2 * e x c ( i - T p h ) - 0 . 9 9 2 6 * e x c h ( i - 1 ) - 0 . 7 6 7 9 * e x c h ( i - 2 )$$

20

## 【0092】

有声の励振は、その場合、それら2つの成分の和である。

$$E x c ( i ) = e x c b ( i ) + e x c h ( i )$$

## 【0093】

・非有声であるフレームの場合には、励振信号  $e x c$  もまた、係数〔0.15、0.7、  
 0.15〕で等級3のLTPフィルタリングにおいて得られるのであるが、それは、10  
 個のサンプル全てで基本波周期を1に等しい値だけ増やし、兆候を0.2の確率で逆転さ  
 せることで、周期性をなくしている。

30

## 【0094】

8. 3で計算されたLPCフィルタにおける励振信号  $e x c$  を導入した代替サンプルの合  
 成。

## 【0095】

9. 合成信号のエネルギーのレベルの制御

エネルギーは、最初の代替フレームが合成された時点から事前に定められたレベルに向か  
 って徐々に近づいていく傾向がある。このレベルを規定するのは、例えば、消失に先行す  
 る最後の5秒間を通じて見いだされる最も弱い出力のフレームのエネルギーとして、規定  
 することが可能である。我々の場合は、二つの、ゲインの適合化法則を規定したが、該法  
 則の選択は4で計算されたフラッグ  $D i m i n F l a g$  に応じて行われる。エネルギー減  
 少の速度はまた、基本波周期によっても左右される。さらに根本的な第三の適合化法則が  
 存在するが、それが用いられるのは、生成された信号の始まりが、後で説明するように  
 (11参照)、最初の信号にうまく対応しないことが検出される場合である。

40

## 【0096】

10. この章の始めで説明したように、8で合成された信号においてTDAC変換が行われ  
 る。得られたTDAC係数は失われたTDAC係数の代わりとなる。そしてTDAC逆  
 変換を行い、出力フレームを得る。これらの演算には三つの目的がある：

・失われたのが最初のウィンドウである場合には、この方法で、正確に受信された先行す  
 るウィンドウの情報を利用し、該ウィンドウの中において、擾乱された最初のフレームを  
 再構成するのに必要なデータの半分がある(図6)。

50

・後に続くフレームを復号化するために復号器のメモリを更新する（符号化器と復号器の同期化、パラグラフ 5.1.4 参照）。

・正確に受信された最初の二進法フレームが、上記に示した技術（パラグラフ 5.1.3 参照）によって再構成した消失された周期の後に到着する場合には、出力信号が（断絶なしに）連続推移を自動的に保証する。

【0097】

11. 加算 - 被覆の技術により、合成された有声信号が最初の信号によく対応しているかいないかを検証できるようになるが、その理由は、失われた最初のフレームの先の方のものについては、正確に受信した最後のウィンドウのメモリの重みがさらに大きいからである（図 6）。

それゆえに、合成された最初のフレームの方のもの、T D A C と逆 T D A C 演算の後で得られたフレームの先の方のものとの間の相関関係を取ることによって、失われたフレームと代替フレームとの間の相似を算定することができる。相関関係が弱い（ $< 0.65$ ）ということは、元の信号が、代替方法によって得られた信号とはかなり異なっているということになり、この後者の信号のエネルギーを最小のレベルに向かって急速に減少させた方がいいということになる。

【0098】

5.2.2.2.2 消失された区域の最初のフレームの後に続く失われたフレーム前のパラグラフの 1 から 6 は、消失された最初のフレームに先行する復号化された信号の分析に関するものであり、その信号の合成モデルの構成（L P C と場合によっては L T P）を可能にする。後に続く消失されたフレームについては、その分析をやり直すことはせず、失われた信号の代替は、最初のフレームが消失された際に計算したパラメータ（係数 L P C、pitch、Max Corr、Res Mem）に基づいて行われる。それゆえに、その信号の合成とその復号器の同期化に対応する演算のみを行うのであるが、そこに、消失された最初のフレームに対し、以下のような修正を加える。

・（前記 7 及び 8 の）合成部分において、320 個の新しいサンプルだけを生成するのだが、その理由は、T D A C 変換のウィンドウの範囲に含まれるのは先行する消失されたフレームの時に生成された後の方のもの 320 個のサンプルと、これらの新しい 320 個のサンプルだからである。

・消失の周期が比較的長くなる場合に、重要となるのは、合成パラメータを、ホワイトノイズのパラメータに向かって、または、規定ノイズのパラメータに向かって、展開していくことである（パラグラフ 3.2.2.2 の 5 参照）。この例で示されるシステムには V A D / C N G は含まれていないので、我々には、例えば、次のような一つまたは幾つかの修正を行える可能性がある：

・L P C フィルタをフラット・フィルタとを段階的に内挿することにより、合成された信号を色彩の弱いものにする。

・ピッチの値を徐々に増大させる。

・有声モードでは、一定時間の後に（例えばエネルギーが最小値に達した時に）、非有声であるモードに切り換える。

【0099】

5.3 音楽信号の特定処理。そのシステムに含まれているモジュールが言葉 / 音楽の区別を可能にするものである場合は、音楽の合成モードを選択した後で、音楽信号の特定処理を実施することができる。図 7 では、音楽合成モジュールは 15 という参照番号を付され、言葉合成モジュールは 16 という参照番号を付され、言葉 / 音楽切り換え器は 17 という参照番号になっている。

そのような処理は、例えば、音楽合成モジュールについては、図 8 に示されるような、以下の手順を活用するものである。

【0100】

1. 現行のスペクトル包絡線の算定

このスペクトル包絡線の計算は、L P C フィルタ [ R A B I N E R ] [ K L E I J N ] の

10

20

30

40

50



形で行われる。分析は従来技術で行われている（〔KLEIJN〕）。健全な周期で記憶されたサンプルをウィンドウ化した後、LPC分析を実施し、フィルタLPC A(Z)を計算する（手順19）。この分析で用いる等級は高度のもの（>100）であり、それにより、音楽信号について高性能を実現する。

#### 【0101】

##### 2. 欠けているサンプルの合成：

代替サンプルの合成は、手順19で計算された合成フィルタLPC(1/A(z))の中に励振信号を導入することにより行われる。この - 手順20で計算される - 励振信号は、ホワイトノイズであり、その振幅の選択は、健全な周期で記憶された後のもののN個のサンプルのエネルギーと同じエネルギーを有する信号が得られるように、行われる。図 10

残留信号の振幅制御の例：

励振が、ゲインによって増倍させられた一様ホワイトノイズとしての外観を呈する場合は、このゲインGは次のようにして計算可能である：

LPCフィルタのゲインの算定：

ダービンのアルゴリズムによって残留信号のエネルギーが求められる。残留信号のエネルギーはまたモデル化によっても認識しうるものであり、それによって、LPCフィルタのゲイン $G_{LPC}$ を、これら二つのエネルギーの比として算定する。

標的エネルギーの計算：

健全な周期で記憶された後のもののN個のサンプルのエネルギーに等しい標的エネルギーを算定する（Nは、典型的にはLPC分析用の信号の長さよりも小さい）。 20

合成された信号のエネルギーは $G^2$ と $G_{LPC}$ によるホワイトノイズのエネルギーとの積である。

Gの選択は、このエネルギーが標的エネルギーと等しくなるように選択した。

#### 【0102】

##### 3. 合成信号のエネルギー制御

言葉信号についてと同様であるが、合成信号のエネルギーの減少速度はずっとゆっくりしており、その速度は（実在しない）基本波周期には左右されない。

合成信号のエネルギー制御は、サンプルごとに計算され適合化させられたゲインを用いて行われる。消失周期が比較的長い場合には、合成信号のエネルギーを段階的に下げることが必要である。ゲインの適合化法則は、様々に異なるパラメータに応じて、消失前に記憶されたエネルギーの値として、そして切断時のその信号の局所的定常性として、計算可能である。 30

#### 【0103】

##### 4. 合成の手順を時間に沿って辿っていく

言葉信号についてと同様に

消失周期が比較的長い場合には、合成パラメータもまた進展させていくことが可能である。そのシステムが連結されている装置が、（〔REC-G.723.1A〕、〔SALAMI-2〕、〔BENYASSINE〕のような）ノイズのパラメータを算定して、発声活性の検出または音楽信号の検出をする装置である場合には、再構成すべき信号を生成するパラメータを算定されたノイズのパラメータに近づけていくことが特に有益である。それが特にそういえるのは、（時間の経過とともにそのノイズのフィルタが得られるまで進展していく内挿係数で、LPCフィルタを算定されたノイズのフィルタと内挿する）スペクトル包絡線のレベルと（例えばウィンドウ化によってノイズのレベルに向かって徐々に進展していくレベルの）エネルギーのレベルにおいてである。 40

#### 【0104】

##### 6. 全般的考察

やがて了解されることと思うが、以上に説明した技術の利点は、どのようなタイプの符号化器とも使用可能であるということであり、特に、以上に説明した技術により、言葉信号と音楽信号について、時間的符号化器あるいは変換値を用いた符号化器で問題となる、ビ 50

ット・パケットが紛失するという問題を克服することが可能になる。事実、本技術においては、伝送されたデータが健全な周期に際して記憶された信号のみが、復号化器から発信されるサンプルとなり、どのような構造の符号化を用いているかにかかわらず、入手可能な情報となる。

【0105】

#### 7. 参考文献

- [AT&T] AT&T (D. A. Kapi low, R. V. Cox) 《  
A high quality low-complexity algorithm  
for frame erasure concealment (FEC) wit  
h G. 711》, Delayed Contribution D. 249 10  
(WP3/16), ITU, may 1999.
- [ATAL] B. S. Atal et M. R. Schroeder. "P  
redictive coding of speech signal and su  
bjectives error criteria". IEEE Trans. o  
n Acoustics, Speech and Signal Processin  
g, 27:247-254, juin 1979.
- [BENYASSINE] A. Benyassine, E. Shlomot e  
t H. Y. Su. "ITU-T recommendation G. 729  
Annex B: A silence compression scheme  
for use with G. 729 optimized for V. 70 20  
digital simultaneous voice and data appl  
ications". IEEE Communication Magazine,  
septembre 97, PP. 56-63.
- [BRANDENBURG] K. H. Brandenburg et M. Bo  
ssi. "Overview of MPEG audio: current a  
nd future standards for low-bit-rate aud  
io coding". Journal of Audio Eng. Soc.,  
Vol. 45-1/2, janvier/fevrier 1997, PP. 4  
-21.
- [CHEN] J. H. Chen, R. V. Cox, Y. C. Lin, 30  
N. Jayant et M. J. Melchner. "A low-del  
ay CELP coder for the CCITT 16 kb/s spee  
ch coding standard". IEEE Journal on Sel  
ected Areas on Communications, Vol. 10-5  
, juin 1992, PP. 830-849.
- [CHEN-2] J. H. Chen, C. R. Watkins. "Lin  
ear prediction coefficient generation du  
ring frame erasure or packet loss". Brev  
et US5574825, EP0673018.
- [CHEN-3] J. H. Chen, C. R. Watkins. "Lin 40  
ear prediction coefficient generation du  
ring frame erasure or packet loss". Brev  
et 884010.
- [CHEN-4] J. H. Chen, C. R. Watkins. "Fra  
me erasure or packet loss compensation m  
ethod". Brevet US5550543, EP0707308.
- [CHEN-5] J. H. Chen. "Excitation signal  
synthesis during frame erasure or packet  
loss". Brevet US5615298, EP0673017.
- [CHEN-6] J. H. Chen. "Computational comp 50

- lexity reduction during frame erasure of packet loss". Brevet US 5 717 822.
- [CHEN-7] J. H. Chen. "Computational complexity reduction during frame erasure or packet loss". Brevet US 9 402 124 35, EP 0 673 015.
- [COX] R. V. Cox. "Three new speech coders from the ITU cover a range of applications". IEEE Communication Magazine, Septembre 97, PP. 40-47. 10
- [COX-2] R. V. Cox. "An improved frame erasure concealment method for ITU-T Rec. G728". Delayed contribution D. 107 (WP3/16), ITU-T, janvier 1998.
- [COMBESCORE] 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. 20
- [DAUMER] W. R. Daumer, P. Mermelstein, X. Maitre et I. Tokizawa. "Overview of the ADPCM coding algorithm". Proc. of GLOBECOM 1984, PP. 23.1.1-23.1.4.
- [ERDOL] N. Erdol, 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. 30
- [FINGSCHIEDT] 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, decembre 1986, PP. 1440-1448. 40
- [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., fevrier 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.
- [HELLWIG] K. Hellwig, P. Vary, D. Massal 50

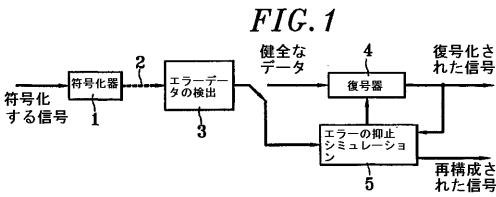
- oux, J. P. Petit, C. Galand et M. Rosso. "Speech codec for the European mobile radio system". GLOBECOM conference, 1989, PP. 1065-1069.
- [HONKANEN] T. Honkanen, J. Vainio, P. Kapanen, P. Haavisto, R. Salami, C. Laflamme et J. P. Adoul. "GSM enhanced full rate speech codec". Proc. of ICASSP conference, 1997, PP. 771-774.
- [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. 10
- [KROON2] P. Kroon, "Linear prediction coefficient generation during frame erasure or packet loss". Brevet US5450449, EP0673016.
- [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. 20
- [MAHIEUX-2] Y. Mahieux, "Dissimulation erreurs de transmission". brevet 92 06720 depose le 3 juin 1992.
- [MAITRE] X. Maitre. "7 kHz audio coding within 64 kbit/s". IEEE Journal on Selected Areas on Communications, Vol. 6-2, fevrier 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. 30
- [PICTEL] PictureTel Corporation, "Detailed Description of the PTC (PictureTel Transform Coder), Contribution ITU-T, SG15/WP2/Q6, 8-9 Octobre 1996 Baltimore meeting, TD7
- [RABINER] L. R. Rabiner, R. W. Schafer. "Digital processing of speech signals". Bell Laboratories inc., 1978. 40
- [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"
- [SALAMI] R. Salami, C. Laflamme, J. P. Adoul, A. Kataoka, S. Hyashi, T. Moriya, C. Lamblin, D. Massaloux, S. Proust, P. 50

- 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.
- [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. 10
- [TRAMAIN] T. E. Tremain. "The government standard linear predictive coding algorithm: LPC 10". Speech technology, avril 1982, PP. 40-49.
- [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.
- 【図面の簡単な説明】 20
- 【図1】本発明で可能な実施態様に沿った伝送システムを示す一覧図。
- 【図2】本発明で可能な実施態様に沿った活用法を示す一覧図。
- 【図3】本発明で可能な実施態様に沿った活用法を示す一覧図。
- 【図4】本発明で可能な活用法に沿ったエラーの抑止シミュレーション方法で用いられるウィンドウを概略的に示す図。
- 【図5】本発明で可能な活用法に沿ったエラーの抑止シミュレーション方法で用いられるウィンドウを概略的に示す図。
- 【図6】本発明で可能な活用法に沿ったエラーの抑止シミュレーション方法で用いられるウィンドウを概略的に示す図。
- 【図7】音楽信号の場合に使用可能な本発明による活用方法を概略的に示す図。 30
- 【図8】音楽信号の場合に使用可能な本発明による活用方法を概略的に示す図。
- 【符号の説明】
- 1 符号化器
  - 2 伝送路
  - 3 エラーデータの検出
  - 4 復号器
  - 5 エラーの抑止シミュレーション
  - 6 復号されたサンプルのメモリ化
  - 7 欠損サンプルの合成
  - 8 復号信号/再構成信号の平滑 40
  - 9 復号器の更新
  - 10 LPC分析
  - 11 LPCフィルタリング  $1/A(Z)$
  - 12 LTP分析と有声/非有声検出
  - 13 励振信号の計算
  - 14 LTPフィルタリング
  - 15 音楽合成
  - 16 言葉合成
  - 17 言葉/音楽切り替え器
  - 19 LPC分析 50

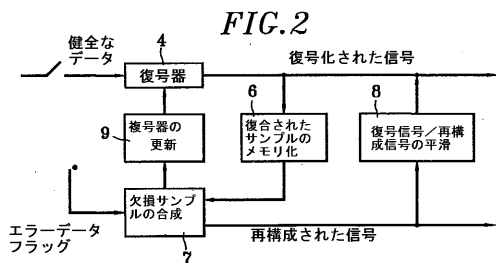
2 0 励振信号の計算

2 1 L P C フィルタリング  $1 / A ( Z )$

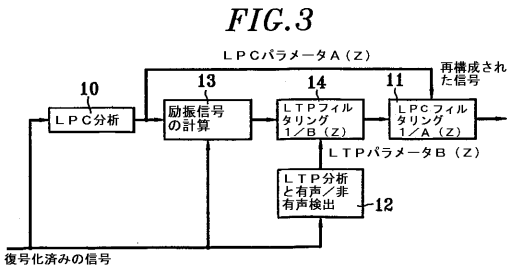
【 図 1 】



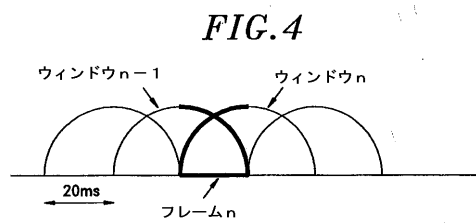
【 図 2 】



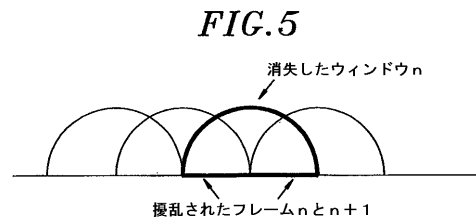
【 図 3 】



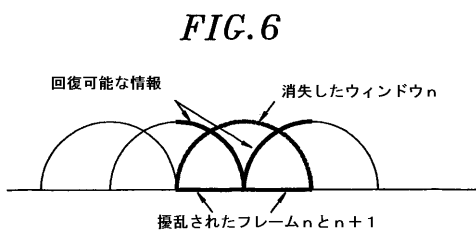
【 図 4 】



【 図 5 】

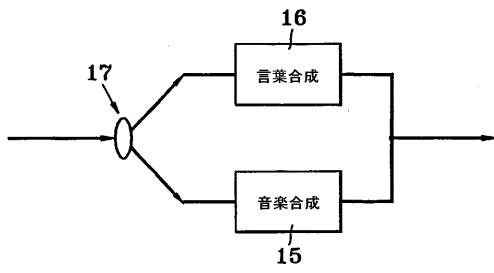


【 図 6 】



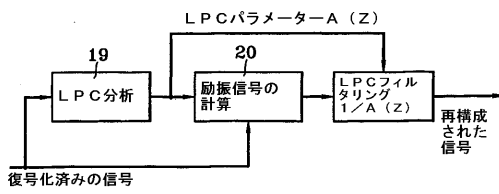
【 図 7 】

FIG.7



【 図 8 】

FIG.8



【国際公開パンフレット】

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété Intellectuelle  
Bureau international



(43) Date de la publication internationale  
14 mars 2002 (14.03.2002)

PCT

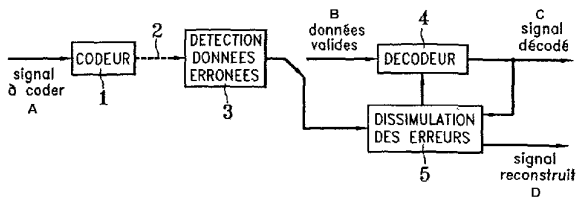
(10) Numéro de publication internationale  
WO 02/21515 A1

- (51) Classification internationale des brevets<sup>7</sup> : G10L 19/00, H04L 1/00
- (72) Inventeurs; et (75) Inventeurs/Déposants (pour US seulement) : KOVESI, Balazs [HU/FR], 12 Rue Corlay, F-22300 Lannion (FR). MASSALOUX, Dominique [FR/FR]; 53, rue du Pré de Saint-Maur, F-22700 Perros-Guirec (FR). DELEAM, David [FR/FR], 11, rue du Maréchal Leclerc, F-22700 Perros Guirec (FR).
- (21) Numéro de la demande internationale : PCT/FR01/02747
- (22) Date de dépôt international : 5 septembre 2001 (05.09.2001)
- (25) Langue de dépôt : français
- (74) Mandataires : MARTIN, Jean-Jacques etc.; Cabinet Regimbeau, 20, rue de Chazelles, F-75847 Paris Cedex 17 (FR).
- (26) Langue de publication : français
- (30) Données relatives à la priorité : 06/11285 5 septembre 2000 (05.09.2000) FR
- (81) États désignés (national) : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK,
- (71) Déposant (pour tous les États désignés sauf US) : FRANCE TELECOM [FR/FR], 6, place d'Alleray, F-75015 Paris (FR).

[Suite sur la page suivante]

(54) Title: TRANSMISSION ERROR CONCEALMENT IN AN AUDIO SIGNAL

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



WO 02/21515 A1

- A...SIGNAL TO BE CODED
- B...VALID DATA
- C...DECODED SIGNAL
- D...RECONSTRUCTED SIGNAL
- 1...ENCODER
- 3...DETECTING ERRONEOUS DATA
- 4...DECODER
- 5...ERROR CONCEALMENT

(57) Abstract: The invention concerns a method for concealing transmission errors in a digital audio signal which consists in receiving a decoded signal after transmission, storing the decoded samples when the transmitted data are valid, estimating at least a short-term prediction operator and a long-term prediction operator on the basis of the stored valid samples and in generating possible missing or erroneous samples in the decoded signals with the estimated operators. The invention is characterised in that it consists in controlling the energy of the synthesis signal thus generated with a calculated and adapted gain sample by sample.

[Suite sur la page suivante]



WO 02/21515 A1



SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,  
ZW.

**Publiée :**  
— avec rapport de recherche internationale

**(84) États désignés (régional) :** brevet ARIPO (GH, GM, KE,  
LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), brevet eurasien  
(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen  
(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,  
MC, NL, PT, SE, TR), brevet OAPI (BF, BJ, CF, CG, CI,  
CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

*En ce qui concerne les codes à deux lettres et autres abrévia-  
tions, se référer aux "Notes explicatives relatives aux codes et  
abréviations" figurant au début de chaque numéro ordinaire de  
la Gazette du PCT.*

**(57) Abrégé :** 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 un opérateur de prédiction à long terme en fonction des échantillons marquants ou erronés dans le signal décodé à l'aide des opérateurs ainsi estimés, 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.

WO 02/21515

PCT/FR01/02747

DISSIMULATION D'ERREURS DE TRANSMISSION DANS UN SIGNAL  
AUDIO

1. DOMAINE TECHNIQUE

5

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.

10

On distingue classiquement deux grandes catégories de codeurs :

- 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

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.

20

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]).

25

Pour tous ces codeurs, les valeurs codées sont ensuite transformées en un train binaire qui sera transmis sur un

30

35

WO 02/21515

PCT/FR01/02747

2

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 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.

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 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.

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 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

WO 02/21515

PCT/FR01/02747

3

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  
5 perdues.

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  
10 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.

## 15 2. ETAT DE LA TECHNIQUE ANTERIEURE

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],  
20 [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  
25 d'effacement 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  
30 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  
35 prédiction à long terme pour les sons voisés, par exemple). Du fait de cette corrélation, il est beaucoup

WO 02/21515

PCT/FR01/02747

4

plus avantageux de réutiliser les paramètres de la dernière trame valide pour synthétiser la trame effacée que d'utiliser des paramètres erronés ou aléatoires.

5 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 :

10 - 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]).

15 - 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 :

20 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 :

25 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  
30 l'excitation passée.

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,  
35 comme le module de synthèse du signal. Ils utilisent

WO 02/21515

PCT/FR01/02747

5

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.

5

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.

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, MAHEUX-2], visent à régénérer les coefficients transformés perdus à partir des valeurs prises par ces coefficients avant l'effacement.

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.

30

Enfin, il existe une famille de techniques de dissimulation des trames effacées développées conjointement avec le codage canal. Ces méthodes, telle celle décrite dans [FINGSCHIEDT], se servent d'informations fournies par le décodeur canal, par exemple d'informations concernant le degré de fiabilité

WO 02/21515

PCT/FR01/02747

6

des paramètres reçus. Elles sont fondamentalement différentes de la présente invention qui ne présuppose pas l'existence d'un codeur canal.

5 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  
10 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  
15 voisins, 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  
20 de l'effacement, ce qui engendrait également des artefacts gênants.

### 3. PRESENTATION DE L'INVENTION

25 L'invention 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.

30 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  
35 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

WO 02/21515

PCT/FR01/02747

7

des échantillons valides mémorisés et on génère d'éventuels échantillons manquants ou erronés dans le signal décodé à l'aide des opérateurs ainsi estimés.

5 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.

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

Notamment, le gain pour le contrôle du signal de  
15 synthèse est avantageusement 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  
20 spectre de fréquences.

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.

25 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  
30 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 correspondants à des sons non stationnaires.

35



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.

5

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.

10

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.

15

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 5.2.2.2.1 point 10).

20

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 limité à la composante non harmonique.

25

Notamment, la composante harmonique est avantageusement obtenue en mettant en oeuvre un filtrage

30

35

WO 02/21515

PCT/FR01/02747

9

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.

5

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.

10

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.

15

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.

20

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

25

Egalement, 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.

30

De façon préférentielle encore, on estime l'enveloppe spectrale du bruit des échantillons décodés

35

WO 02/21515

PCT/FR01/02747

10

valides et on génère un signal synthétisé évoluant vers un signal possédant la même enveloppe spectrale.

L'invention propose également un procédé de  
5 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  
10 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.

L'invention concerne en outre un dispositif de  
15 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.

20 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  
25 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.

30

#### 4. PRESENTATION DES FIGURES

D'autres caractéristiques et avantages de l'invention  
35 ressortiront encore de la description qui suit, laquelle

WO 02/21515

PCT/FR01/02747

11

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 ;

15 - 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 ;

20 - 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.

20

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

##### 5.1 Principe d'un mode de réalisation possible

25

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 30 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.

35 On notera que ce module 5, outre l'indication de données effacées, reçoit le signal décodé en période valide et

WO 02/21515

PCT/FR01/02747

12

transmet au décodeur des signaux utilisés pour sa mise à jour.

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

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);
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

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).

#### 5.1.2 Pendant un bloc de données effacées.

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

1. Estimation de l'enveloppe spectrale courante :

WO 02/21515

PCT/FR01/02747

13

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:

15 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.

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].

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

WO 02/21515

PCT/FR01/02747

14

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 :

10

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).

15

### 4. Synthèse des échantillons manquants :

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:

20

25

#### 4.1 En zone voisée :

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

30

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

35

3.

WO 02/21515

PCT/FR01/02747

15

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.

5

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.

10

4.2 En zone non voisée :

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.

20 4.3 Contrôle de l'amplitude du signal résiduel :

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.

25

5. Contrôle de l'énergie du signal de synthèse

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

30

35



WO 02/21515

PCT/FR01/02747

16

avant l'effacement (voir en 1), période fondamentale, et stationnarité locale du signal au moment de la coupure.

5 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.

10 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 15 actuelle). Cette information peut être également utilisée pour le calcul du gain adaptatif.

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

20 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 25 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 30 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).

35

5.1.3 Au rétablissement de la transmission

WO 02/21515

PCT/FR01/02747

17

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.

15 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).

20 5.1.4 Mise à jour des mémoires du décodeur

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.

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.

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.

35

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.

5 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.

10 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.

15

#### 5.2. Description d'exemples de réalisation particuliers

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

25

Système de codage/décodage numérique par transformée de type TDAC.

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

30

Trame de 20 ms (320 échantillons).

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

35

faisant la transformation inverse TDAC, on obtient une trame de sortie de 20 ms qui est la somme de la deuxième

WO 02/21515

PCT/FR01/02747

20

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

5 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

10 suivante (figure 6), pour la reconstruction de ces deux trames.

#### 5.2.2 Mise en oeuvre

15 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).

20

##### *5.2.2.1 En période valide*

Correspondant au paragraphe 5.1.2, on met à jour la mémoire des échantillons décodés. Cette mémoire est

25 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

30 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

35 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

WO 02/21515

PCT/FR01/02747

21

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

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

##### 5.2.2.2.1 Première trame binaire perdue après une période valide

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).

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  $\text{Corr}(T)$  correspondant à un retard  $T$  en n'utilisant que  $2 \cdot T_m$  échantillons à la fin du signal mémorisé :

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

25 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")  $\text{MaxPitch}$ .

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

WO 02/21515

PCT/FR01/02747

23

On calcule  $\text{Corr}(T)$  pour  $T = 2, \dots, \text{MaxPitch}$ . Si  $T'$  est le plus petit retard tel que  $\text{Corr}(T') < 0$  (on élimine ainsi les corrélations à très court terme), alors on cherche  $\text{MaxCorr}$ , maximum de  $\text{Corr}(T)$  pour  $T' < T \leq \text{MaxPitch}$ . Soit  $T_p$  la période correspondant à  $\text{MaxCorr}$  ( $\text{Corr}(T_p) = \text{MaxCorr}$ ). On cherche également  $\text{MaxCorrMP}$ , maximum de  $\text{Corr}(T)$  pour  $T' < T \leq 0.75 * \text{MinPitch}$ . Si  $T_p < \text{MinPitch}$  ou  $\text{MaxCorrMP} > 0.7 * \text{MaxCorr}$  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  $T_p = \text{MaxPitch}/2$ , et le coefficient de corrélation  $\text{MaxCorr}$  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  $\text{MinPitch}$  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  $\text{DiminFlag}=1$ ).

Dans le cas où  $\text{MaxCorr} > 0.6$ , 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  $\text{MaxCorrL} = \text{Corr}(T_1)$ . Si  $T_1 > \text{MinPitch}$  et  $\text{MaxCorrL} > 0.75 * \text{MaxCorr}$ , on choisit  $T_1$  comme nouvelle période fondamentale.

Si  $T_p$  est inférieur à  $\text{MaxPitch}/2$ , on peut vérifier s'il s'agit vraiment d'une trame voisée en cherchant le maximum local de la corrélation autour de  $2 * T_p$  ( $2TP$ ) et



WO 02/21515

PCT/FR01/02747

24

en vérifiant si  $\text{Corr}(T_p) > 0.4$ . Si  $\text{Corr}(T_p) < 0.4$  et si l'énergie du signal diminue, on met  $\text{DiminFlag}=1$  et on diminue la valeur de  $\text{MaxCorr}$ , sinon on cherche le maximum local suivant entré le  $T_p$  actuel et  $\text{MaxPitch}$ .

5

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.

10 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  $\text{DiminFlag}=1$  et on fait décroître la valeur de  $\text{MaxCorr}$  en fonction de degré de diminution.

15

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

20

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

25

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

6. Egalisation de l'énergie du signal résiduel. Dans le cas d'un signal non voisé ou faiblement voisé ( $\text{MaxCorr} < 0.7$ ), l'énergie du signal résiduel stocké dans  $\text{ResMem}$

35

WO 02/21515

PCT/FR01/02747

25

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 :

10   ▪ 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.

15   ▪ On cherche l'amplitude maximale MaxAmplSv de chaque sous-vecteur. Si  $\text{MaxAmplSv} > 1.5 * \text{MeanAmpl}$ , on multiplie le sous-vecteur par  $1.5 * \text{MeanAmpl} / \text{MaxAmplSv}$ .

20 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.

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

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

30 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

35 comme la partie entière d'une valeur réelle aléatoire

WO 02/21515

PCT/FR01/02747

26

Tpa. La valeur initiale de Tpa est égale à Tp 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 :

5     $exch(i) = -0.0635 * (exc(i-Tph-1) + exc(i-Tph+1)) +$   
           $0.1182 * exc(i-Tph) - 0.9926 * exch(i-1) -$   
           $0.7679 * exch(i-2)$

L'excitation voisée est alors la somme de ces 2

10 composantes :

Exc(i) = excb(i) + 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

15 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.

20 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.

25 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

30 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

WO 02/21515

PCT/FR01/02747

27

début du signal généré ne correspond pas bien au signal originel, comme expliqué ultérieurement (voir point 11).

- 5 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
- 10 buts :
- 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
  - 15 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).
  - On assure automatiquement la transition continue (sans
  - 20 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).
- 25 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
- 30 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  $\otimes$  TDAC inverse, on peut estimer la similitude entre la trame
- 35 perdue et la trame de remplacement. Une corrélation faible ( $< 0.65$ ) signale que le signal originel est assez

WO 02/21515

PCT/FR01/02747

28

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

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 :

- 20   ▪ 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.
- 25   ▪ 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 :
- 30   ▪ Interpolation progressive du filtre LPC avec un filtre plat pour rendre le signal synthétisé moins coloré.
- 35   ▪ Augmentation progressive de la valeur du pitch.

WO 02/21515

PCT/FR01/02747

29

- En mode voisé, on bascule en mode non-voisé après un certain temps (par exemple quand l'énergie minimale est atteinte).

5 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  
10 la figure 7, le module de synthèse de musique a été  
référéncé par 15, celui de synthèse de parole par 16 et  
le commutateur parole/musique par 17.

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

1. Estimation de l'enveloppe spectrale courante :

On calcule cette enveloppe spectrale sous la forme d'un  
filtre LPC [RABINER][KLEIJUN]. L'analyse est effectuée par  
20 des méthodes classiques ([KLEIJUN]). 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  
25 signaux musicaux.

2. Synthèse des échantillons manquants :

La synthèse des échantillons de remplacement s'effectue  
en introduisant un signal d'excitation dans le filtre de  
30 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  
35 8, l'étape de filtrage est référencée par 21.

WO 02/21515

PCT/FR01/02747

30

Exemple du contrôle de l'amplitude du signal résiduel :

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

5 Estimation du gain du filtre LPC:  
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  
10 rapport de ces deux énergies.  
Calcul de l'énergie cible :  
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  
15 l'analyse LPC).  
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.

20 3. Contrôle de l'énergie du signal de synthèse  
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) :  
25 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  
30 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.

WO 02/21515

PCT/FR01/02747

31

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

Comme pour les signaux de parole :

Dans le cas de périodes d'effacement relativement  
5 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  
10 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  
15 temps jusqu'à obtention du filtre du bruit) et de  
l'énergie (niveau évoluant progressivement vers celui du  
bruit, par exemple par fenêtrage).

#### 6. REMARQUE GENERALE

20 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  
25 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  
30 disponible quelle que soit la structure de codage  
utilisée.

#### 7. REFERENCES BIBLIOGRAPHIQUES

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



WO 02/21515

PCT/FR01/02747

32

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

WO 02/21515

PCT/FR01/02747

33

- [CHEN-5] J. H. Chen. "Excitation signal synthesis during frame erasure or packet loss". Brevet US5615298, EP0673017.
- 5 [CHEN-6] J. H. Chen. "Computational complexity reduction during frame erasure of packet loss". Brevet US5717822.
- [CHEN-7] J. H. Chen. "Computational complexity reduction during frame erasure or packet loss". Brevet US940212435, EP0673015.
- 10 [COX] R. V. Cox. "Three new speech coders from the ITU cover a range of applications". IEEE Communication Magazine, septembre 97, PP. 40-47.
- 15 [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.
- 20 [COMBESCURE] P. Combescure, 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.
- 25 [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.
- 30 [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.
- 35

WO 02/21515

PCT/FR01/02747

34

- [FINGSCHIEDT] T. Fingscheidt, P. Vary, "Robust speech decoding: a universal approach to bit error concealment", Proc. of ICASSP conference, 1997, pp. 1667-1670.
- 5 [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.
- 10 [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.
- 15 [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.
- [HELLWIG] K. Hellwig, P. Vary, D. Massaloux, J. P. Petit, 20 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. 25 Haavisto, R. Salami, C. Laflamme et J. P. Adoul. "GSM enhanced full rate speech codec ". Proc. of ICASSP conference, 1997, PP.771-774.
- [KROON] P. Kroon, B.S. Atal. "On the use of pitch 30 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 35 US5450449, EP0673016. generation during frame erasure or packet loss". Brevet US5450449, EP0673016.

WO 02/21515

PCT/FR01/02747

35

- [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.
- 5 [MAHIEUX-2] Y. Mahieux, "Dissimulation erreurs de transmission", brevet 92 06720 déposé le 3 juin 1992.
- [MAITRE] X. Maitre. "7 kHz audio coding within 64 kbit/s". IEEE Journal on Selected Areas on  
10 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  
15 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  
20
- [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  
25 "Silence compression scheme for dual rate speech coder for multimedia communications transmitting at 5.3 & 6.3 kbit/s"
- [SALAMI] R. Salami, C. Laflamme, J. P. Adoul, A. Kataoka,  
30 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.
- 35

WO 02/21515

PCT/FR01/02747

36

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

15

## REVENDICATIONS

1. 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 pour les sons voisés 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, 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.

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.

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.

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

généérés à la suite de données valides correspondants à des sons non stationnaires.

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.

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.

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.

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.

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.

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.

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.

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.

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



WO 02/21515

PCT/FR01/02747

40

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
- 10 signal décodé, caractérisé en ce que ce dispositif de dissimulation d'erreurs est un dispositif selon la revendication 17.

1/3

FIG. 1

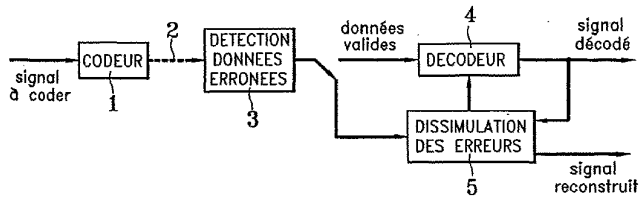


FIG. 2

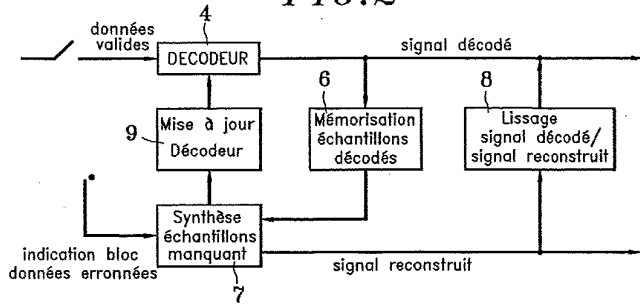
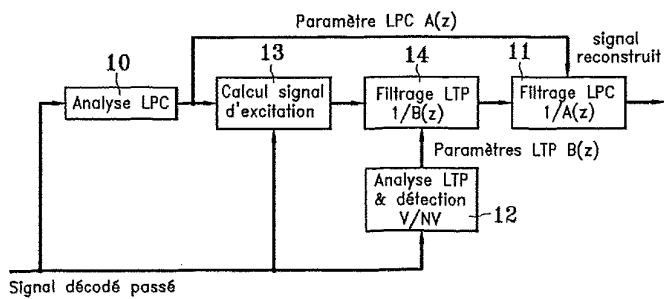


FIG. 3



2/3

FIG. 4

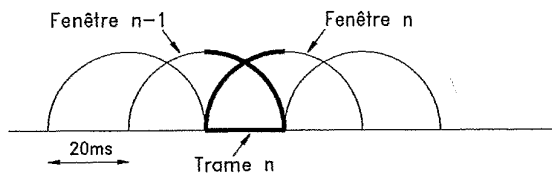


FIG. 5

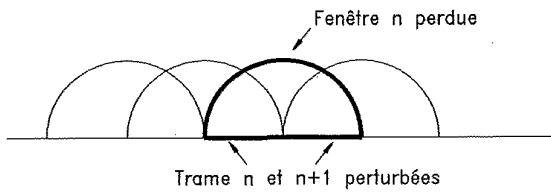


FIG. 6

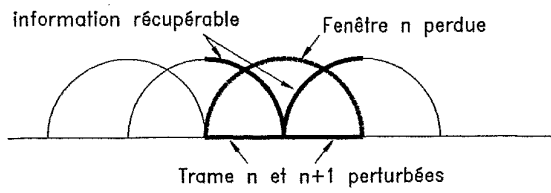


FIG. 7

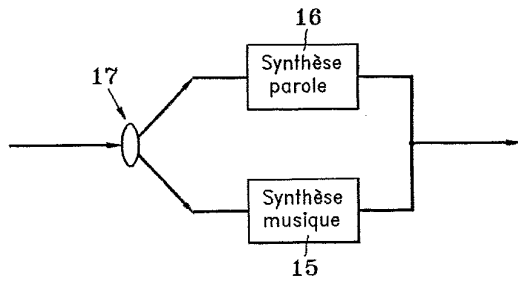
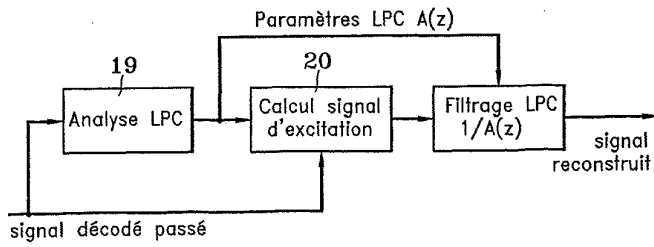


FIG. 8



## 【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/FR 01/02747
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 G10L19/00 H04L1/00		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 G10L H04L		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the International search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, INSPEC		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 884 010 A (CHEN ET AL) 16 March 1999 (1999-03-16) abstract; figures 3,5,9	1-18
A	US 5 717 822 A (CHEN) 10 February 1998 (1998-02-10) cited in the application abstract; figures 1,3 -/-	1-18
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents: *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *S* document member of the same patent family		
Date of the actual completion of the international search 19 November 2001		Date of mailing of the international search report 26/11/2001
Name and mailing address of the ISA European Patent Office, P.B. 8016 Patentlaan 2 NL - 2200 HV Rijswijk Tel. (+31-70) 346-3046, Tx. 31 651 epo NL Fax (+31-70) 346-3016		Authorized officer Quélavoine, R

Form PCT/ISA/210 (second sheet) (July 1999)

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/FR 01/02747

C/(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category <sup>o</sup>	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	COMBESURE 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 March 1999 (1999-03-15), pages 5-8, XP000898251 ISBN: 0-7803-6042-1 cited in the application abstract page 7, right-hand column, last paragraph -page 8, left-hand column, paragraph 1 ---	1-18
A	FR 2 774 827 A (FRANCE TELECOM) 13 August 1999 (1999-08-13) abstract; figures 1,2 ---	1,2,5-7, 17
A	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 July 1993 (1993-07-01), pages 295-303, XP000388573 ISSN: 1063-6676 abstract -----	1,2,4, 14,17

Form PCT/ISA/210 (continuation of second sheet) (July 1982)

INTERNATIONAL SEARCH REPORT			International Application No	
Information on patent family members			PCT/FR 01/02747	
Patent document cited in search report	Publication date	Patent family member(s)	Publication date	
US 5884010	A	16-03-1999	US 5574825 A	12-11-1996
			AU 683126 B2	30-10-1997
			AU 1367595 A	21-09-1995
			AU 685902 B2	29-01-1998
			AU 1471395 A	21-09-1995
			CA 2142398 A1	15-09-1995
			CA 2144102 A1	15-09-1995
			DE 69522979 D1	08-11-2001
			EP 0673018 A2	20-09-1995
			JP 7311596 A	28-11-1995
JP 8063200 A	08-03-1996			
US 5717822	A	10-02-1998	AU 683125 B2	30-10-1997
			AU 1367495 A	21-09-1995
			CA 2142391 A1	15-09-1995
			EP 0673015 A2	20-09-1995
			JP 7325594 A	12-12-1995
FR 2774827	A	13-08-1999	FR 2774827 A1	13-08-1999
			AU 2170699 A	23-08-1999
			BR 9904776 A	08-03-2000
			CN 1263625 T	16-08-2000
			EP 1051703 A1	15-11-2000
			WO 9940573 A1	12-08-1999
			JP 2001511917 T	14-08-2001

RAPPORT DE RECHERCHE INTERNATIONALE		Demande Internationale No PCT/FR 01/02747
A. CLASSEMENT DE L'OBJET DE LA DEMANDE CIB 7 G10L19/00 H04L1/00		
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB		
B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE Documentation minimale consultée (système de classification suivi des symboles de classement) CIB 7 G10L H04L		
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche		
Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés) EPO-Internal, MPI Data, INSPEC		
C. DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	US 5 884 010 A (CHEN ET AL) 16 mars 1999 (1999-03-16) abrégé; figures 3,5,9 ---	1-18
A	US 5 717 822 A (CHEN) 10 février 1998 (1998-02-10) cité dans la demande abrégé; figures 1,3 ---	1-18
	-/-	
<input checked="" type="checkbox"/> Voir la suite du cadre C pour le fin de la liste des documents <input checked="" type="checkbox"/> Les documents de familles de brevets sont indiqués en annexe		
* Catégories spéciales de documents cités:		
*A* document définissant l'état général de la technique, non considéré comme particulièrement pertinent		** document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la filière conduisant la base de l'invention
*E* document antérieur, mais publié à la date de dépôt international ou après cette date		*X* document particulièrement pertinent: l'inventeur revendiqué ne peut être considéré comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément
*I* document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (elle qu'indiquée)		*Y* document particulièrement pertinent: l'inventeur revendiqué ne peut être considéré comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier
*O* document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens		*Z* document qui fait partie de la même famille de brevets
*P* document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée		
Date à laquelle la recherche internationale a été effectivement achevée	Date d'expédition du présent rapport de recherche internationale	
19 novembre 2001	26/11/2001	
Nom et adresse postale de l'administration chargée de la recherche internationale Office Européen des Brevets, P.B. 5818 Patentamt 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2340, Tx. 31 051 epo nl, Fax: (+31-70) 340-3010	Fonctionnaire autorisé  Quélavoine, R	

Formulaire PCT/ISA/210 (système tenu) (juillet 1992)



## RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale No  
PCT/FR 01/02747

D.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	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 abrégé page 7, colonne de droite, dernier alinéa -page 8, colonne de gauche, alinéa 1	1-18
A	FR 2 774 827 A (FRANCE TELECOM) 13 août 1999 (1999-08-13) abrégé; figures 1,2	1,2,5-7, 17
A	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 abrégé	1,2,4, 14,17

Formulaire PCT/ISA210 suite de la deuxième feuille (juillet 1992)

## RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande Internationale No

PCT/FR 01/02747

Document brevet cité au rapport de recherche		Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 5884010	A	16-03-1999	US 5574825 A	12-11-1996
			AU 683126 B2	30-10-1997
			AU 1367595 A	21-09-1995
			AU 685902 B2	29-01-1998
			AU 1471395 A	21-09-1995
			CA 2142398 A1	15-09-1995
			CA 2144102 A1	15-09-1995
			DE 69522979 D1	08-11-2001
			EP 0673018 A2	20-09-1995
			JP 7311596 A	28-11-1995
US 5717822	A	10-02-1998	AU 683125 B2	30-10-1997
			AU 1367495 A	21-09-1995
			CA 2142391 A1	15-09-1995
			EP 0673015 A2	20-09-1995
			JP 7325594 A	12-12-1995
FR 2774827	A	13-08-1999	FR 2774827 A1	13-08-1999
			AU 2170699 A	23-08-1999
			BR 9904776 A	08-03-2000
			CN 1263625 T	16-08-2000
			EP 1051703 A1	15-11-2000
			WO 9940573 A1	12-08-1999
			JP 2001511917 T	14-08-2001

---

フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PH,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

(72)発明者 デレアム,ダヴィッド

フランス共和国,エフ - 2 2 7 0 0 ペロス - ギレク,リュ デュ マレシャル ルクレール,1  
1

Fターム(参考) 5D045 AA02