

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4169320号  
(P4169320)

(45) 発行日 平成20年10月22日(2008.10.22)

(24) 登録日 平成20年8月15日(2008.8.15)

(51) Int.Cl.		F I		
<b>G 1 0 L</b>	<b>21/02</b>	<b>(2006.01)</b>	<b>G 1 0 L</b>	<b>21/02 1 0 4</b>
<b>G 1 0 L</b>	<b>19/00</b>	<b>(2006.01)</b>	<b>G 1 0 L</b>	<b>19/00 2 2 0 Z</b>
<b>H 0 4 B</b>	<b>3/23</b>	<b>(2006.01)</b>	<b>H 0 4 B</b>	<b>3/23</b>
<b>H 0 4 R</b>	<b>3/02</b>	<b>(2006.01)</b>	<b>H 0 4 R</b>	<b>3/02</b>

請求項の数 2 (全 10 頁)

(21) 出願番号	特願2002-103778 (P2002-103778)	(73) 特許権者	000004226
(22) 出願日	平成14年4月5日(2002.4.5)		日本電信電話株式会社
(65) 公開番号	特開2003-295900 (P2003-295900A)		東京都千代田区大手町二丁目3番1号
(43) 公開日	平成15年10月15日(2003.10.15)	(74) 代理人	100127683
審査請求日	平成16年7月26日(2004.7.26)		弁理士 根岸 裕一
審判番号	不服2007-12730 (P2007-12730/J1)	(74) 代理人	100121669
審判請求日	平成19年5月2日(2007.5.2)		弁理士 本山 泰
		(72) 発明者	守谷 健弘
			東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内
		(72) 発明者	田中 雅史
			東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 音声処理方法、音声処理プログラム

(57) 【特許請求の範囲】

【請求項1】

所定数の信号のサンプルにより構成されるフレーム単位で、信号の前処理を行う前処理手段と、

信号の分析を行うことにより複数の符号化モードから一つの符号化モードを選択するモード選択手段と、

前記選択された符号化モードで前記前処理後の信号を符号化する符号化手段と、

を組み合わせる音声信号を処理する音声処理方法において、

各フレーム毎に、

フレームの前記モード選択手段における符号化モードの選択を、前記前処理前の信号を用いて前記前処理手段における前処理と並行して行い、

当該フレームの前記符号化手段における符号化処理を、当該フレームの前記前処理と前記符号化モードの選択とが終了した時点で、当該フレームの前記前処理後の信号に対して前記選択された符号化モードに従って実行する

ことを特徴とする音声処理方法。

【請求項2】

コンピュータが読み取り可能な符号によって記述され、前記請求項1記載の音声処理方法をコンピュータで実行する音声処理プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、例えばエコーキャンセラと符号化とを併用するような音声処理方法、音声処理プログラムに関する。

【0002】

【従来の技術】

図3にこの発明で対象とする音声処理装置の一例を示す。この発明で対象とする音声処理装置は拡声系の通信システムを想定している。端末1Aと端末1BはそれぞれマイクMとスピーカSPを有し、端末1A側のマイクMから入力された音声信号はエコーキャンセラ2Aで反響抑圧処理などの前処理が施され、符号器3Aで圧縮符号化処理されて相手の端末1Bに送られ、相手の端末1Bでは復号器4Bで復号されてスピーカSPから放音される。

10

【0003】

図3に示したような拡声系の通信システムでは図示したように、エコーキャンセラ2Aと、例えば圧縮符号化及び復号化を組み合わせることで実現される。エコーキャンセラ2Aには逐次適応型と時間的にオーバーラップする音声信号をフレーム単位で周波数領域に変換し、周波数領域で反響抑圧処理を行う処理型とに分けることができるが、時間的にオーバーラップするフレーム単位で反響抑圧処理を行う型式の方が効率が高い。

一方、圧縮符号化では複数の符号化モードをフレームごとに適応的に選択することで少ない情報量で高品質な音声を送信することができる圧縮符号化方法が提案されている。

【0004】

20

以下に時間的にオーバーラップするフレーム単位で反響抑圧処理を行う形式のエコーキャンセラと、複数の符号モードをフレームごとに適応的に選択して符号化を行う符号器の一例を説明する。

図4はそのエコーキャンセラの構成を示す。図4において、まず、送信者が発した音声は受話信号 $X(k)$ として受信者側の受信端2から入力され、受信端2から受信端1へと送出される経路において、遅延器303、エコー経路伝搬遅延推定部302、エコー経路601へそれぞれ入力される。なお、ここでエコー経路601は、実際の音響エコー経路もしくは、回線エコー経路(ハイブリッド回路)を意図している。

受話信号 $X(k)$ は、エコー経路601を伝搬することによりエコー信号 $b(k)$ として出力される。一方、送信端1からは受信者側の音声を送話信号 $s(k)$ として入力される。

30

【0005】

エコー経路601を伝搬した後のエコー信号 $b(k)$ 及び受信者側からの送話信号 $s(k)$ は加算器602に入力され、加算器602によりこれら信号は重畳されてエコー重畳信号 $y(k)$ となり、続くエコーキャンセラ301内のエコー経路伝搬遅延推定部302に入力される。

エコー経路伝搬遅延推定部302は、先ほど入力された受話信号 $X(k)$ と今回入力されたエコー重畳信号 $y(k)$ との特徴量の相関から、エコー経路601の伝搬遅延量を算出して出力する。

遅延器303は、エコー経路伝搬遅延推定部302から入力された伝搬遅延量に基づいてFIRフィルタを作成し、作成したFIRフィルタに受話信号 $x(k)$ を通過させる。FIRフィルタ通過後の受話信号 $X(k)$ は、参照信号 $X_r(k)$ となり、続く適応フィルタ部304及びエコー抑圧ゲイン算出部305へ出力される。

40

【0006】

適応フィルタ部304では、入力される参照信号 $X_r(k)$ 及び、エコー重畳信号 $y(k)$ を用いて、エコー重畳信号 $y(k)$ の振幅と位相を推定した推定エコー信号 $y_e(k)$ を生成する。加算器307ではエコー重畳信号 $y(k)$ から推定エコー信号 $y_e(k)$ を差し引くことでエコーを低減した残留エコー信号 $y_{re}(k)$ を生成する。

エコー抑圧ゲイン算出部305では、参照信号 $X_r(k)$ 、エコー重畳信号 $y(k)$ を用いて、エコー重畳信号 $y(k)$ の振幅を推定し、エコー重畳信号 $y(k)$ の振幅を抑圧す

50

るようにエコー抑圧ゲイン  $G$  を算出する。

【 0 0 0 7 】

乗算器 3 0 6 は、エコー重畳信号  $y(k)$  にエコー抑圧ゲイン算出部 3 0 5 から出力されるエコー抑圧ゲイン  $G$  を乗ずることにより、入力される残留エコー信号  $y_{re}(k)$  のエコー信号  $b(k)$  を抑圧し、送信端 2 から入力された送信信号  $s(k)$  にできるだけ近い信号を相手の端末へと出力する。

以上は主にエコーキャンセラ 3 0 1 の動作状況であるが、この発明で対象とするエコーキャンセラは特にエコーキャンセラ 3 0 1 へ入力される参照信号  $X_r(k)$  とエコー重畳信号  $y_{re}(k)$  を周波数分析部 4 0 1 で周波数領域の係数に変換し、加算器 3 0 7 と乗算器 3 0 6 では分析された周波数領域係数毎に反響抑圧処理を施している点を特徴とするものである。なお、エコーキャンセラの範囲にハウリングを防いだり、送受話音声の音量を適切に調整するための利得調整器 3 1 0 を含めることもあるが、以降の説明では、利得調整器を省略する。

【 0 0 0 8 】

周波数分析部 4 0 1 は例えば、図 5 に示すようにフレーム分割部 4 0 1 - 1 と、窓掛け部 4 0 1 - 2 と、時間領域信号を周波数領域係数に変換する変換部 4 0 1 - 3 とによって構成することができる。フレーム分割部 4 0 1 - 1 は入力端子 1 1 から入力される音声データ列を  $N$  サンプル毎にフレーム分割する。窓掛け部 4 0 1 - 2 は例えば 2 フレーム分のデータ列に対して窓関数を乗算する。変換部 4 0 1 - 3 は窓掛けされたデータ列を周波数領域係数に変換し、出力端子 1 2 に周波数領域係数を出力する。

【 0 0 0 9 】

変換部 4 0 1 - 3 は例えば FFT (高速フーリエ変換) 或は DCT (Discrete cosine Transform: 離散コサイン変換)、或は MDCT (Modified Discrete Cosine Transform: 変形離散コサイン変換) などを用いることができる。尚、DCT については K. R. Rao. P. Yip 著、安田、藤原訳 “画像符号化技術 - DCT とその国際標準” 第 2 章 (オーム社)、MDCT については ISO / IEC 標準 IS - 11172 - 3 に述べられている。

窓掛け部 4 0 1 - 2 で施される窓掛け演算は 2 フレーム分の演算処理を 1 フレーム分づつずらしながら実行され、その演算結果を変換部 1 6 で周波数領域係数に変換する。このことから時間的にオーバーラップする窓を使って反響抑圧処理を行うエコーキャンセラと称せられている。

【 0 0 1 0 】

図 6 は周波数合成部 4 0 3 の構成を示す。周波数合成部 4 0 3 は逆変換部 4 0 3 - 1 と、窓掛け部 4 0 3 - 2 と、オーバーラップ加算部 4 0 3 - 3 とによって構成される。入力端子 1 3 に入力された周波数領域係数は逆変換部 4 0 3 - 1 により時間軸上の信号に変換され、その時間軸上の信号に窓掛け部 4 0 3 - 2 は窓掛け演算を施すオーバーラップ加算部 4 0 3 - 3 は 2 フレーム分のデータ列を前の 2 フレームの後半のデータと、後の 2 フレームの前半のデータを重ね合わせて加算し、その重ね合わせて加算したフレームの波形を確定し、出力端子 1 4 から音声信号を出力する。

【 0 0 1 1 】

次に、フレームごとに圧縮符号モードを選択して可及的に少ない情報量により高品位の音声を伝送し、又は記憶することができる符号器について説明する。

例えば MDCT に基づく圧縮符号化では窓の形状を選択する場合、信号の先読みが必要となる。すなわち、フレームの信号の時間的変動の大きさから、長い窓を使うか、短い窓を複数使うかを選択したり、或は信号の性質から窓の形状を選択することが行われている。また、モード選択の他の例としては有声無声音、有声無音、背景雑音の有無などの音声や楽音信号の特徴に適應させたものがある。図 7 はその符号器 2 0 の概略の構成を示す。この例では、有声無声音、有声無音、背景雑音の有無などにより符号化モードを選択する形式の符号器の場合を示す。

【 0 0 1 2 】

10

20

30

40

50

入力端子 2 1 に入力された原音声は、周波数分析部 4 0 1 で周波数領域係数に変換されて音声特徴分析部 2 2 に送られ、ここにおいて音声の特徴を示す種々のパラメータが計算される。このパラメータの代表的な例としては、音声パワ、変形相関関数の最大値、変形相関関数が最大となる遅れ時間（オープンループピッチ）、スペクトル変化量がある。これら計算されたパラメータはモード決定部 2 3 に送られる。モード決定部 2 3 においては、これらのオープンループにより求めた特徴パラメータと、バッファ部 2 5 に記憶される過去のフレームにおける量子化パラメータ例えば前フレームの適応符号帳のピッチ周期、合成音声のパワ、前フレームが属したモード（以下ここではこれらのデータを先読みデータと称す）を使用し、現在のフレームの音声は予め決められたモードの内の何れのモードに属するかを決定し、決定結果に基づいて切替えスイッチ 2 6 及び 2 7 を切替えて符号化部 2 8 - 1 ないし 2 8 - 4 の何れかを選択する。

10

#### 【 0 0 1 3 】

図 8 は 4 モードの場合の音声区間のモード分類の仕方を説明するフローチャートを示す。まず、入力されたパラメータから、フレーム毎に非音声区間（モード 1）であるか或は非音声区間であるかを判定する。これには通常音声パワに着目して判定する。ある閾値を決め、パワがこれより大きい時は音声区間であり、これ以下のときは非音声区間であると判定する。

次に、音声区間であると判定されたフレームについて、有声音であるか、或は無声音であるかを判定する。通常この判定は、音声パワと変形相関関数の最大値を使用して行い、音声パワがある値、例えば 1 サンプルあたりの平均パワが数万～十万以下であり且つ変形相関関数の最大値が或る閾値、例えば 0 . 1 ~ 0 . 3 以下の場合は無声音であると判定する。それ以外の場合は、有声音であると判定する。この場合も判定は音声学的に厳密なものである必要はなく、品質の劣化につながらなければ良い。

20

更に、有声と判定された音声区間について、過渡部であるか、或は定常部であるかを判定する。この判定は、隣接するフレーム間の特徴量の変化分を調べ、変化が少なければ定常部であり、大きければ過渡部（非定常）であると判定する。

#### 【 0 0 1 4 】

図 8 におけるモード 1 及びモード 2 に対応するこれら非音声或は無声音の区間においては、音声のピッチに相当する周期成分がないので、有声音区間と比較して少ないビット数を雑音符号帳及び重み符号長、線形予測パラメータに割り当てることができる。

30

モード 3 に対応する有声音の過渡部については、音声を符号化するのに最も多くのビット数を必要とする。

一方、モード 4 に対応する有声音の定常部については、ピッチ周波数の変化はゆるやかであるから適応符号帳の周期符号は過渡部よりも少ないビット数を割り当てることができる。

#### 【 0 0 1 5 】

##### 【 発明が解決しようとする課題 】

上述したように、前処理手段として周波数領域で反響抑圧処理を施すエコーキャンセラと数フレーム前の先読みデータを利用して符号化モードを選択して符号化処理を行う符号器とを継続に接続して音声を処理する場合、前処理から符号化処理が完了するまでの処理ステップにより遅延が大きく発生してしまう欠点がある。

40

図 9 に周波数領域で反響抑圧処理を施すエコーキャンセラと符号化モードを選択して符号化処理を行う符号器で構成した音声処理装置の一例を示す。図 9 に示すように、送話信号  $x_k(k)$  は周波数分析部 4 0 1 で周波数分析され、周波数領域係数に変換されてエコーキャンセラ 3 0 1 で反響抑圧処理が施され、更に周波数合成部 4 0 3 で時間軸信号に変換されて符号器 2 0 に入力される。

#### 【 0 0 1 6 】

符号器 2 0 でも図 7 で説明したように、周波数分析部 4 0 1 で周波数領域係数に変換され、符号化すべきフレームの信号の状態を予め分析して圧縮符号化モードを決定し、その決定に従って次のフレームのタイミングで符号化処理を実行する。

図 1 0 及び図 9 に示した組合せに関わる処理スケジュールを示す。図 1 0 で右側程新しく

50

、左が過去になる。一番右のフレームF 0を現在のフレームとすると、現在のフレームF 0に実時間で音声データが蓄えられる。このフレームF 0に音声データが蓄積される間にすでにデータが蓄えられている過去のフレームF 1～F 5の信号に対する処理が全て行われる。尚、図10に示す処理スケジュールでは伝送遅延は0としている。

【0017】

まず、ひとつ前のフレームF 1と2つ前のフレームF 2にまたがった信号に対してエコーキャンセラ301が前処理を行う。この前処理が終わった時点でエコーキャンセラ301の後段に配置した周波数合成部403は2つ前のフレームF 2の信号の重ね合わせ加算が終了し、エコーキャンセル処理したフレームF 2の信号の波形を確定する。

符号器20では、このフレームF 2の信号の波形を使って符号化モードの選択のためのモード決定処理を行う。

10

【0018】

このモード決定に基づき、3つ前と4つ前のフレームF 3とF 4にまたがる信号に対して符号化処理を行う。すなわち、3つ前と4つ前のフレームF 3とF 4に対応する圧縮ビット列が作成される。

このビット列が瞬時に復号器に渡り十分に高速に復号できたとすると、4つ前のフレームF 4の信号を確定することができる。つまり、フレームF 0のデータがフレームバッファに対して蓄積が終了した時点でフレームF 4の波形が確定し、再生が可能となる。

【0019】

従って、従来は伝送速度を除くと、フレーム長の5倍の遅れが生じる不都合が生じる。この発明の目的は図9に示した縦続接続の組合せにおいて、遅延量を可及的に短縮することができる音声処理方法及び音声処理プログラムを提案しようとするものである。

20

【0020】

【課題を解決するための手段】

この発明では、所定数の信号のサンプルにより構成されるフレーム単位で、信号の前処理を行う前処理手段と、

信号の分析を行うことにより複数の符号化モードから一つの符号化モードを選択するモード選択手段と、

前記選択された符号化モードで前記前処理後の信号を符号化する符号化手段と、

を組み合わせる音声信号を処理する音声処理方法において、

30

各フレーム毎に、

フレームの前記モード選択手段における符号化モードの選択を、前記前処理前の信号を用いて前記前処理手段における前処理と並行して行い、

当該フレームの前記符号化手段における符号化処理を、当該フレームの前記前処理と前記符号化モードの選択とが終了した時点で、当該フレームの前記前処理後の信号に対して前記選択された符号化モードに従って実行する音声処理方法を提案する。

この発明では、コンピュータが読み取り可能な符号によって記述され、音声処理方法をコンピュータで実行する音声処理プログラムを提案する。

【0021】

作用

40

この発明による音声処理方法によれば、符号器で行われるモード決定処理を前処理手段（エコーキャンセラ）で処理される以前の信号で行う方法を探るから、モードが決定された時点で前処理手段から出力される信号はモード決定処理したフレームの信号である。従って、このフレームの信号に対して直ちに符号化処理を開始することにより1フレームの遅延を削除することができることになる。

従って、この発明によれば全体で4フレームの遅延に短縮できる利点を得られる。

【0022】

【発明の実施の形態】

図1にこの発明の音声処理方法を実行することができる音声処理装置の実施例を示す。図1において図4乃至図7と対応する部分には同一符号を付して示す。図1に示す実施例

50

ではエコーキャンセラ 301 の前段側から音声信号を取り出して符号器 20 の音声特徴分析部 22 に入力し、エコーキャンセラ 301 で行われる反響抑圧処理と並行してモード決定を実行する構成とし、更にエコーキャンセラ 301 で前処理された信号を符号化部 28 - 1 ~ 28 - 4 の何れかで圧縮符号化処理する構成とした場合を示す。

【0023】

つまり、モード決定部 23 で符号化モードを決定するための先読みをエコーキャンセラ 301 の前段側から取り出して行うことにより、エコーキャンセラ 301 で反響抑圧処理された信号より 1 フレーム前のフレームを先読みすることができる。

この結果、モード決定部 23 で先読みしたフレームの信号に対するモード決定処理が終了する時点で、そのフレームの信号の波形が周波数合成部 403 から出力され始める。従って、符号化モードが決定されたフレームの信号をスイッチ 26 を通じてモード決定部 23 で決定した符号化部 28 - 1 ~ 28 - 4 の何れかに入力し、符号化処理が行われる。符号化処理された信号は符号送出部 24 から伝送路に送出される。これと共に、バッファ部 25 に先読みデータとして蓄積される。

10

【0024】

図 2 及び図 1 に示した実施例の処理スケジュールを示す。ここでは上述したようにエコーキャンセラ 301 の処理が終了した波形でなく、エコーキャンセラ 301 の処理前の波形（フレーム F1 の信号）を近似的に使う符号化前のモード決定処理を行う。これにより 2 つ前のフレーム F2 に対してエコーキャンセラ 301 の処理が終了すれば、符号化部 28 - 1 ~ 28 - 4 の何れかはフレーム F2 と F3 にまたがる信号の符号化処理を開始し、フレーム F2 と F3 のフレームに対応する圧縮ビット列を作成する。

20

【0025】

このビット列が瞬時に復号器に渡り、充分高速に復号できたとすると、フレーム F3 の波形を確定することができるので、このフレーム F3 の信号を再生することができる。従って、この発明によればフレーム長の 4 倍の時間の遅れで信号波形を再生することができる。

上述したこの発明による音声処理方法はコンピュータが読み取り可能な符号によって記述されたプログラムをコンピュータで実行することにより実現することができる。プログラムは例えば CD-ROM のような記憶媒体からコンピュータにインストールされて実行するか、又は通信回線を通じてコンピュータにインストールして実行させることができる。

30

【0026】

【発明の効果】

以上説明したように、この発明によれば拡声系の通信システムの遅延を削減することができ、双方向通話の質を改善できる効果が得られる。

【図面の簡単な説明】

【図 1】この発明による音声処理方法を実行することができる音声処理装置の一実施例を説明するためのブロック図。

【図 2】図 1 の動作を説明するためのタイミングチャート。

【図 3】従来の技術を説明するためのブロック図。

【図 4】従来から用いられている周波数領域で反響抑圧処理を実行するエコーキャンセラの構成及び動作を説明するためのブロック図。

40

【図 5】図 4 に示した周波数分析部の内部の構成を説明するためのブロック図。

【図 6】図 4 に示した周波数合成部の内部の構成を説明するためのブロック図。

【図 7】モード選択手段を具備した符号器の構成及びその動作を説明するためのブロック図。

【図 8】図 7 に示した符号器の動作を説明するためのフローチャート。

【図 9】図 4 に示したエコーキャンセラと図 7 に示した符号器を組み合わせた構成を説明するためのブロック図。

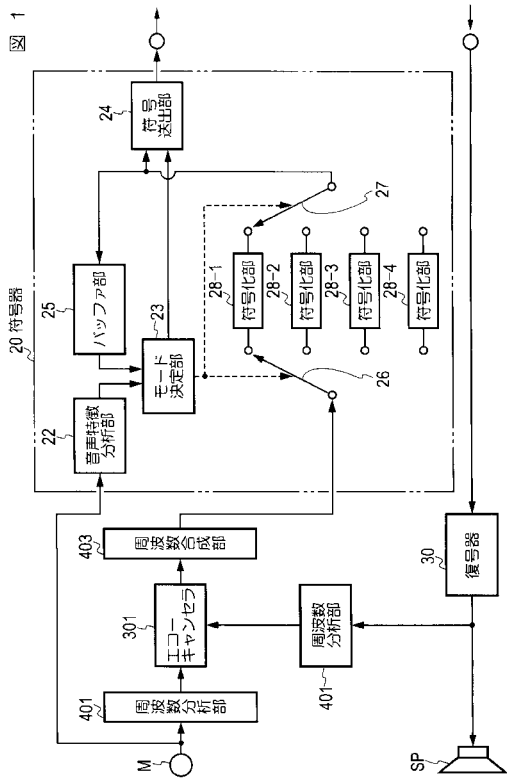
【図 10】図 9 に示した構成の動作を説明するためのタイミングチャート。

【符号の説明】

50

M	マイクロホン	26、27	スイッチ
SP	スピーカ	28-1 ~ 28-4	符号化部
20	符号器	30	復号器
22	音声特徴分析部	301	エコーキャンセラ
23	モード決定部	401	周波数分析部
24	符号送出部	403	周波数合成部
25	バッファ部		

【図1】



【図2】

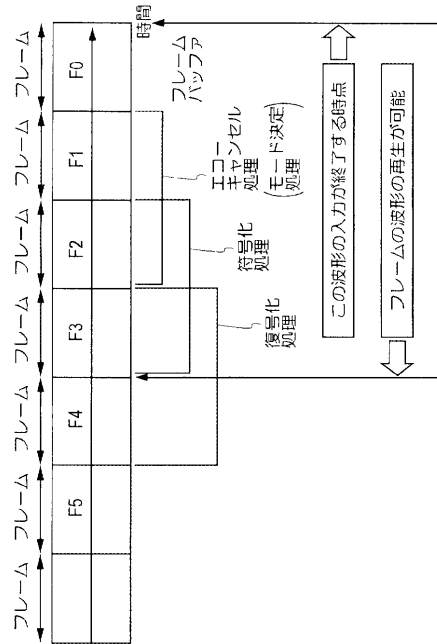
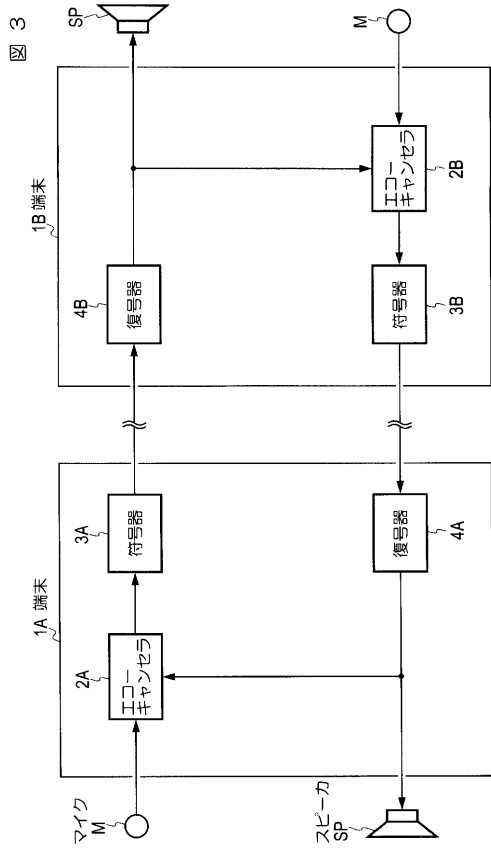
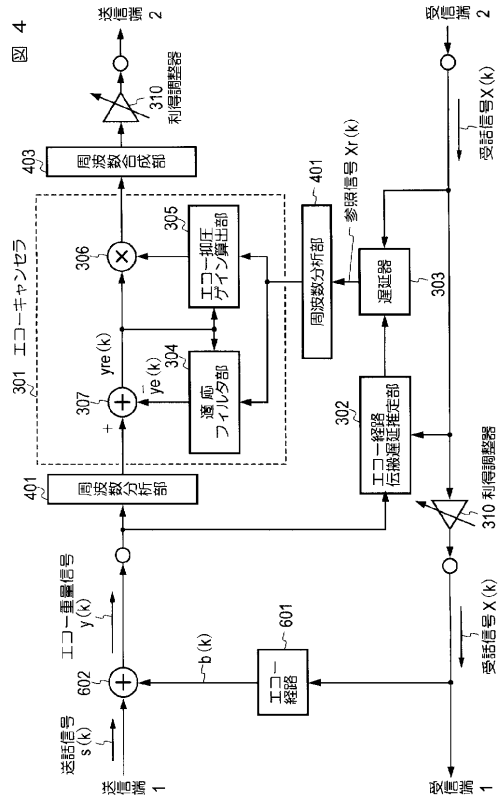


図 2

【図3】



【図4】



【図5】

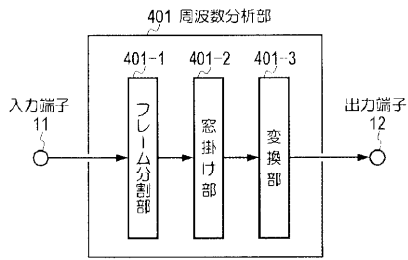
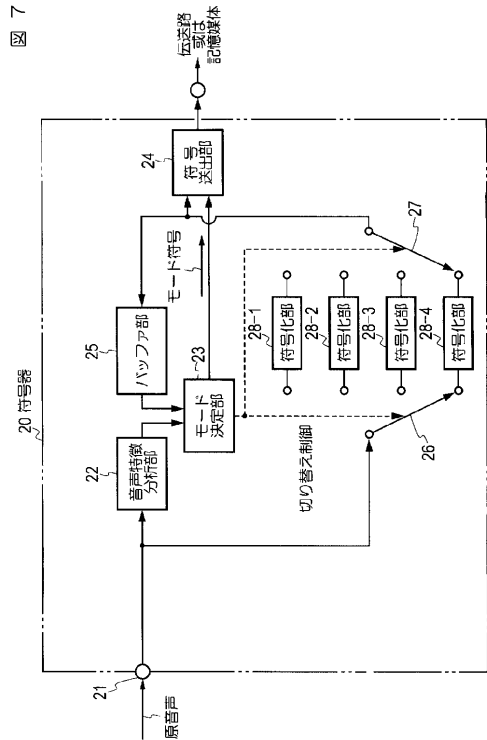


図 5

【図7】



【図6】

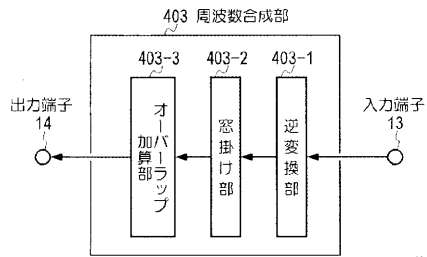


図 6



【 図 8 】

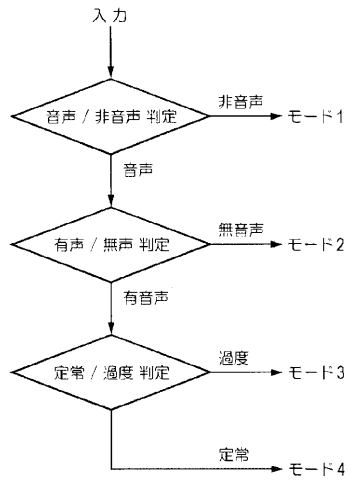


図 8

【 図 9 】

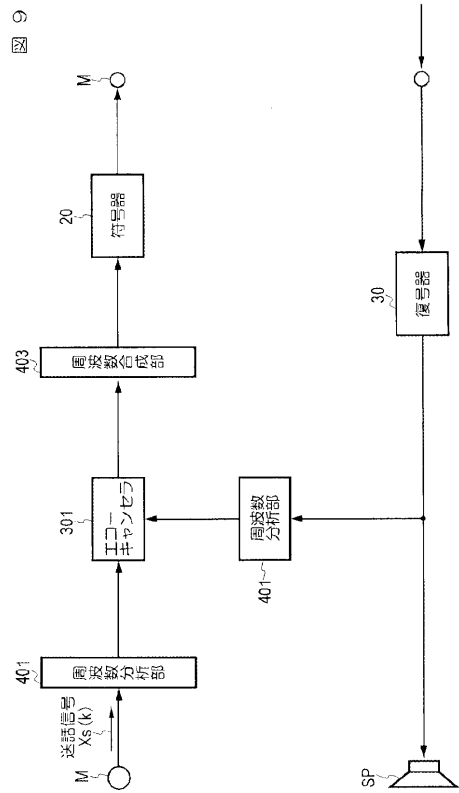


図 9

【 図 10 】

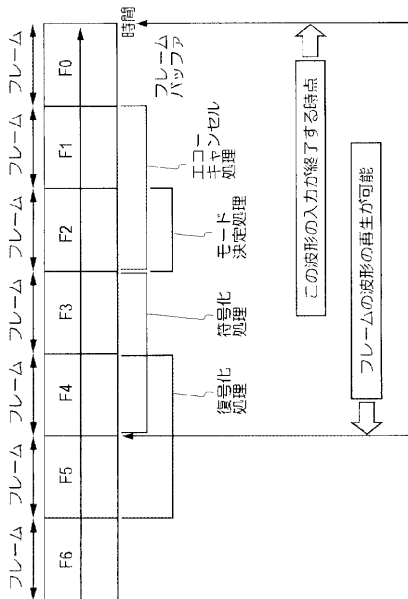


図 10

---

フロントページの続き

合議体

審判長 西山 昇

審判官 伊藤 隆夫

審判官 原 光明

(56)参考文献 特開2001-318694(JP,A)  
特開平8-263098(JP,A)

(58)調査した分野(Int.Cl., DB名)

G10L19/00-21/02

H04B14/04

H03M7/30