

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2004 年 11 月 11 日 (11.11.2004)

PCT

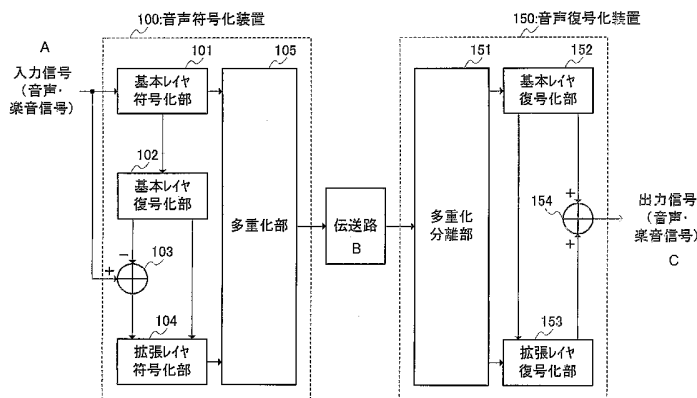
(10) 国際公開番号
WO 2004/097796 A1

- (51) 国際特許分類: **G10L 19/04**, 19/12, H03M 7/30 (74) 代理人: 鷺田 公一 (WASHIDA, Kimihito); 〒2060034 東京都多摩市鶴牧1丁目24-1 新都市センタービル5階 Tokyo (JP).
- (21) 国際出願番号: PCT/JP2004/006294
- (22) 国際出願日: 2004 年 4 月 30 日 (30.04.2004)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願2003-125665 2003 年 4 月 30 日 (30.04.2003) JP
- (71) 出願人 (米国を除く全ての指定国について): 松下電器産業株式会社 (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) [JP/JP]; 〒5718501 大阪府門真市大字門真1006番地 Osaka (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ユーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE,

[続葉有]

(54) Title: AUDIO ENCODING DEVICE, AUDIO DECODING DEVICE, AUDIO ENCODING METHOD, AND AUDIO DECODING METHOD

(54) 発明の名称: 音声符号化装置、音声復号化装置及びこれらの方法



100...AUDIO ENCODING DEVICE
A...INPUT SIGNAL (AUDIO/MUSIC COMPOSITION SIGNAL)
101...BASIC LAYER ENCODING SECTION
102...BASIC LAYER DECODING SECTION
104...EXTENDED LAYER ENCODING SECTION
105...MULTIPLEXING SECTION
B...TRANSMISSION PATH
150...AUDIO DECODING DEVICE
151...MULTIPLEXING/DEMULTIPLEXING SECTION
152...BASIC LAYER DECODING SECTION
153...EXTENDED LAYER DECODING SECTION
C...OUTPUT SIGNAL (AUDIO/MUSIC COMPOSITION SIGNAL)

(57) Abstract: A basic layer encoding section (101) encodes an input signal and acquires basic layer encoded information. A basic layer decoding section (102) decodes the basic layer encoded information and acquires a basic layer decoded signal and long-term prediction information (pitch lag). An adder (103) reverses polarity of the basic layer decoded signal and adds it to the input signal so as to acquire a residual signal. An extended layer encoding section (104) encodes a long-term prediction coefficient calculated by using the long-term prediction information and the residual signal

[続葉有]



IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

添付公開書類:
— 国際調査報告書

and acquires extended layer encoded information. A basic layer decoding section (152) decodes the basic layer encoded information and acquires a basic layer decoded signal and long-term prediction information. An extended layer decoding section (153) uses the long-term prediction information to decode the extended layer encoded information and acquires an extended layer decoded signal. An adder (154) adds the basic layer decoded signal to the extended layer decoded signal so as to acquire audio/music composition signal. Thus, it is possible to realize scalable encoding with a small amount of calculation and a small amount of encoded information.

(57) 要約: 基本レイヤ符号化部101は入力信号を符号化し基本レイヤ符号化情報を得る。基本レイヤ復号化部102は基本レイヤ符号化情報を復号化し基本レイヤ復号化信号及び長期予測情報(ピッチラグ)を得る。加算部103は入力信号に基本レイヤ復号化信号を極性反転して加算し残差信号を得る。拡張レイヤ符号化部104は長期予測情報及び残差信号を用いて算出した長期予測係数を符号化し拡張レイヤ符号化情報を得る。基本レイヤ復号化部152は基本レイヤ符号化情報を復号化し基本レイヤ復号化信号及び長期予測情報を得る。拡張レイヤ復号化部153は長期予測情報を利用して拡張レイヤ符号化情報を復号化し拡張レイヤ復号化信号を得る。加算部154は基本レイヤ復号化信号と拡張レイヤ復号化信号とを加算し音声・楽音信号を得る。これにより、少ない計算量及び符号化情報量でスケーラブル符号化を実現することができる。

明 細 書

音声符号化装置、音声復号化装置及びこれらの方法

5 技術分野

本発明は、音声・楽音信号を符号化して伝送する通信システムに使用される音声符号化装置、音声復号化装置及びこれらの方法に関する。

背景技術

- 10 デジタル無線通信や、インターネット通信に代表されるパケット通信、あるいは音声蓄積などの分野においては、電波などの伝送路容量や記憶媒体の有効利用を図るため、音声信号の符号化／復号化技術が不可欠であり、これまでに多くの音声符号化／復号化方式が開発されてきた。その中で、CELP方式の音声符号化／復号化方式が主流の方式として実用化されている。
- 15 CELP方式の音声符号化装置は、予め記憶された音声モデルに基づいて入力音声をコード化する。具体的には、デジタル化された音声信号を20ms程度のフレームに区切り、フレーム毎に音声信号の線形予測分析を行い、線形予測係数と線形予測残差ベクトルを求め、線形予測係数と線形予測残差ベクトルをそれぞれ個別に符号化する。
- 20 低ビットレートの通信を実行するためには、記憶できる音声モデルの量が限られるため、従来のCELPタイプの音声符号化／復号化方式では、主に発声音のモデルを記憶している。
- また、インターネット通信のようなパケットを伝送する通信システムでは、ネットワークの状態によりパケット損失が起こるため、符号化情報の一部が欠損した場合であっても符号化情報の残りの一部から音声、楽音を復号化できる
- 25 ことが望ましい。同様に、通信容量に応じてビットレートを変化させる可変レート通信システムにおいては、通信容量が低下した場合、符号化情報の一部の

みを伝送することにより通信容量の負担を軽減させることが容易であることが望ましい。このように、符号化情報の全てもしくは符号化情報の一部のみを用いて音声、楽音を復号化できる技術として、最近、スケーラブル符号化技術が注目を浴びている。従来にもいくつかのスケーラブル符号化方式が開示されて

5 いる。

スケーラブル符号化方式は、一般的に、基本レイヤと拡張レイヤとからなり、各レイヤは、基本レイヤを最も下位のレイヤとし、階層構造を形成している。そして、各レイヤでは、より下位のレイヤの入力信号と出力信号との差である残差信号について符号化が行われる。この構成により、全レイヤの符号化情報
10 もしくは下位レイヤの符号化情報のみを用いて、音声・楽音信号を復号化することができる。

しかしながら、従来のスケーラブル符号化方式では、基本レイヤおよび拡張レイヤの符号化方式としてCELPタイプの音声符号化／復号化方式を用いるため、計算量、符号化情報共に相応の量が必要となる。

15

発明の開示

本発明の目的は、少ない計算量及び符号化情報量でスケーラブル符号化を実現することができる音声符号化装置、音声復号化装置及びこれらの方法を提供することである。

20 この目的は、長期予測を行う拡張レイヤを設け、音声・楽音の長期的な相関の性質を利用して拡張レイヤにおいて残差信号の長期予測を行うことにより復号化信号の品質の向上を図り、基本レイヤの長期予測情報を利用して長期予測ラグを求めることにより演算量の削減を図ることにより達成される。

25 図面の簡単な説明

図1は、本発明の実施の形態1に係る音声符号化装置／音声復号化装置の構成を示すブロック図、

図 2 は、上記実施の形態に係る基本レイヤ符号化部の内部構成を示すブロック図、

図 3 は、上記実施の形態に係る基本レイヤ符号化部の内のパラメータ決定部が適応音源符号帳から生成される信号を決定する処理を説明するための図、

5 図 4 は、上記実施の形態に係る基本レイヤ復号化部の内部構成を示すブロック図、

図 5 は、上記実施の形態に係る拡張レイヤ符号化部の内部構成を示すブロック図、

10 図 6 は、上記実施の形態に係る拡張レイヤ復号化部の内部構成を示すブロック図、

図 7 は、本発明の実施の形態 2 に係る拡張レイヤ符号化部の内部構成を示すブロック図、

図 8 は、上記実施の形態に係る拡張レイヤ復号化部の内部構成を示すブロック図、及び、

15 図 9 は、本発明の実施の形態 3 に係る音声信号送信装置／音声信号受信装置の構成を示すブロック図である。

発明を実施するための最良の形態

以下、本発明の実施の形態について、図面を用いて説明する。なお、以下の
20 各本実施の形態では、基本レイヤと拡張レイヤとで構成される二階層の音声符号化／復号化方法において拡張レイヤで長期予測を行う場合について説明する。ただし、本発明は階層について制限はなく、三階層以上の階層的な音声符号化／復号化方法において下位レイヤの長期予測情報を利用して上位レイヤで長期予測を行う場合についても適用することができる。階層的な音声符号化方法と
25 は、残差信号（下位レイヤの入力信号と下位レイヤの復号化信号との差）を長期予測によって符号化して符号化情報を出力する音声符号化方法が上位レイヤに複数存在して階層構造を成している方法である。また、階層的な音声復号化

方法とは、残差信号を復号化する音声復号化方法が上位レイヤに複数存在して階層構造を成している方法である。ここで、最下のレイヤに存在する音声・楽音符号化／復号化方法を基本レイヤとする。また、基本レイヤより上位レイヤに存在する音声・楽音符号化／復号化方法を拡張レイヤとする。

- 5 また、本発明の各実施の形態では、基本レイヤがC E L Pタイプの音声符号化／復号化を行う場合を例にして説明する。

(実施の形態1)

図1は、本発明の実施の形態1に係る音声符号化装置／音声復号化装置の構成を示すブロック図である。

- 10 図1において、音声符号化装置100は、基本レイヤ符号化部101と、基本レイヤ復号化部102と、加算部103と、拡張レイヤ符号化部104と、多重化部105とから主に構成される。また、音声復号化装置150は、多重化分離部151と、基本レイヤ復号化部152と、拡張レイヤ復号化部153と、加算部154とから主に構成される。

- 15 基本レイヤ符号化部101は、音声・楽音信号を入力し、C E L Pタイプの音声符号化方法を用いて入力信号を符号化し、符号化によって求められる基本レイヤ符号化情報を基本レイヤ復号化部102に出力するとともに、多重化部105に出力する。

- 20 基本レイヤ復号化部102は、C E L Pタイプの音声復号化方法を用いて基本レイヤ符号化情報を復号化し、復号化によって求められる基本レイヤ復号化信号を加算部103に出力する。また、基本レイヤ復号化部102は、ピッチラグを基本レイヤの長期予測情報として拡張レイヤ符号化部104に出力する。

- 25 ここで、「長期予測情報」とは、音声・楽音信号が有する長期的な相関を表す情報である。また、「ピッチラグ」とは、基本レイヤで特定される位置情報であり、詳細な説明は後述する。

加算部103は、入力信号に、基本レイヤ復号化部102から出力された基本レイヤ復号化信号を極性反転してから加算し、加算結果である残差信号を拡

張レイヤ符号化部 104 に出力する。

拡張レイヤ符号化部 104 は、基本レイヤ復号化部 102 より出力された長期予測情報及び加算部 103 より出力された残差信号を用いて長期予測係数を算出し、長期予測係数を符号化し、符号化によって求められる拡張レイヤ符号化情報を多重化部 105 に出力する。

多重化部 105 は、基本レイヤ符号化部 101 から出力された基本レイヤ符号化情報と、拡張レイヤ符号化部 104 から出力された拡張レイヤ符号化情報と、を多重化して多重化情報として伝送路を介して多重化分離部 151 に出力する。

- 10 多重化分離部 151 は、音声符号化装置 100 から伝送された多重化情報を、基本レイヤ符号化情報と拡張レイヤ符号化情報とに分離し、分離された基本レイヤ符号化情報を基本レイヤ復号化部 152 に出力し、また、分離された拡張レイヤ符号化情報を拡張レイヤ復号化部 153 に出力する。

- 15 基本レイヤ復号化部 152 は、CELP タイプの音声復号化方法を用いて基本レイヤ符号化情報を復号化し、復号化によって求められる基本レイヤ復号化信号を加算部 154 に出力する。また、基本レイヤ復号化部 152 は、ピッチラグを基本レイヤの長期予測情報として拡張レイヤ復号化部 153 に出力する。

- 20 拡張レイヤ復号化部 153 は、長期予測情報を利用して拡張レイヤ符号化情報を復号化し、復号化によって求められる拡張レイヤ復号化信号を加算部 154 に出力する。

加算部 154 は、基本レイヤ復号化部 152 から出力された基本レイヤ復号化信号と拡張レイヤ復号化部 153 から出力された拡張レイヤ復号化信号とを加算し、加算結果である音声・楽音信号を後工程の装置に出力する。

- 25 次に、図 1 の基本レイヤ符号化部 101 の内部構成を図 2 のブロック図を用いて説明する。

基本レイヤ符号化部 101 の入力信号は、前処理部 200 に入力される。前処理部 200 は、DC 成分を取り除くハイパスフィルタ処理や後続する符号化

6

処理の性能改善につながるような波形整形処理やプリアンファシス処理を行い、これらの処理後の信号 (X_{in}) をLPC分析部201および加算器204に出力する。

5 LPC分析部201は、 X_{in} を用いて線形予測分析を行い、分析結果（線形予測係数）をLPC量子化部202へ出力する。LPC量子化部202は、LPC分析部201から出力された線形予測係数(LPC)の量子化処理を行い、量子化LPCを合成フィルタ203へ出力するとともに量子化LPCを表す符号(L)を多重化部213へ出力する。

10 合成フィルタ203は、量子化LPCに基づくフィルタ係数により、後述する加算器210から出力される駆動音源に対してフィルタ合成を行うことにより合成信号を生成し、合成信号を加算器204へ出力する。

加算器204は、合成信号の極性を反転させて X_{in} に加算することにより誤差信号を算出し、誤差信号を聴覚重み付け部211へ出力する。

15 適応音源符号帳205は、過去に加算器210によって出力された駆動音源信号をバッファに記憶しており、パラメータ決定部212から出力された信号によって特定される過去の駆動音源信号サンプルから1フレーム分のサンプルを適応音源ベクトルとして切り出して乗算器208へ出力する。

量子化利得生成部206は、パラメータ決定部212から出力された信号によって特定される適応音源利得と固定音源利得とをそれぞれ乗算器208と209へ出力する。

固定音源符号帳207は、パラメータ決定部212から出力された信号によって特定される形状を有するパルス音源ベクトルに拡散ベクトルを乗算して得られた固定音源ベクトルを乗算器209へ出力する。

25 乗算器208は、量子化利得生成部206から出力された量子化適応音源利得を、適応音源符号帳205から出力された適応音源ベクトルに乗じて、加算器210へ出力する。乗算器209は、量子化利得生成部206から出力された量子化固定音源利得を、固定音源符号帳207から出力された固定音源ベク

トルに乗じて、加算器 210 へ出力する。

加算器 210 は、利得乗算後の適応音源ベクトルと固定音源ベクトルとをそれぞれ乗算器 208 と乗算器 209 から入力し、これらをベクトル加算し、加算結果である駆動音源を合成フィルタ 203 および適応音源符号帳 205 へ出力する。なお、適応音源符号帳 205 に入力された駆動音源は、バッファに記憶される。

聴覚重み付け部 211 は、加算器 204 から出力された誤差信号に対して聴覚的な重み付けをおこない、聴覚重み付け領域での X_{in} と合成信号との歪みを算出し、パラメータ決定部 212 へ出力する。

10 パラメータ決定部 212 は、聴覚重み付け部 211 から出力された符号化歪みを最小とする適応音源ベクトル、固定音源ベクトル及び量子化利得を、各々適応音源符号帳 205、固定音源符号帳 207 及び量子化利得生成部 206 から選択し、選択結果を示す適応音源ベクトル符号 (A)、音源利得符号 (G) 及び固定音源ベクトル符号 (F) を多重化部 213 に出力する。なお、適応音
15 源ベクトル符号 (A) は、ピッチラグに対応する符号である。

多重化部 213 は、LPC 量子化部 202 から量子化 LPC を表す符号 (L) を入力し、パラメータ決定部 212 から適応音源ベクトルを表す符号 (A)、固定音源ベクトルを表す符号 (F) および量子化利得を表す符号 (G) を入力し、これらの情報を多重化して基本レイヤ符号化情報として出力する。

20 以上が、図 1 の基本レイヤ符号化部 101 の内部構成の説明である。

次に、図 3 を用いて、パラメータ決定部 212 が、適応音源符号帳 205 から生成される信号を決定する処理を簡単に説明する。図 3 において、バッファ 301 は適応音源符号帳 205 が備えるバッファであり、位置 302 は適応音源ベクトルの切り出し位置であり、ベクトル 303 は、切り出された適応音源
25 ベクトルである。また、数値「41」、「296」は、切り出し位置 302 を動かす範囲の下限と上限とに対応している。

切り出し位置 302 を動かす範囲は、適応音源ベクトルを表す符号 (A) に

割り当てるビット数を「8」とする場合、「2 5 6」の長さの範囲（例えば、4 1～2 9 6）に設定することができる。また、切り出し位置 3 0 2 を動かす範囲は、任意に設定することができる。

- 5 パラメータ決定部 2 1 2 は、切り出し位置 3 0 2 を設定された範囲内で動かし、適応音源ベクトル 3 0 3 をそれぞれフレームの長さだけ切り出す。そして、パラメータ決定部 2 1 2 は、聴覚重み付け部 2 1 1 から出力される符号化歪みが最小となる切り出し位置 3 0 2 を求める。

このように、パラメータ決定部 2 1 2 によって求められるバッファの切り出し位置 3 0 2 が「ピッチラグ」である。

- 10 次に、図 1 の基本レイヤ復号化部 1 0 2（1 5 2）の内部構成について図 4 を用いて説明する。

- 図 4 において、基本レイヤ復号化部 1 0 2（1 5 2）に入力された基本レイヤ符号化情報は、多重化分離部 4 0 1 によって個々の符号（L、A、G、F）に分離される。分離された L P C 符号（L）は L P C 復号化部 4 0 2 に出力され、分離された適応音源ベクトル符号（A）は適応音源符号帳 4 0 5 に出力され、分離された音源利得符号（G）は量子化利得生成部 4 0 6 に出力され、分離された固定音源ベクトル符号（F）は固定音源符号帳 4 0 7 へ出力される。
- 15

L P C 復号化部 4 0 2 は、多重化分離部 4 0 1 から出力された符号（L）から L P C を復号し、合成フィルタ 4 0 3 に出力する。

- 20 適応音源符号帳 4 0 5 は、多重化分離部 4 0 1 から出力された符号（A）で指定される過去の駆動音源信号サンプルから 1 フレーム分のサンプルを適応音源ベクトルとして取り出して乗算器 4 0 8 へ出力する。また、適応音源符号帳 4 0 5 は、ピッチラグを長期予測情報として拡張レイヤ符号化部 1 0 4（拡張レイヤ復号化部 1 5 3）に出力する。

- 25 量子化利得生成部 4 0 6 は、多重化分離部 4 0 1 から出力された音源利得符号（G）で指定される適応音源ベクトル利得と固定音源ベクトル利得を復号し乗算器 4 0 8 及び乗算器 4 0 9 へ出力する。

固定音源符号帳407は、多重化分離部401から出力された符号(F)で指定される固定音源ベクトルを生成し、乗算器409へ出力する。

乗算器408は、適応音源ベクトルに適応音源ベクトル利得を乗算して、加算器410へ出力する。乗算器409は、固定音源ベクトルに固定音源ベクトル利得を乗算して、加算器410へ出力する。

加算器410は、乗算器408、409から出力された利得乗算後の適応音源ベクトルと固定音源ベクトルの加算を行って駆動音源ベクトルを生成し、これを合成フィルタ403及び適応音源符号帳405に出力する。

合成フィルタ403は、加算器410から出力された駆動音源ベクトルを駆動信号として、LPC復号化部402によって復号されたフィルタ係数を用いて、フィルタ合成を行い、合成した信号を後処理部404へ出力する。

後処理部404は、合成フィルタ403から出力された信号に対して、ホルマント強調やピッチ強調といったような音声の主観的な品質を改善する処理や、定常雑音の主観的な品質を改善する処理などを施し、基本レイヤ復号化信号として出力する。

以上が、図1の基本レイヤ復号化部102(152)の内部構成の説明である。

次に、図1の拡張レイヤ符号化部104の内部構成について図5のブロック図を用いて説明する。

拡張レイヤ符号化部104では、残差信号をNサンプルずつ区切り(Nは自然数)、Nサンプルを1フレームとしてフレーム毎に符号化を行う。以下、残差信号を $e(0) \sim e(X-1)$ と表し、符号化の対象となるフレームを $e(n) \sim e(n+N-1)$ と表すこととする。ここで、Xは残差信号の長さであり、Nはフレームの長さに相当する。また、nは各フレームの先頭に位置するサンプルであり、nはNの整数倍に相当する。なお、あるフレームの信号を過去に生成された信号から予測して生成する方法は長期予測と呼ばれる。また、長期予測を行うフィルタはピッチフィルタ、コムフィルタ等と呼ばれる。

図5において、長期予測ラグ指示部501は、基本レイヤ復号化部102で求められる長期予測情報 t を入力し、これに基づいて拡張レイヤの長期予測ラグ T を求め、これを長期予測信号記憶部502に出力する。なお、基本レイヤと拡張レイヤとの間でサンプリング周波数の違いが生じる場合、長期予測ラグ

5 T は、以下の式(1)により求めることができる。なお、式(1)において、 D は拡張レイヤのサンプリング周波数、 d は基本レイヤのサンプリング周波数である。

$$T = D \times t / d \quad \cdots \text{式(1)}$$

長期予測信号記憶部502は、過去に生成された長期予測信号を記憶するバッファを備える。バッファの長さを M とした場合、バッファは過去に生成された長期予測信号の系列 $s(n-M-1) \sim s(n-1)$ で構成される。長期予測信号記憶部502は、長期予測ラグ指示部501より長期予測ラグ T を入力すると、バッファに記憶されている過去の長期予測信号の系列から長期予測ラグ T だけ遡った長期予測信号 $s(n-T) \sim s(n-T+N-1)$ を切り出し、

10 これを長期予測係数計算部503及び長期予測信号生成部506に出力する。また、長期予測信号記憶部502は、長期予測信号生成部506から長期予測信号 $s(n) \sim s(n+N-1)$ を入力し、以下の式(2)によりバッファの更新を行う。

$$\begin{aligned} \hat{s}(i) &= s(i+N) \quad (i=n-M-1, \dots, n-1) \\ s(i) &= \hat{s}(i) \quad (i=n-M-1, \dots, n-1) \end{aligned} \quad \cdots \text{式(2)}$$

20 なお、長期予測ラグ T がフレーム長 N より短く、長期予測信号記憶部502が長期予測信号を切り出すことができない場合、長期予測ラグ T をフレーム長 N より長くなるまで整数倍することにより長期予測信号を切り出すことができる。あるいは、長期予測ラグ T だけ遡った長期予測信号 $s(n-T) \sim s(n-T+N-1)$ を繰り返して、フレーム長 N の長さまで充当させることにより

25 切り出すことができる。

長期予測係数計算部503は、残差信号 $e(n) \sim e(n+N-1)$ 及び長

期予測信号 $s(n-T) \sim s(n-T+N-1)$ を入力し、これらを用いて以下の式(3)により、長期予測係数 β を算出し、これを長期予測係数符号化部504に出力する。

$$\beta = \frac{\sum_{i=0}^{N-1} e(n+i)s(n-T+i)}{\sum_{i=0}^{N-1} s(n-T+i)^2} \quad \cdots \text{式(3)}$$

- 5 長期予測係数符号化部504は、長期予測係数 β を符号化し、符号化によって求められる拡張レイヤ符号化情報を長期予測係数復号化部505に出力し、伝送路を介して拡張レイヤ復号化部153に出力する。なお、長期予測係数 β の符号化方法として、スカラ量子化により行う方法等が知られている。

- 長期予測係数復号化部505は、拡張レイヤ符号化情報を復号化し、これに
10 よって求められる復号化長期予測係数 β_q を長期予測信号生成部506に出力する。

- 長期予測信号生成部506は、復号化長期予測係数 β_q 及び長期予測信号 $s(n-T) \sim s(n-T+N-1)$ を入力し、これらを用いて以下の式(4)により、長期予測信号 $s(n) \sim s(n+N-1)$ を算出し、これを長期予測
15 信号記憶部502に出力する。

$$s(n+i) = \beta_q \times s(n-T+1) \quad (i=0, \dots, N-1) \quad \cdots \text{式(4)}$$

以上が、図1の拡張レイヤ符号化部104の内部構成の説明である。

次に、図1の拡張レイヤ復号化部153の内部構成について図6のブロック図を用いて説明する。

- 20 図6において、長期予測ラグ指示部601は、基本レイヤ復号化部152から出力された長期予測情報を用いて拡張レイヤの長期予測ラグ T を求め、これを長期予測信号記憶部602に出力する。

長期予測信号記憶部602は、過去に生成された長期予測信号を記憶するバッファを備える。バッファの長さを M とした場合、バッファは過去に生成され

た長期予測信号の系列 $s(n-M-1) \sim s(n-1)$ で構成される。長期予測信号記憶部 602 は、長期予測ラグ指示部 601 より長期予測ラグ T を入力すると、バッファに記憶されている過去の長期予測信号の系列から長期予測ラグ T だけ遡った長期予測信号 $s(n-T) \sim s(n-T+N-1)$ を切り出し、

5 これを長期予測信号生成部 604 に出力する。また、長期予測信号記憶部 602 は、長期予測信号生成部 604 から長期予測信号 $s(n) \sim s(n+N-1)$ を入力し、上記式 (2) によりバッファの更新を行う。

長期予測係数復号化部 603 は、拡張レイヤ符号化情報を復号化し、復号化によって求められる復号化長期予測係数 β_q を長期予測信号生成部 604 に出力する。

10

長期予測信号生成部 604 は、復号化長期予測係数 β_q 及び長期予測信号 $s(n-T) \sim s(n-T+N-1)$ を入力し、これらを用いて上記式 (4) により、長期予測信号 $s(n) \sim s(n+N-1)$ を算出し、これを長期予測信号記憶部 602 及び加算部 153 に拡張レイヤ復号化信号として出力する。

15 以上が、図 1 の拡張レイヤ復号化部 153 の内部構成の説明である。

このように、長期予測を行う拡張レイヤを設け、音声・楽音の長期的な相関の性質を利用して残差信号を拡張レイヤにおいて長期予測することにより、少ない符号化情報で周波数帯域の広い音声・楽音信号を効果的に符号化／復号化することができ、また、演算量の削減を図ることができる。

20 このとき、長期予測ラグを符号化／復号化するのではなく、基本レイヤの長期予測情報を利用して長期予測ラグを求めることにより、符号化情報の削減を図ることができる。

また、基本レイヤ符号化情報を復号化することによって、基本レイヤの復号化信号のみを得ることができ、CELP タイプの音声符号化／復号化方法において、符号化情報の一部からでも音声・楽音を復号化できる機能（スケーラブル符号化）を実現することができる。

25

また、長期予測においては、音声・楽音が有する長期的な相関を利用し、現

フレームとの相関が最も高いフレームをバッファから切り出し、切り出したフレームの信号を用いて現フレームの信号を表現する。しかしながら、現フレームとの相関が最も高いフレームをバッファから切り出す手段において、ピッチラグなどの音声・楽音が有する長期的な相関を表わす情報が無い場合には、バ

- 5 ッファからフレームを切り出す際の切り出し位置を変化させながら、切り出したフレームと現フレームとの自己相関関数を計算し、最も相関が高くなるフレームを探索する必要がある、探索に掛かる計算量は非常に大きくなってしま

ところが、基本レイヤ符号化部 101 で求めたピッチラグを用いて切り出し位置を一意に定めることにより、通常の長期予測を行う際に掛かる計算量を大

- 10 幅に削減することができる。

なお、本実施の形態で説明した拡張レイヤ長期予測方法では、基本レイヤ復号化部より出力される長期予測情報がピッチラグである場合について説明したが、本発明はこれに限られず、音声・楽音が有する長期的な相関を表す情報であれば長期予測情報として用いることができる。

- 15 また、本実施の形態では、長期予測信号記憶部 502 がバッファから長期予測信号を切り出す位置を長期予測ラグ T とする場合について説明したが、これを長期予測ラグ T 付近の位置 $T + \alpha$ (α は微小な数であり、任意に設定可能) とする場合についても本発明は適用することができ、長期予測ラグ T に微小な誤差が生じる場合でも本実施の形態と同様の作用・効果を得ることができる。

- 20 例えば、長期予測信号記憶部 502 は、長期予測ラグ指示部 501 より長期予測ラグ T を入力し、バッファに記憶されている過去の長期予測信号の系列から $T + \alpha$ だけ遡った長期予測信号 $s(n - T - \alpha) \sim s(n - T - \alpha + N - 1)$ を切り出し、以下の式 (5) を用いて判定値 C を算出し、判定値 C が最大となる α を求め、これを符号化する。復号化を行う場合、長期予測信号記憶部 60
- 25 2 は、 α の符号化情報を復号化して α を求め、また、長期予測ラグ T を用いて長期予測信号 $s(n - T - \alpha) \sim s(n - T - \alpha + N - 1)$ を切り出す。

$$C = \frac{\left[\sum_{i=0}^{N-1} e(n+i) s(n-T-\alpha+i) \right]^2}{\sum_{i=0}^{N-1} s(n-T-\alpha+i)^2} \quad \dots \text{式(5)}$$

- また、本実施の形態では、音声・楽音信号を用いて長期予測を行う場合について説明したが、MDCT、QMF等の直交変換を用いて音声・楽音信号を時間領域から周波数領域へ変換し、変換後の信号（周波数パラメータ）を用いて
- 5 長期予測を行う場合についても本発明は適用することができ、本実施の形態と同様の作用・効果を得ることができる。例えば、音声・楽音信号の周波数パラメータで拡張レイヤ長期予測を行う場合には、図5において、長期予測係数計算部503に、長期予測信号 $s(n-T) \sim s(n-T+N-1)$ を時間領域から周波数領域へ変換する機能及び残差信号を周波数パラメータへ変換する機能
- 10 能を新たに設け、長期予測信号生成部506に、長期予測信号 $s(n) \sim s(n+N-1)$ を周波数領域から時間領域へ逆変換する機能を新たに設ける。また、図6において、長期予測信号生成部604に、長期予測信号 $s(n) \sim s(n+N-1)$ を周波数領域から時間領域へ逆変換する機能を新たに設ける。

- また、通常の音声・楽音符号化／復号化方法では、伝送路において誤り検出
- 15 もしくは誤り訂正に用いる冗長ビットを符号化情報に付加させて、冗長ビットを含む符号化情報を伝送することが一般的であるが、本発明では、基本レイヤ符号化部101より出力される符号化情報（A）と拡張レイヤ符号化部104より出力される符号化情報（B）とに割り当てる冗長ビットのビット配分を符号化情報（A）に重みを付けて振り分けることができる。

20 (実施の形態2)

実施の形態2では、残差信号と長期予測信号との差（長期予測残差信号）の符号化／復号化を行う場合について説明する。

- 本実施の形態の音声符号化装置／音声復号化装置は、構成が図1と同様であり、拡張レイヤ符号化部104及び拡張レイヤ復号化部153の内部構成のみ
- 25 が異なる。

図7は、本実施の形態に係る拡張レイヤ符号化部104の内部構成を示すブロック図である。なお、図7において、図5と共通する構成部分には図5と同一符号を付して説明を省略する。

図7の拡張レイヤ符号化部104は、図5と比較して、加算部701、長期
5 予測残差信号符号化部702、符号化情報多重化部703、長期予測残差信号復号化部704及び加算部705を追加した構成を採る。

長期予測信号生成部506は、算出した長期予測信号 $s(n) \sim s(n+N-1)$ を加算部701及び加算部705に出力する。

加算部701は、以下の式(6)に示すように、長期予測信号 $s(n) \sim s$
10 $(n+N-1)$ の極性を反転させて残差信号 $e(n) \sim e(n+N-1)$ に加算し、加算結果である長期予測残差信号 $p(n) \sim p(n+N-1)$ を長期予測残差信号符号化部702に出力する。

$$p(n+i) = e(n+i) - s(n+i) \quad (i=0, \dots, N-1) \quad \dots \text{式(6)}$$

長期予測残差信号符号化部702は、長期予測残差信号 $p(n) \sim p(n+N-1)$ の符号化を行い、符号化によって求められる符号化情報（以下、「長期予測残差符号化情報」という）を符号化情報多重化部703及び長期予測残差信号復号化部704に出力する。なお、長期予測残差信号の符号化は、ベクトル量子化が一般的である。

ここで、長期予測残差信号 $p(n) \sim p(n+N-1)$ の符号化方法について8ビットでベクトル量子化を行う場合を例に説明する。この場合、長期予測
20 残差信号符号化部702の内部には、予め作成された256種類のコードベクトルが格納されたコードブックが用意される。このコードベクトル $\text{CODE}(k)$ $(0) \sim \text{CODE}(k)(N-1)$ は、Nの長さのベクトルである。また、kはコードベクトルのインデクスであり、0から255までの値をとる。長期予測
25 残差信号符号化部702は、以下の式(8)により長期予測残差信号 $p(n) \sim p(n+N-1)$ とコードベクトル $\text{CODE}(k)(0) \sim \text{CODE}(k)(N-1)$ との二乗誤差 er を求める。

$$er = \sum_{i=0}^{N-1} (p(n+i) - CODE^{(k)}(i))^2 \quad \cdots \text{式(7)}$$

そして、長期予測残差信号符号化部 7 0 2 は、二乗誤差 er が最小となる k の値を長期予測残差符号化情報として決定する。

5 符号化情報多重化部 7 0 3 は、長期予測係数符号化部 5 0 4 より入力した拡張レイヤ符号化情報と、長期予測残差信号符号化部 7 0 2 より入力した長期予測残差符号化情報を多重化し、多重化後の情報を伝送路を介して拡張レイヤ復号化部 1 5 3 に出力する。

10 長期予測残差信号復号化部 7 0 4 は、長期予測残差符号化情報の復号化を行い、復号化によって求められた復号化長期予測残差信号 $p_q(n) \sim p_q(n+N-1)$ を加算部 7 0 5 に出力する。

15 加算部 7 0 5 は、長期予測信号生成部 5 0 6 より入力した長期予測信号 $s(n) \sim s(n+N-1)$ と長期予測残差信号復号化部 7 0 4 より入力した復号化長期予測残差信号 $p_q(n) \sim p_q(n+N-1)$ とを加算し、加算結果を長期予測信号記憶部 5 0 2 に出力する。この結果、長期予測信号記憶部 5 0 2 は、以下の式 (8) によりバッファの更新を行う。

$$\left. \begin{aligned} \hat{s}(i) &= s(i+N) \quad (i=n-M-1, \dots, n-N-1) \\ \hat{s}(i) &= s(i+N) + p, (i-N) \quad (i=n-N, \dots, n-1) \\ s(i) &= \hat{s}(i) \quad (i=n-M-1, \dots, n-1) \end{aligned} \right\} \quad \cdots \text{式(8)}$$

以上が、本実施の形態に係る拡張レイヤ符号化部 1 0 4 の内部構成の説明である。

20 次に、本実施の形態に係る拡張レイヤ復号化部 1 5 3 の内部構成について、図 8 のブロック図を用いて説明する。なお、図 8 において、図 6 と共通する構成部分には図 6 と同一符号を付して説明を省略する。

図 8 の拡張レイヤ復号化部 1 5 3 は、図 6 と比較して、符号化情報分離部 8 0 1、長期予測残差信号復号化部 8 0 2 及び加算部 8 0 3 を追加した構成を採る。

符号化情報分離部 801 は、伝送路より受信した多重化されている符号化情報を、拡張レイヤ符号化情報と長期予測残差符号化情報とに分離し、拡張レイヤ符号化情報を長期予測係数復号化部 603 に出力し、長期予測残差符号化情報を長期予測残差信号復号化部 802 に出力する。

- 5 長期予測残差信号復号化部 802 は、長期予測残差符号化情報を復号化して復号化長期予測残差信号 $p_q(n) \sim p_q(n+N-1)$ を求め、これを加算部 803 に出力する。

- 加算部 803 は、長期予測信号生成部 604 より入力した長期予測信号 $s(n) \sim s(n+N-1)$ と長期予測残差信号復号化部 802 より入力した復号化長期予測残差信号 $p_q(n) \sim p_q(n+N-1)$ とを加算し、加算結果を長期予測信号記憶部 602 に出力し、加算結果を拡張レイヤ復号化信号として出力する。
- 10

以上が、本実施の形態に係る拡張レイヤ復号化部 153 の内部構成の説明である。

- 15 このように、残差信号と長期予測信号との差（長期予測残差信号）を符号化／復号化することにより、上記実施の形態 1 よりもさらに高品質な復号化信号を得ることができる。

- なお、本実施の形態では、ベクトル量子化により長期予測残差信号の符号化を行う場合について説明したが、本発明は符号化方法に制限はなく、例えば、
- 20 形状・利得 VQ、分割 VQ、変換 VQ、多段階 VQ により符号化を行ってもよい。

- 以下、13 ビットで形状 8 ビット、利得 5 ビットの形状・利得 VQ により符号化を行う場合について説明する。この場合、コードブックは形状コードブック、利得コードブックの二種類が用意される。形状コードブックは 256 種類の形状コードベクトルから成り、形状コードベクトル $SCODE(k1)(0) \sim$
- 25 $SCODE(k1)(N-1)$ は、N の長さのベクトルである。ここで、 $k1$ は形状コードベクトルのインデクスであり、0 から 255 までの値をとる。また、利得コードブックは 32 種類の利得コードから成り、利得コード $GCODE(k$

- 2) はスカラの値をとる。ここで、 k_2 は利得コードのインデクスであり、0 から 31 までの値をとる。長期予測残差信号符号化部 702 は、以下の式 (9) により長期予測残差信号 $p(n) \sim p(n+N-1)$ の利得 **gain** と形状ベクトル **shape**(0) $\sim$ **shape**(N-1) を求め、以下の式 (10) により利得 **gain** と利得コード **GCODE**(k_2) との利得誤差 **gainer** と、形状ベクトル **shape**(0) $\sim$ **shape**(N-1) と形状コードベクトル **SCODE**(k_1)(0) $\sim$ **SCODE**(k_1)(N-1) との二乗誤差 **shapeer** とを求める。

$$\text{gain} = \sqrt{\sum_{i=0}^{N-1} p(n+i)^2} \quad \dots \text{式(9)}$$

$$\text{shape}(i) = \frac{p(n+i)}{\text{gain}} \quad (i=0, \dots, N-1)$$

$$\text{gainer} = |\text{gain} - \text{GCODE}^{(k_2)}| \quad \dots \text{式(10)}$$

$$\text{shapeer} = \sum_{i=0}^{N-1} (\text{shape}(i) - \text{SCODE}^{(k_1)}(i))^2$$

- 10 そして、長期予測残差信号符号化部 702 は、利得誤差 **gainer** が最小となる k_2 の値と二乗誤差 **shapper** が最小となる k_1 の値とを求め、これらの求めた値を長期予測残差符号化情報とする。

- 次に、8ビットで分割VQにより符号化を行う場合について説明する。この場合、コードブックは第1分割コードブック、第2分割コードブックの二種類
15 が用意される。第1分割コードブックは16種類の第1分割コードベクトル **SPCODE**(k_3)(0) $\sim$ **SPCODE**(k_3)(N/2-1) から成り、第2分割コードブック **SPCODE**(k_4)(0) $\sim$ **SPCODE**(k_4)(N/2-1) は16種類の第2分割コードベクトルから成り、それぞれコードベクトルはN/2の長さのベクトルである。ここで、 k_3 は第1分割コードベクトルのイン
20 デクスであり、0から15までの値をとる。また、 k_4 は第2分割コードベクトルのインデクスであり、0から15までの値をとる。長期予測残差信号符号化部 702 は、以下の式 (11) により長期予測残差信号 $p(n) \sim p(n+$

- N-1) を、第1分割ベクトル $s_{p1}(0) \sim s_{p1}(N/2-1)$ と第2分割ベクトル $s_{p2}(0) \sim s_{p2}(N/2-1)$ とに分割し、以下の式(12)により第1分割ベクトル $s_{p1}(0) \sim s_{p1}(N/2-1)$ と第1分割コードベクトル $SPCODE(k3)(0) \sim SPCODE(k3)(N/2-1)$ との
- 5 二乗誤差 **spliter1** と、第2分割ベクトル $s_{p2}(0) \sim s_{p2}(N/2-1)$ と第2分割コードブック $SPCODE(k4)(0) \sim SPCODE(k4)(N/2-1)$ との二乗誤差 **spliter2** とを求める。

$$\begin{aligned} sp_1(i) &= p(n+1) \quad (i=0, \dots, N/2-1) \\ sp_2(i) &= p(n+N/2+i) \quad (i=0, \dots, N/2-1) \end{aligned} \quad \dots \text{式(11)}$$

$$\begin{aligned} \text{spliter}_1 &= \sum_{i=0}^{N/2-1} (sp_1(i) - SPCODE_{k3}^{(k3)}(i))^2 \\ \text{spliter}_2 &= \sum_{i=0}^{N/2-1} (sp_2(i) - SPCODE_{k4}^{(k4)}(i))^2 \end{aligned} \quad \dots \text{式(12)}$$

- 10 そして、長期予測残差信号符号化部702は、二乗誤差 **spliter1** が最小となる $k3$ の値と二乗誤差 **spliter2** が最小となる $k4$ の値とを求め、これらの求めた値を長期予測残差信号符号化情報とする。

- 次に、8ビットで離散フーリエ変換を用いた変換VQにより符号化を行う場合について説明する。この場合、256種類の変換コードベクトルから成る変換コードブックが用意され、変換コードベクトル $TCODE(k5)(0) \sim TCODE(k5)(N/2-1)$ はNの長さのベクトルである。ここで、 $k5$ は変換コードベクトルのインデクスであり、0から255までの値をとる。長期予測残差信号符号化部702は、以下の式(13)により長期予測残差信号
- 15 $p(n) \sim p(n+N-1)$ を離散フーリエ変換して変換ベクトル $tp(0) \sim tp(N-1)$ を求め、以下の式(14)により変換ベクトル $tp(0) \sim tp(N-1)$ と変換コードベクトル $TCODE(k5)(0) \sim TCODE(k5)(N/2-1)$ との二乗誤差 **transer** を求める。

$$tp(i) = \sum_{n=0}^{N-1} p(n+i) e^{-j \frac{2\pi \sigma i}{N}} \quad (i=0, \dots, N-1) \quad \dots \text{式(13)}$$

$$transer = \sum_{i=0}^{N-1} (tp(i) - TCODE^{(k3)}(i))^2 \quad \dots \text{式(14)}$$

そして、長期予測残差信号符号化部 702 は、二乗誤差 **transer** が最小となる k_5 の値を求め、この値を長期予測残差信号化情報とする。

次に、13ビットで一段目5ビット、二段目8ビットの二段VQにより符号化を行う場合について説明する。この場合、一段目コードブック、二段目コードブックの二種類のコードブックを用意する。一段目コードブックは32種類の一段目コードベクトル **PHCODE 1** (k_6) (0) ~ **PHCODE 1** (k_6) (N-1) から成り、二段目コードブックは256種類の二段目コードベクトル **PHCODE 2** (k_7) (0) ~ **PHCODE 2** (k_7) (N-1) から成り、それぞれコードベクトルはNの長さのベクトルである。ここで、 k_6 は一段目コードベクトルのインデクスであり、0から31までの値をとる。また、 k_7 は二段目コードベクトルのインデクスであり、0から255までの値をとる。長期予測残差信号符号化部 702 は、以下の式 (15) により長期予測残差信号 $p(n) \sim p(n+N-1)$ と一段目コードベクトル **PHCODE 1** (k_6) (0) ~ **PHCODE 1** (k_6) (N-1) との二乗誤差 **phaseer 1** を求め、二乗誤差 **phaseer 1** が最小となる k_6 の値を求め、この値を k_{max} とする。

$$phaseer_1 = \sum_{i=0}^{N-1} (tp(i) - TCODE^{(k3)}(i))^2 \quad \dots \text{式(15)}$$

そして、長期予測残差信号符号化部 702 は、以下の式 (16) により誤差ベクトル **ep** (0) ~ **ep** (N-1) を求め、以下の式 (17) により誤差ベクトル **ep** (0) ~ **ep** (N-1) と二段目コードベクトル **PHCODE 2** (k_7) (0) ~ **PHCODE 2** (k_7) (N-1) との二乗誤差 **phaseer 2** を求め、二乗誤差 **phaseer 2** が最小となる k_7 の値を求め、この値と k_{max} とを長期予測残差信号化情報とする。

$$ep(i) = p(n+i) - PHCODE_1^{(k_{max})}(i) \quad (i=0, \dots, N-1) \quad \dots \text{式(16)}$$

$$phase_{er_2} = \sum_{i=0}^{N-1} (ep(i) - PHCODE_{E_2}^{(K3)}(i))^2 \quad \cdots \text{式(17)}$$

(実施の形態 3)

図 9 は、上記実施の形態 1、2 で説明した音声符号化装置及び音声復号化装置を含む音声信号送信装置および音声信号受信装置の構成を示すブロック図である。

図 9 において、音声信号 901 は入力装置 902 によって電氣的信号に変換され A/D 変換装置 903 に出力される。A/D 変換装置 903 は入力装置 902 から出力された（アナログ）信号をデジタル信号に変換し音声符号化装置 904 へ出力する。音声符号化装置 904 は、図 1 に示した音声符号化装置 100 を実装し、A/D 変換装置 903 から出力されたデジタル音声信号を符号化し符号化情報を RF 変調装置 905 へ出力する。RF 変調装置 905 は音声符号化装置 904 から出力された音声符号化情報を電波等の伝播媒体に載せて送出するための信号に変換し送信アンテナ 906 へ出力する。送信アンテナ 906 は RF 変調装置 905 から出力された出力信号を電波（RF 信号）として送出する。なお、図中の RF 信号 907 は送信アンテナ 906 から送出された電波（RF 信号）を表す。以上が音声信号送信装置の構成および動作である。

RF 信号 908 は受信アンテナ 909 によって受信され RF 復調装置 910 に出力される。なお、図中の RF 信号 908 は受信アンテナ 909 に受信された電波を表し、伝播路において信号の減衰や雑音の重畳がなければ RF 信号 907 と全く同じものになる。

RF 復調装置 910 は受信アンテナ 909 から出力された RF 信号から音声符号化情報を復調し音声復号化装置 911 へ出力する。音声復号化装置 911 は、図 1 に示した音声復号化装置 150 を実装し、RF 復調装置 910 から出力された音声符号化情報から音声信号を復号し D/A 変換装置 912 へ出力する。D/A 変換装置 912 は音声復号化装置 911 から出力されたデジタル

音声信号をアナログの電氣的信号に変換し出力装置 9 1 3 へ出力する。

出力装置 9 1 3 は電氣的信号を空気の振動に変換し音波として人間の耳に聴こえるように出力する。なお、図中、参照符号 9 1 4 は出力された音波を表す。以上が音声信号受信装置の構成および動作である。

- 5 無線通信システムにおける基地局装置および通信端末装置に、上記のような音声信号送信装置および音声信号受信装置を備えることにより、高品質な復号化信号を得ることができる。

- 10 以上説明したように、本発明によれば、少ない符号化情報で周波数帯域の広い音声・楽音信号を効果的に符号化／復号化することができ、また、演算量の削減を図ることができる。また、基本レイヤの長期予測情報を利用して長期予測ラグを求めることにより、符号化情報を削減することができる。また、基本レイヤ符号化情報を復号化することによって、基本レイヤの復号化信号のみを得ることができ、C E L P タイプの音声符号化／復号化方法において、符号化情報の一部からでも音声・楽音を復号化できる機能（スケーラブル符号化）を
- 15 実現することができる。

本明細書は、2003年4月30日出願の特願2003-125665に基づくものである。この内容をここに含めておく。

産業上の利用可能性

- 20 本発明は、音声・楽音信号を符号化して伝送する通信システムに使用される音声符号化装置、音声復号化装置に用いるに好適である。

請 求 の 範 囲

1. 入力信号を符号化して第1符号化情報を生成する基本レイヤ符号化手段と、
前記第1符号化情報を復号化して第1復号化信号を生成するとともに、音声・
楽音が有する長期的な相関を表す情報である長期予測情報を生成する基本レイ
5 ヤ復号化手段と、前記入力信号と前記第1復号化信号との差分である残差信号
を求める加算手段と、前記長期予測情報及び前記残差信号を用いて長期予測係
数を算出し、前記長期予測係数を符号化して第2符号化情報を生成する拡張レイ
ヤ符号化手段と、を具備する音声符号化装置。
2. 基本レイヤ復号化手段は、駆動音源信号サンプルから切り出された適応音
10 源ベクトルの切り出し位置を示す情報を長期予測情報とする請求項1記載の音
声符号化装置。
3. 拡張レイヤ符号化手段は、前記長期予測情報に基づいて拡張レイヤの長期
予測ラグを求める手段と、バッファに記憶されている過去の長期予測信号系列
から前記長期予測ラグだけ遡った長期予測信号を切り出す手段と、前記残差信
15 号及び前記長期予測信号を用いて長期予測係数を算出する手段と、前記長期予
測係数を符号化することにより前記拡張レイヤ符号化情報を生成する手段と、
前記拡張レイヤ符号化情報を復号化して復号化長期予測係数を生成する手段と、
前記復号化長期予測係数及び前記長期予測信号を用いて新たな長期予測信号を
算出し、前記新たな長期予測信号を用いて前記バッファを更新する手段と、を
20 有する請求項1記載の音声符号化装置。
4. 拡張レイヤ符号化手段は、前記残差信号と前記長期予測信号との差分であ
る長期予測残差信号を求める手段と、前記長期予測残差信号を符号化すること
により長期予測残差符号化情報を生成する手段と、前記長期予測残差符号化情
報を復号化して復号化長期予測残差信号算出する手段と、前記新たな長期予測
25 信号と前記復号化長期予測残差信号とを加算し、加算結果を用いて前記バッ
ファを更新する手段と、をさらに有する請求項3記載の音声符号化装置。
5. 請求項1記載の音声符号化装置から第1符号化情報及び第2符号化情報を

受信して音声を復号化する音声復号化装置であって、

前記第 1 符号化情報を復号化して第 1 復号化信号を生成するとともに、音声・楽音が有する長期的な相関を表す情報である長期予測情報を生成する基本レイヤ復号化手段と、前記長期予測情報を用いて前記第 2 符号化情報を復号化
5 して第 2 復号化信号を生成する拡張レイヤ復号化手段と、前記第 1 復号化信号と前記第 2 復号化信号とを加算し、加算結果である音声・楽音信号を出力する加算手段と、を具備する音声復号化装置。

6. 基本レイヤ復号化手段は、駆動音源信号サンプルから切り出された適応音源ベクトルの切り出し位置を示す情報を長期予測情報とする請求項 5 記載の音声復号化装置。
10

7. 拡張レイヤ復号化手段は、前記長期予測情報に基づいて拡張レイヤの長期予測ラグを求める手段と、バッファに記憶されている過去の長期予測信号系列から長期予測ラグだけ遡った長期予測信号を切り出す手段と、前記拡張レイヤ符号化情報を復号化して復号化長期予測係数を求める手段と、前記復号化長期
15 予測係数及び長期予測信号を用いて長期予測信号を算出し、前記長期予測信号を用いて前記バッファを更新する手段と、を有し、前記長期予測信号を拡張レイヤ復号化信号とする請求項 5 記載の音声復号化装置。

8. 拡張レイヤ復号化手段は、前記長期予測残差符号化情報を復号化して復号化長期予測残差信号を求める手段と、前記長期予測信号と前記復号化長期予測
20 残差信号とを加算する手段と、を有し、前記加算結果を拡張レイヤ復号化信号とする請求項 7 記載の音声復号化装置。

9. 音声符号化装置を具備する音声信号送信装置であって、

前記音声符号化装置は、入力信号を符号化して第 1 符号化情報を生成する基本レイヤ符号化手段と、前記第 1 符号化情報を復号化して第 1 復号化信号を生成するとともに、音声・楽音が有する長期的な相関を表す情報である長期予測
25 情報を生成する基本レイヤ復号化手段と、前記入力信号と前記第 1 復号化信号との差分である残差信号を求める加算手段と、前記長期予測情報及び前記残差

信号を用いて長期予測係数を算出し、前記長期予測係数を符号化して第2符号化情報を生成する拡張レイヤ符号化手段と、を具備する。

10 10. 請求項1記載の音声符号化装置から第1符号化情報及び第2符号化情報を受信して音声を復号化する音声復号化装置を具備する音声信号受信装置であ
5 5 って、

前記第1符号化情報を復号化して第1復号化信号を生成するとともに、音声・楽音が有する長期的な相関を表す情報である長期予測情報を生成する基本レイヤ復号化手段と、前記長期予測情報を用いて前記第2符号化情報を復号化して第2復号化信号を生成する拡張レイヤ復号化手段と、前記第1復号化信号
10 10 と前記第2復号化信号とを加算し、加算結果である音声・楽音信号を出力する加算手段と、を具備する。

11. 入力信号を符号化して第1符号化情報を生成する工程と、前記第1符号化情報を復号化して第1復号化信号を生成するとともに、音声・楽音が有する長期的な相関を表す情報である長期予測情報を生成する工程と、前記入力信号
15 15 と前記第1復号化信号との差分である残差信号を求める工程と、前記長期予測情報及び前記残差信号を用いて長期予測係数を算出し、前記長期予測係数を符号化して第2符号化情報を生成する工程と、を具備する音声符号化方法。

12. 請求項11記載の音声符号化方法で生成された第1符号化情報及び第2符号化情報を用いて音声を復号化する音声復号化方法であって、

20 20 前記第1符号化情報を復号化して第1復号化信号を生成するとともに、音声・楽音が有する長期的な相関を表す情報である長期予測情報を生成する工程と、前記長期予測情報を用いて前記第2符号化情報を復号化して第2復号化信号を生成する工程と、前記第1復号化信号と前記第2復号化信号とを加算し、加算結果である音声・楽音信号を出力する工程と、を具備する音声復号化方法。

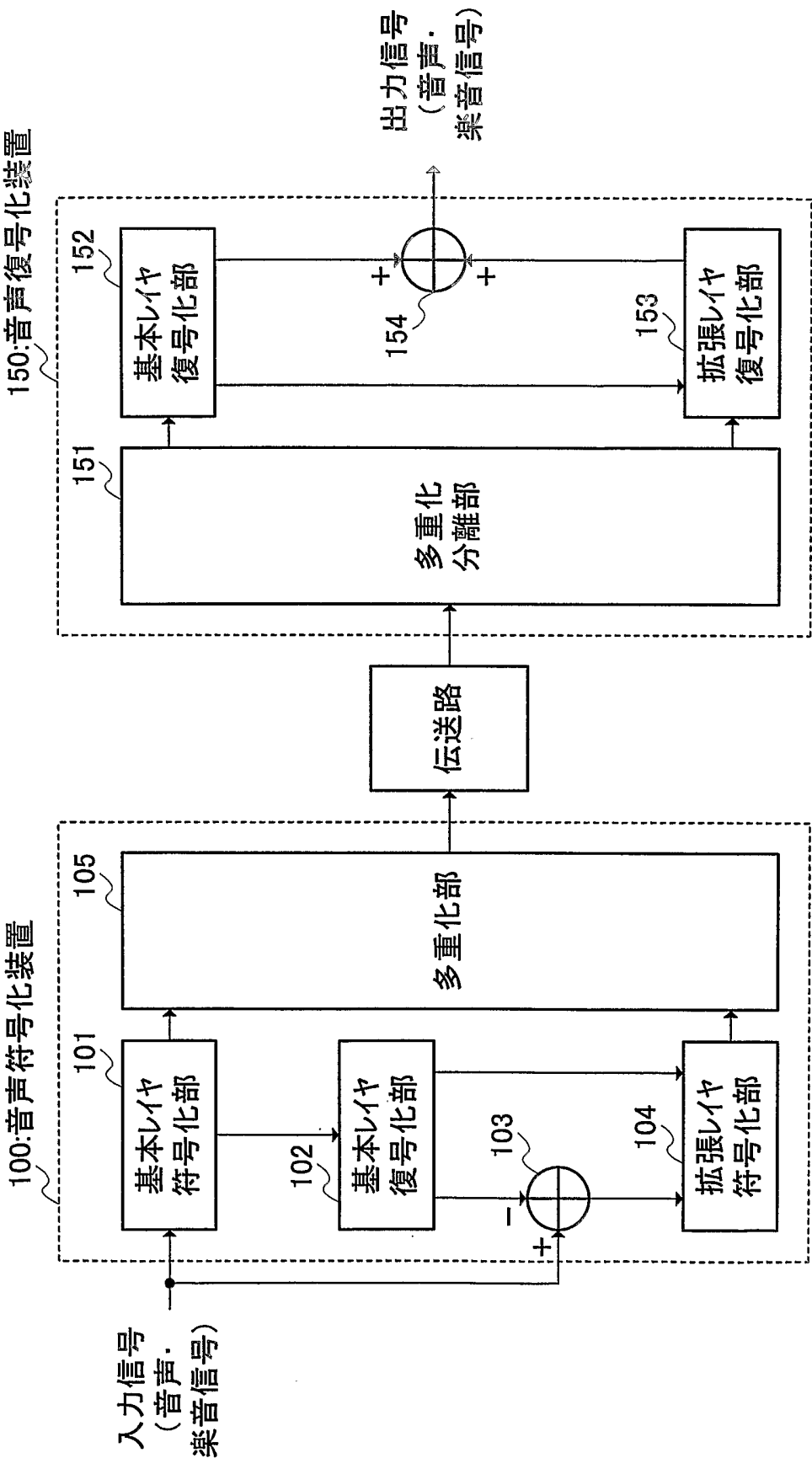


図1

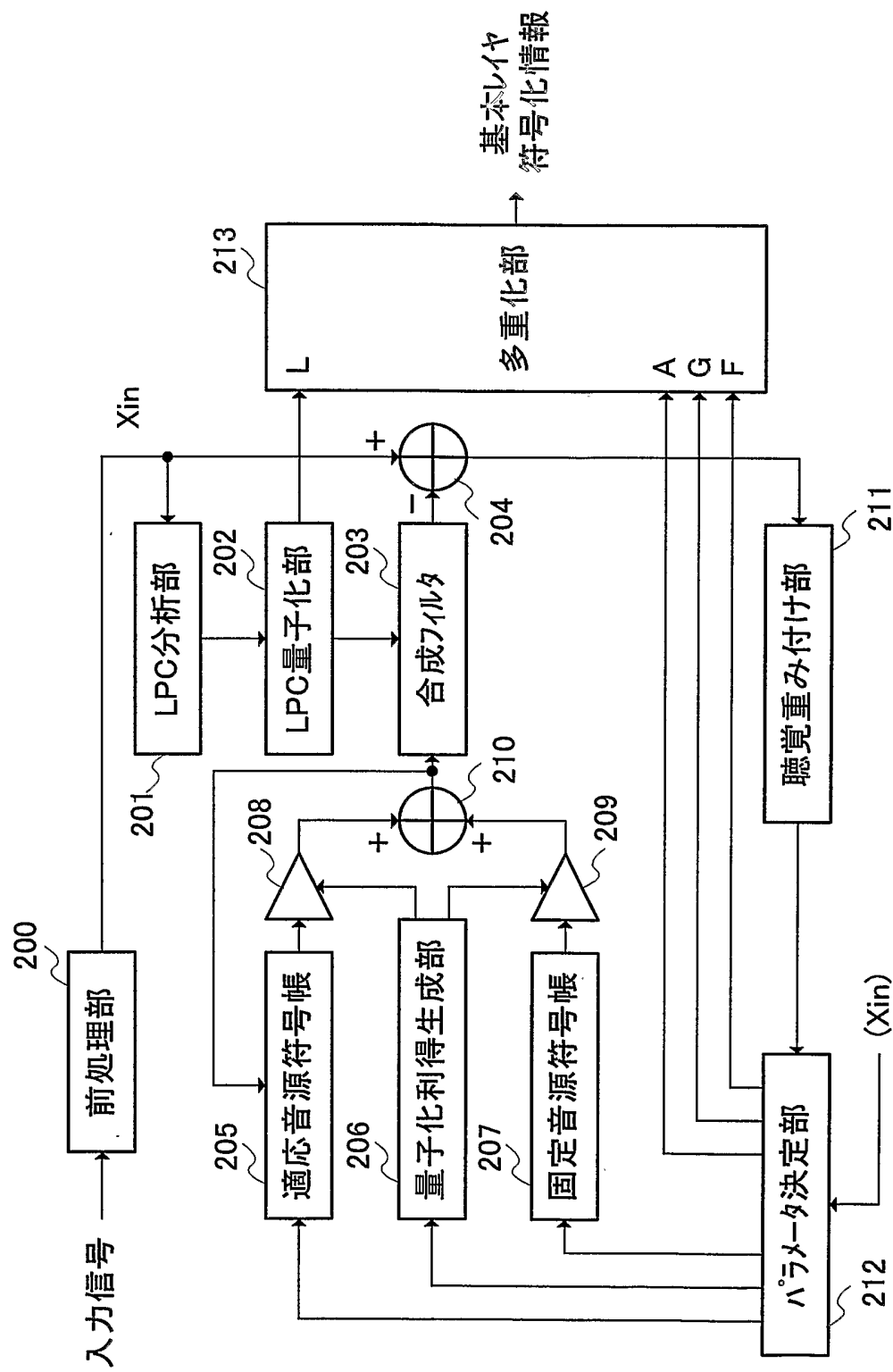


図2

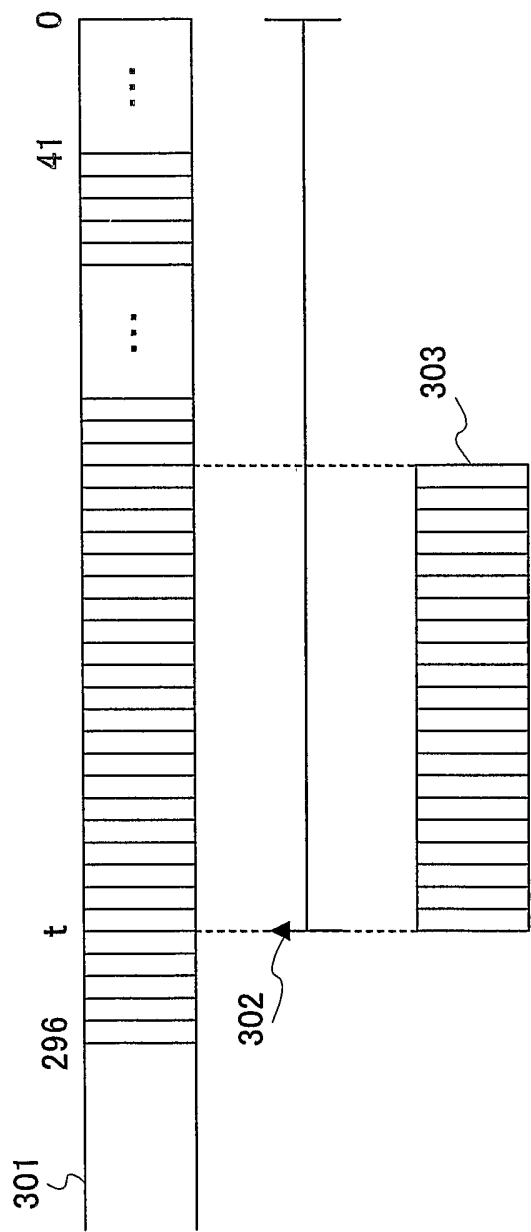


図3

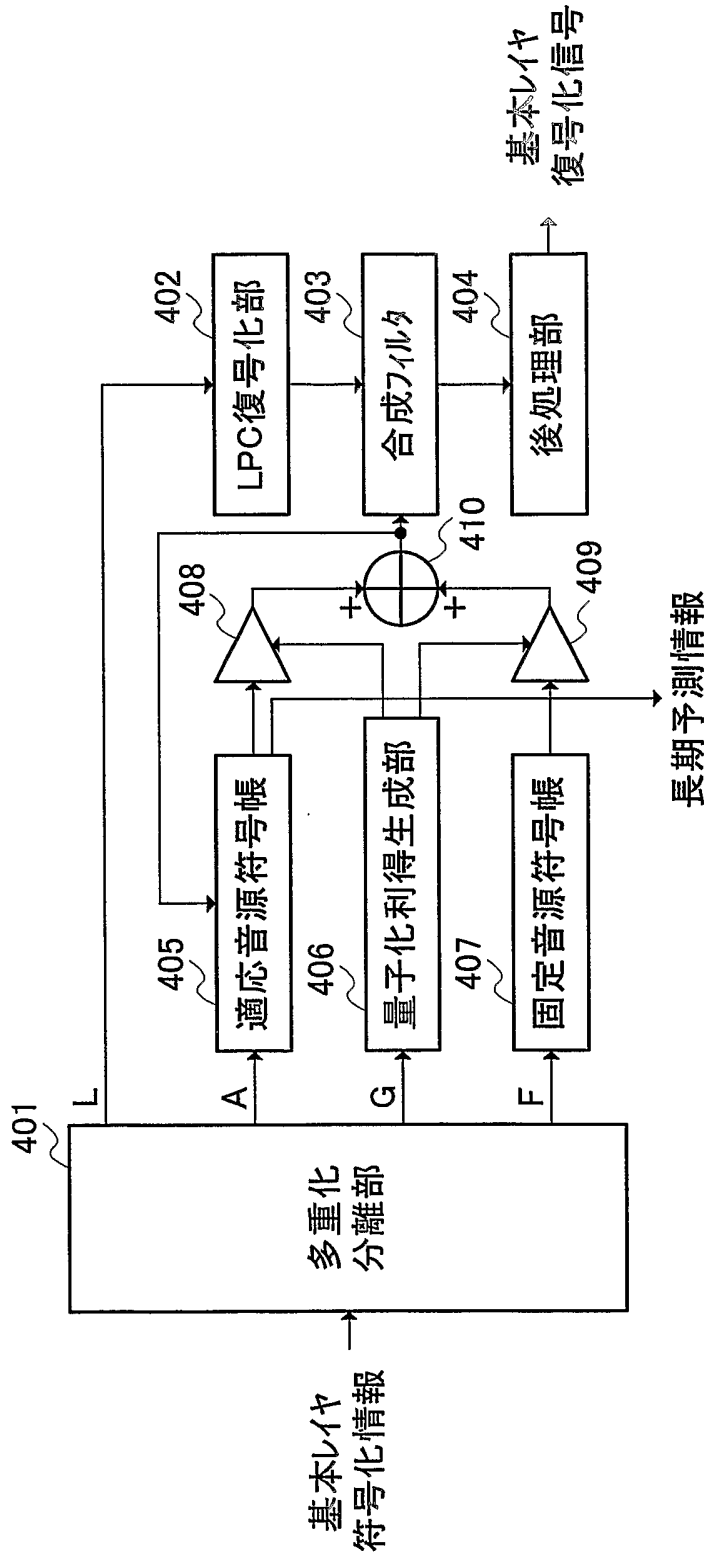


図4

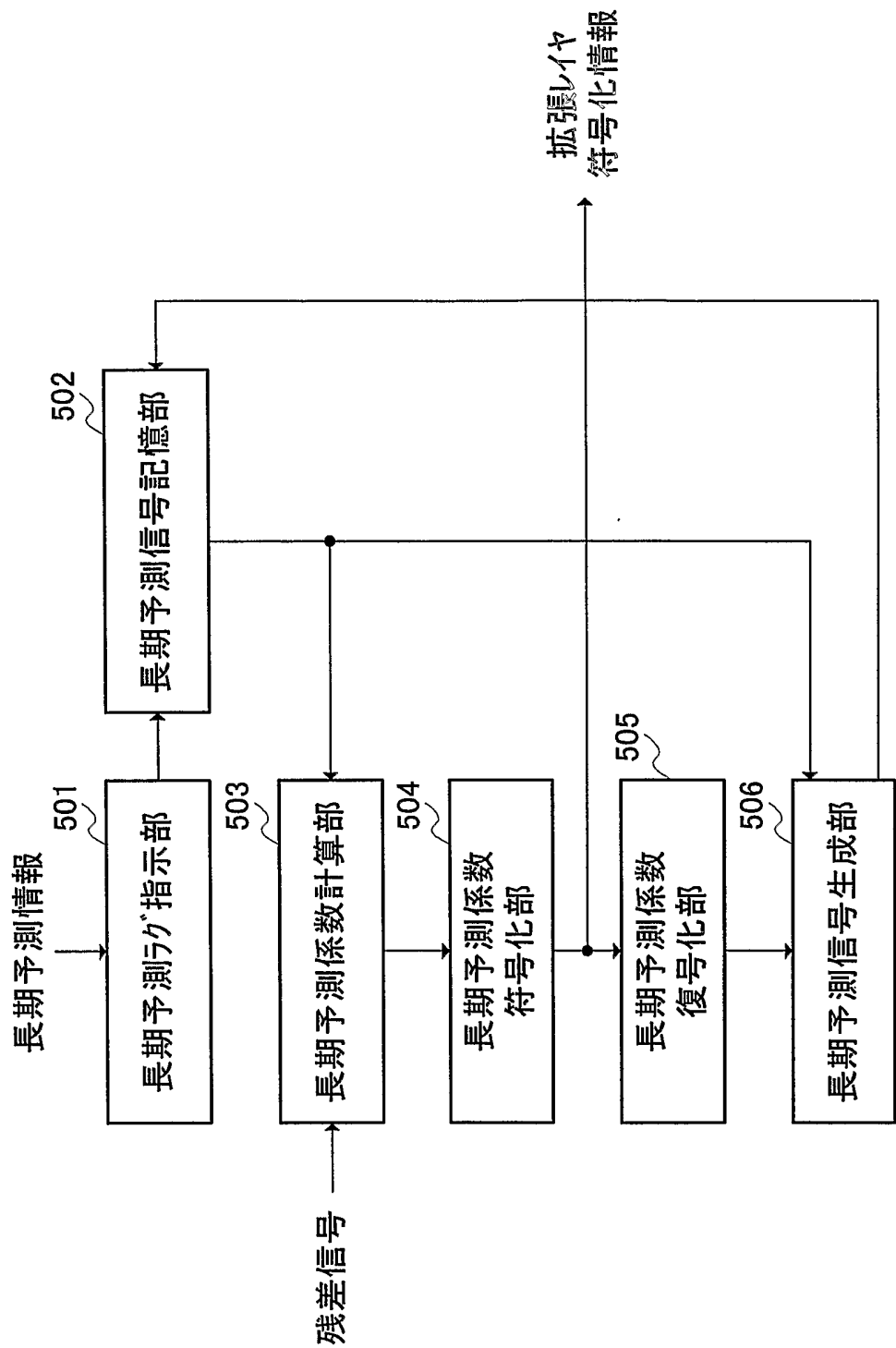


図5

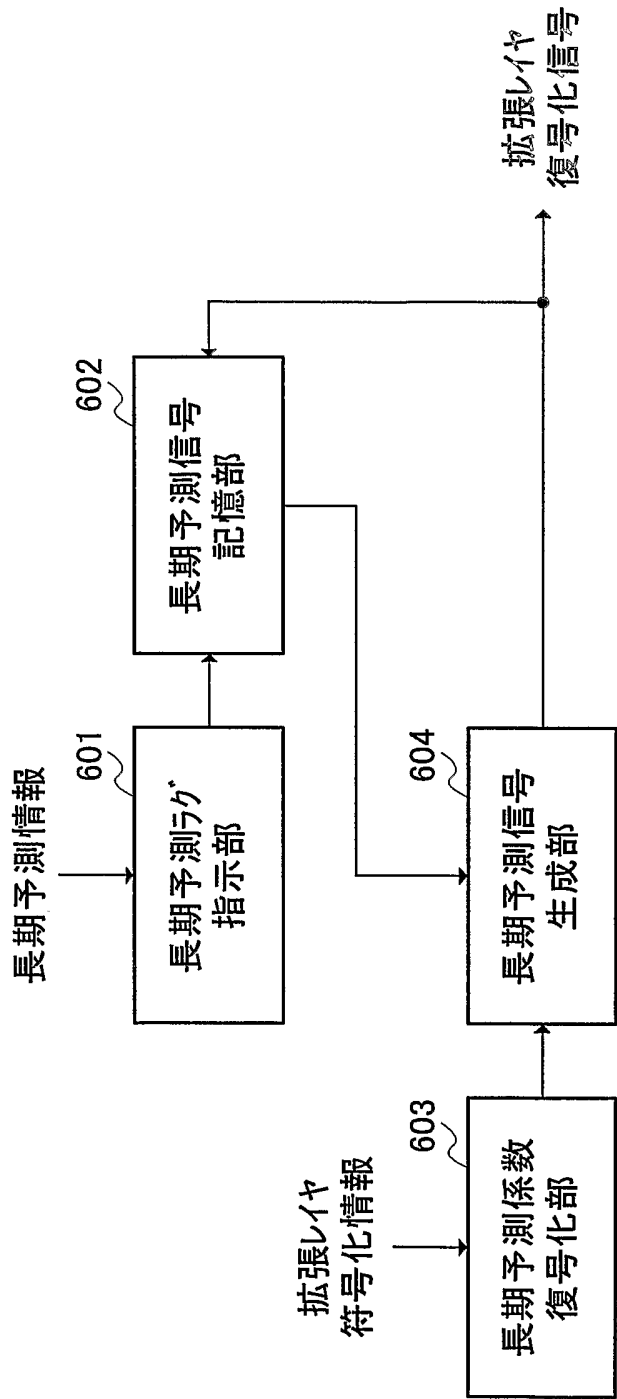


図6

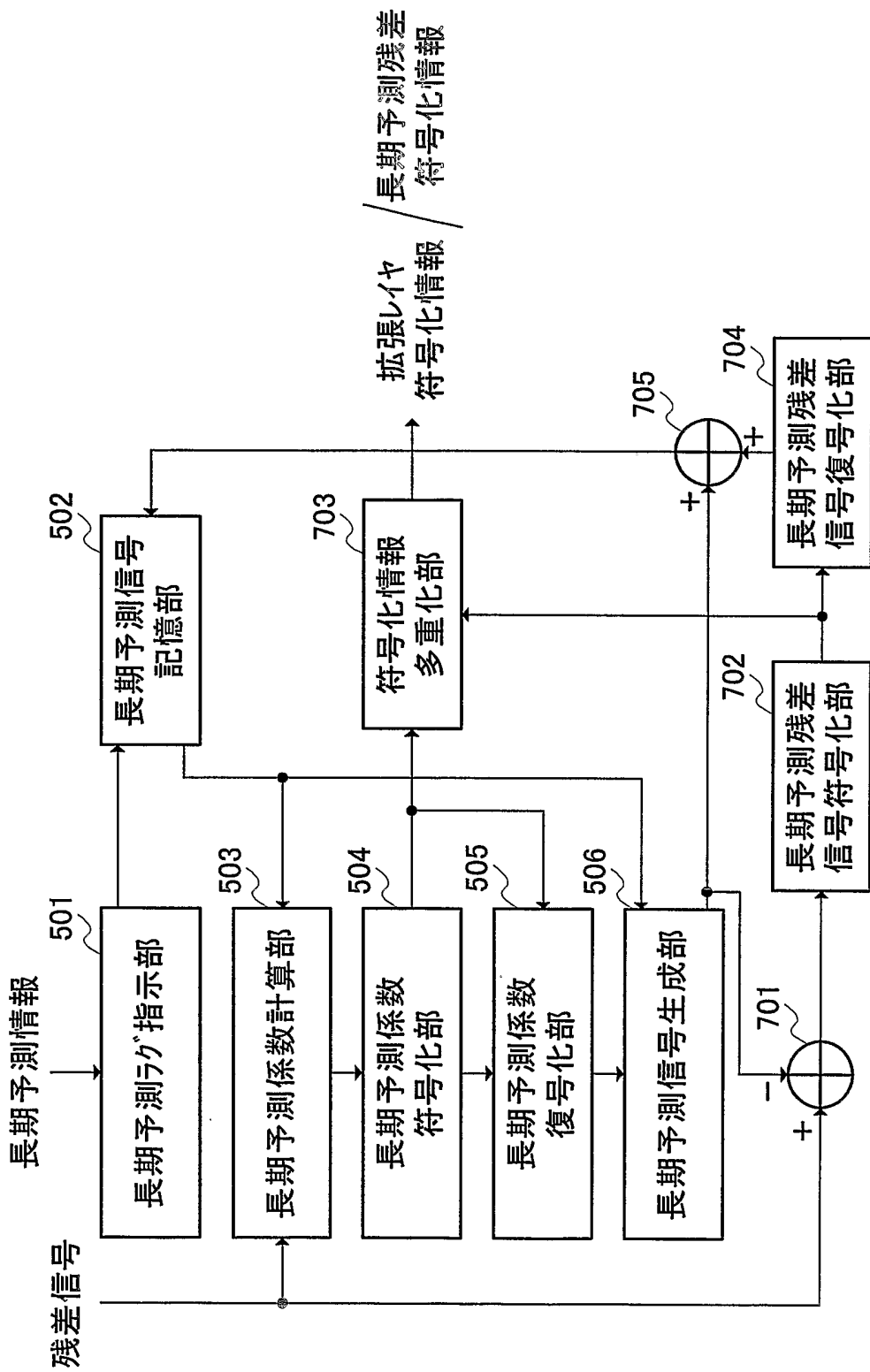


図7

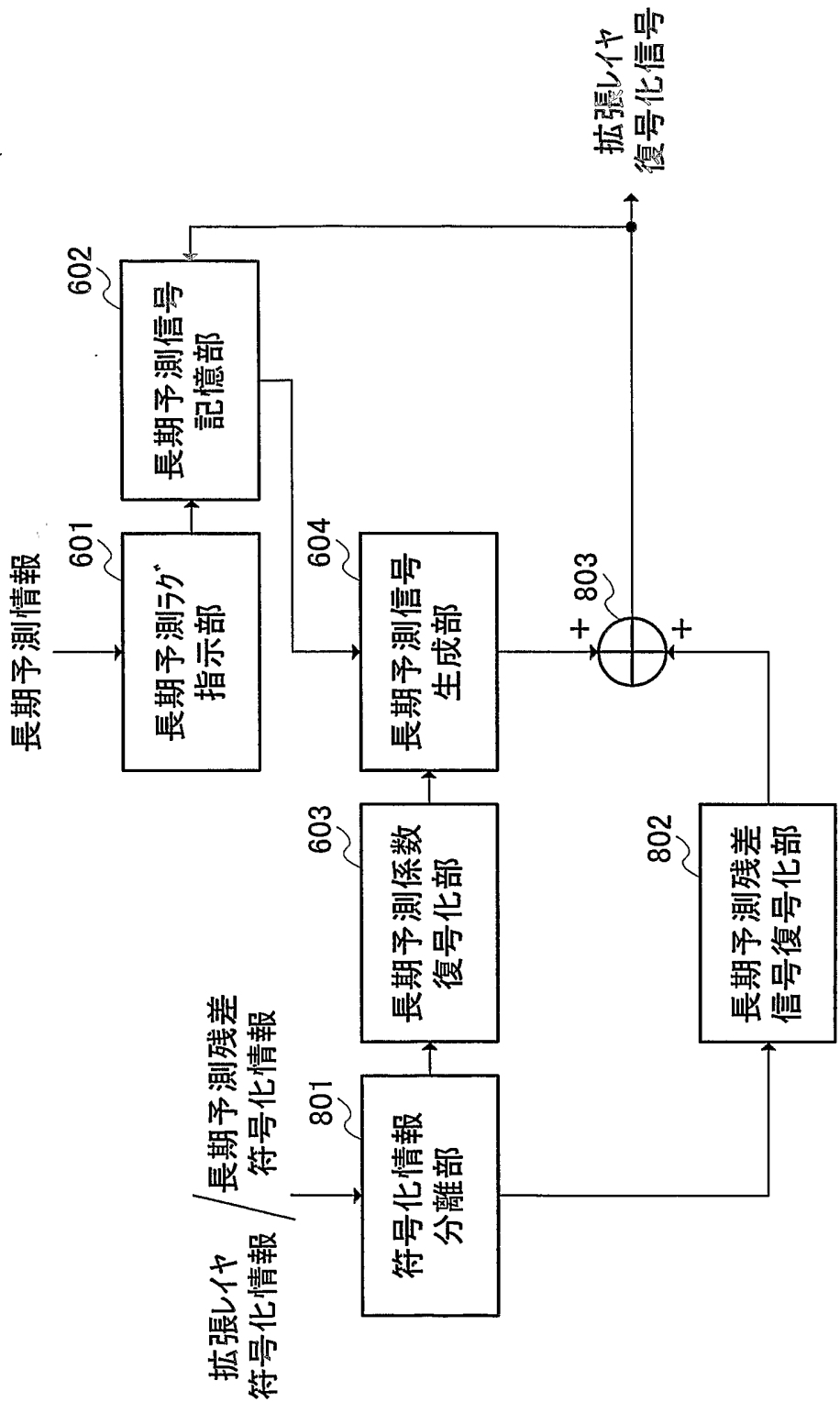


図8

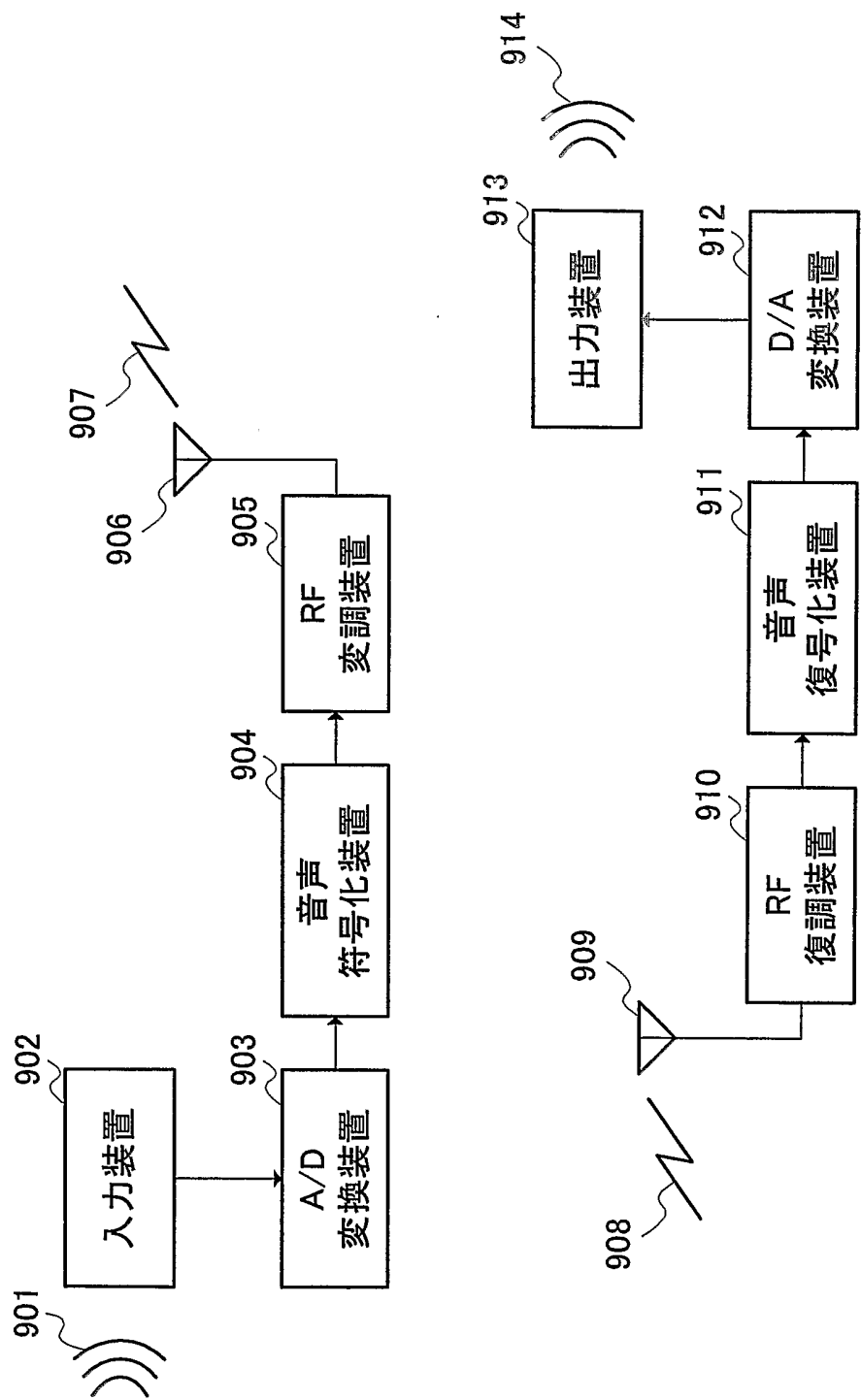


図9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/006294

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G10L19/04, 19/12, H03M7/30

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)

Int.Cl⁷ G10L19/00-19/14, H03M7/30, H04B14/04

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2004
Kokai Jitsuyo Shinan Koho	1971-2004	Toroku Jitsuyo Shinan Koho	1994-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 8-54900 A (Yamaha Corp.), 27 February, 1996 (27.02.96), Full text; Figs. 1 to 6 (Family: none)	1-12
Y	JP 8-328595 A (Sanyo Electric Co., Ltd.), 13 December, 1996 (13.12.96), Full text; Figs. 1 to 6 & JP 8-328596 A & US 5864797 A	1-12
A	JP 10-177399 A (Mitsubishi Electric Corp.), 30 June, 1998 (30.06.98), Full text; Figs. 1 to 15 & JP 2003-29799 A	1-12

☒ Further documents are listed in the continuation of Box C.☐ See patent family 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 application or patent 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

"&" document member of the same patent family

Date of the actual completion of the international search

05 August, 2004 (05.08.04)

Date of mailing of the international search report

24 August, 2004 (24.08.04)

Name and mailing address of the ISA/

Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/006294

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 5-249999 A (Toshiba Corp.), 28 September, 1993 (28.09.93), Full text; Figs. 1 to 9 & US 5671327 A	1-12
A	JP 8-147000 A (Yamaha Corp.), 07 June, 1996 (07.06.96), Full text; Figs. 1 to 7 & US 5797118 A	1-12
A	JP 8-211895 A (Rockwell International Corp.), 20 August, 1996 (20.08.96), Full text; Figs. 1 to 9 & US 5781880 A & EP 713208 A2 & EP 745971 A2	1-12
A	JP 5-73099 A (Oki Electric Industry Co., Ltd.), 26 March, 1993 (26.03.93), Full text; Figs. 1 to 3 (Family: none)	1-12
A	JP 6-102900 A (Fujitsu Ltd.), 15 April, 1994 (15.04.94), Full text; Figs. 1 to 10 (Family: none)	1-12

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ G10L19/04, 19/12, H03M7/30

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ G10L19/00-19/14, H03M7/30, H04B14/04

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年

日本国公開実用新案公報 1971-2004年

日本国実用新案登録公報 1996-2004年

日本国登録実用新案公報 1994-2004年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 8-54900 A (ヤマハ株式会社) 1996.02.27, 全文, 図1-6 (ファミリーなし)	1-12
Y	JP 8-328595 A (三洋電機株式会社) 1996.12.13, 全文, 図1-6 & JP 8-328596 A & US 5864797 A	1-12
A	JP 10-177399 A (三菱電機株式会社) 1998.06.30, 全文, 図1-15 & JP 2003-29799 A	1-12

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

05.08.2004

国際調査報告の発送日

24.8.2004

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

山下 剛史

5C

8946

電話番号 03-3581-1101 内線 3540

C (続き). 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	J P 5-249999 A (株式会社東芝) 1993.09.28, 全文, 図1-9 & US 5671327 A	1-12
A	J P 8-147000 A (ヤマハ株式会社) 1996.06.07, 全文, 図1-7 & US 5797118 A	1-12
A	J P 8-211895 A (ロックウェル・インターナショナル・ コーポレーション), 1996.08.20, 全文, 図1-9 & US 5781880 A & EP 713208 A2 & EP 745971 A2	1-12
A	J P 5-73099 A (沖電気工業株式会社) 1993.03.26, 全文, 図1-3 (ファミリーなし)	1-12
A	J P 6-102900 A (富士通株式会社) 1994.04.15, 全文, 図1-10 (ファミリーなし)	1-12