

(19) 日本国特許庁 (JP)

(12) 特許公報 (B2)

(11) 特許番号

**特許第3946074号**

(P3946074)

(45) 発行日 平成19年7月18日 (2007.7.18)

(24) 登録日 平成19年4月20日 (2007.4.20)

(51) Int. Cl.<sup>8</sup>

F I

G 1 0 L 21/02 (2006.01)

G 1 0 L 21/02 1 0 1 A

G 1 0 L 19/02 (2006.01)

G 1 0 L 21/02 1 0 3 Z

G 1 0 L 19/02 1 2 0

請求項の数 6

(全 2 0 頁)

(21) 出願番号 特願2002-104363 (P2002-104363)  
 (22) 出願日 平成14年4月5日 (2002.4.5)  
 (65) 公開番号 特開2003-295898 (P2003-295898A)  
 (43) 公開日 平成15年10月15日 (2003.10.15)  
 審査請求日 平成16年8月3日 (2004.8.3)

(73) 特許権者 000004226  
 日本電信電話株式会社  
 東京都千代田区大手町二丁目3番1号  
 (74) 代理人 100121706  
 弁理士 中尾 直樹  
 (74) 代理人 100066153  
 弁理士 草野 卓  
 (74) 代理人 100128705  
 弁理士 中村 幸雄  
 (74) 代理人 100100642  
 弁理士 稲垣 稔  
 (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

最終頁に続く

(54) 【発明の名称】 音声処理装置

(57) 【特許請求の範囲】

【請求項 1】

音声信号をNサンプルからなるフレーム時刻ごと (Nは2以上の整数) に2Nサンプルの切り出し窓を乗算してサンプルの取り込みが完了している1フレーム前と2フレーム前の2Nサンプルの窓掛け音声信号を得る第1オーバーラップ窓掛け演算処理部と、

この第1オーバーラップ窓掛け演算処理部が演算処理した2Nサンプルの窓掛け音声信号を周波数領域係数に変換する第1変換部と、

この第1変換部で変換した周波数領域係数を演算処理して反響抑圧処理を施す前処理手段と、

この前処理手段の演算処理結果を時間領域信号に変換する第1逆変換部と、

この第1逆変換部で逆変換した時間領域信号を直接、周波数領域係数に変換する第2変換部と、

この第2変換部で変換した周波数領域係数を量子化し、符号化・復号化する手段と、

この手段の処理結果を時間領域信号に変換する第2逆変換部と、

この第2逆変換部が逆変換した時間領域信号を1フレーム分ずつオーバーラップさせて加算し、オーバーラップしたフレームの音声信号を確定するオーバーラップ加算部と、

によって構成したことを特徴とする音声処理装置。

【請求項 2】

音声信号をNサンプルからなるフレーム時刻ごと (Nは2以上の整数) に2Nサンプルの切り出し窓を乗算してサンプルの取り込みが完了している1フレーム前と2フレーム前の

2 Nサンプルとの窓掛け音声信号を得るオーバーラップ窓掛け演算処理部と、  
 このオーバーラップ窓掛け演算処理部が演算処理した2 Nサンプルの窓掛け音声信号を  
 周波数領域係数に変換する変換部と、  
 この変換部で変換した周波数領域係数を演算処理して反響抑圧処理を施す前処理手段と、  
 この前処理手段の処理結果に直接、符号化・復号化処理を施す手段と、  
 この手段で処理した処理結果を時間領域信号に変換する逆変換部と、  
 この逆変換部が変換した時間領域信号を1フレーム分ずつオーバーラップさせて加算し、  
 オーバーラップされたフレームの音声信号を確定するオーバーラップ加算部と、  
 によって構成したことを特徴とする音声処理装置。

10

【請求項3】

音声信号をNサンプルからなるフレーム時刻ごと(Nは2以上の整数)に2 Nサンプルの  
 切り出し窓を乗算してサンプルの取り込みが完了している1フレーム前と2フレーム前の  
 2 Nサンプルの窓掛け音声信号を得る第1オーバーラップ窓掛け演算処理部と、  
 この第1オーバーラップ窓掛け演算処理部が演算処理した2 Nサンプルの窓掛け音声信  
 号を周波数領域係数に変換する第1変換部と、  
 この第1変換部が変換した周波数領域係数を圧縮符号化及び復号化処理する符号器及び  
 復号器と、  
 復号器が復号した復号信号を時間領域信号に逆変換する第1逆変換部と、  
 この第1逆変換部で逆変換した時間領域信号を直接、周波数領域係数に変換する第2変換部と、  
 この第2変換部で変換した周波数領域係数を後処理する後処理手段と、  
 この後処理手段で処理した処理結果を時間領域信号に変換する第2逆変換部と、  
 この第2逆変換部が変換した時間領域信号を1フレーム分ずつオーバーラップさせて加  
 算し、オーバーラップされたフレームの音声信号を確定するオーバーラップ加算部と、  
 によって構成したことを特徴とする音声処理装置。

20

【請求項4】

音声信号をNサンプルからなるフレーム時刻ごと(Nは2以上の整数)に2 Nサンプルの  
 切り出し窓を乗算してサンプルの取り込みが完了している1フレーム前と2フレーム前の  
 2 Nサンプルとの窓掛け音声信号を得るオーバーラップ窓掛け演算処理部と、  
 このオーバーラップ窓掛け演算処理部が演算処理した2 Nサンプルの窓掛け音声信号を  
 周波数領域係数に変換する変換部と、  
 この変換部で変換した周波数領域係数を演算処理して符号化・復号化処理を施す手段と、  
 この手段の処理結果に直接、雑音抑圧処理を施す後処理手段と、  
 この後処理手段で処理した処理結果を時間領域信号に変換する逆変換部と、  
 この逆変換部が変換した時間領域信号を1フレーム分ずつオーバーラップさせて加算し、  
 オーバーラップされたフレームの音声信号を確定するオーバーラップ加算部と、  
 によって構成したことを特徴とする音声処理装置。

30

【請求項5】

音声信号をNサンプルからなるフレーム時刻ごと(Nは2以上の整数)に2 Nサンプルの  
 切り出し窓を乗算してサンプルの取り込みが完了している1フレーム前と2フレーム前の  
 2 Nサンプルの窓掛け音声信号を得る第1オーバーラップ窓掛け演算処理部と、  
 この第1オーバーラップ窓掛け演算処理部が演算処理した2 Nサンプルの窓掛け音声信  
 号を周波数領域係数に変換する第1変換部と、  
 この第1変換部が変換した周波数領域係数を圧縮符号化及び復号化処理する符号器及び  
 復号器と、  
 復号器が復号した復号信号を時間領域信号に逆変換する逆変換部と、  
 この逆変換部で逆変換した時間領域信号を直接、周波数領域係数に変換する第2変換部と、

40

50

この第 2 変換部で変換した周波数領域係数を参照信号として利用するエコーキャンセラと、

によって構成したことを特徴とする音声処理装置。

【請求項 6】

音声信号を  $N$  サンプルからなるフレーム時刻ごと ( $N$  は 2 以上の整数) に  $2N$  サンプルの切り出し窓を乗算してサンプルの取り込みが完了している 1 フレーム前と 2 フレーム前の  $2N$  サンプルの窓掛け音声信号を得る第 1 オーバーラップ窓掛け演算部と、

この第 1 オーバーラップ窓掛け演算部が演算処理した  $2N$  サンプルの窓掛け音声信号を周波数領域係数に変換する第 1 変換部と、

この第 1 変換部が変換した周波数領域係数を圧縮符号化及び復号化処理する符号器及び復号器と、

この復号器の復号結果を直接、参照信号として利用するエコーキャンセラと、

によって構成したことを特徴とする音声処理装置。

【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などの前処理と符号化及び復号化を後処理とするような音声処理装置、又は符号化、復号化を前処理とし、ノイズキャンセラを後処理とする音声処理装置に関する。

【0002】

【従来の技術】

図 1 4 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 1 A と端末 1 B はそれぞれマイク  $M$  とスピーカ  $SP$  を有し、端末 1 A 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 2 A で反響抑圧処理などの前処理が施され、符号器 3 A で圧縮符号化処理されて相手の端末 1 B では復号器 4 B で復号されてスピーカ  $SP$  から放音される。

【0003】

図 1 4 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 2 A と例えば圧縮符号化及び符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 2 A 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形式の方が効率が高い。

また、音声や音楽の符号化でも時間的にオーバーラップする窓を使って変換符号化を行う場合がある。このように時間的にオーバーラップする窓を使った符号化とエコーキャンセラのような前処理を組み合わせると処理の遅延が大きくなってしまふという欠点があった。

【0004】

その理由を明らかにするために、予め時間的にオーバーラップする窓を使っ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、時間的にオーバーラップする窓を使って符号化及び符号化処理を行う符号器及び符号器について説明する。

図 1 5 はそのエコーキャンセラの構成を示す。同図において、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 へと送られる経路において、遅延器 3 0 3、エコー経路伝搬遅延推定部 3 0 2、エコー経路 6 0 1 へそれぞれ入力される。なお、ここでエコー経路 6 0 1 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。

【0005】

受話信号  $x(k)$  は、エコー経路 6 0 1 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。

エコー経路 6 0 1 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  $s(k)$

10

20

30

40

50

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。

【0006】

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 フィルタに受信信号  $x(k)$  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。

10

【0007】

適応フィルタ部 304 では、入力される参照信号  $X_r(k)$  及び、残留エコー信号  $y(k)$  を用いて、残留エコー信号  $y(k)$  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重畳信号  $y_{re}(k)$  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  $y(k)$  の振幅を抑圧するようにエコー抑圧ゲイン  $G$  を算出する。

【0008】

20

乗算器 306 は、エコー重畳信号  $y(k)$  にエコー抑圧ゲイン算出部 305 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重畳信号  $y_{re}(k)$  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 301 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 301 へ入力される参照信号  $X_r(k)$  とエコー重畳信号  $y(k)$  を周波数分析部 401 で周波数領域の係数に変換し、加算器 307 と乗算器 306 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節器 310 を含めることもあるが、以降の説明では、利得調節器を省略する。

30

【0009】

周波数分析部 401 は例えば、図 16 に示すようにフレーム分割部 14 と、窓掛け部 15 と、時間領域信号を周波数領域係数に変換する変換部 16 とによって構成することができる。フレーム分割部 14 は入力端子 11 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 15 は例えば  $N$  フレーム分のデータ列一般には 2 フレームのデータ列に対して窓関数を乗算する。変換部 16 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 12 に周波数領域係数を出力する。

【0010】

変換部 16 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。

40

窓掛け部 15 で施される窓掛け演算は 2 フレーム分の演算処理を 1 フレーム分ずつシフトさせながら実行され、その演算結果を変換部 16 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。

【0011】

50

図 17 は周波数合成部 403 の構成を示す。周波数合成部 403 は逆変換部 58 と、窓掛け部 59 と、オーバーラップ加算部 61 とによって構成される。入力端子 13 に入力された周波数領域係数は逆変換部 58 により時間領域の信号に変換され、その時間領域の信号に窓掛け部 59 は窓掛け演算を施す。オーバーラップ加算部 61 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、その重ね合わされたフレームの波形を確定し、出力端子 14 から音声信号を出力する。

#### 【0012】

次に時間的にオーバーラップする窓を使って符号化及び復号化を行う符号器及び復号器の例を説明する。図 18 は特許 3087814 号明細書に開示された符号器と復号器の構成を示す。この発明では符号器と復号器の詳細な構成を要件とするものでないから、各部の構成は各ブロック内に名称を記入して示すに留め、ここでは、符号器 10 側の入力部分にフレーム分割部 14 と、窓掛け部 15 と、変換部 16 とによって構成される周波数分析部 401 が格納されている点と、復号器 50 側の出力部分に逆変換部 58 と、窓掛け部 59、オーバーラップ加算部 61 とによって構成される周波数合成部 403 が格納されている点を理解されたい。

#### 【0013】

図 18 に示した構成を簡素化してこの発明の説明に適した必要最小限の構成を図 19 に示す。符号器 10 は周波数分析部 401 と量子化部 411 とを縦続接続して表示することができる。また、復号器 50 は逆量子化部 412 と周波数合成部 403 を縦続接続して表示することができる。

符号器 10 では入力端子 11 から音声信号が入力され、この音声信号が周波数分析部 401 で周波数領域係数に変換される。周波数領域係数は量子化部 411 で量子化処理されて伝送路に送出される。

#### 【0014】

復号器 50 では逆量子化部 412 で逆量子化処理し、その逆量子化された周波数領域係数が周波数合成部 403 で時間領域の信号に変換され、出力端子 91 から音声信号が再生される。

符号器 10 でも 2 フレーム分ずつ窓掛け演算処理が施され、変換部 16 で 2 フレーム分ずつ周波数領域係数に変換され、その 2 フレーム分の周波数領域係数が 1 フレーム分ずつシフトされながら量子化部 411 で量子化される。

量子化部 411 で量子化された周波数領域係数は伝送路を通じて相手側の復号器 50 に伝送され、逆量子化部 412 で逆量子化され、更に逆変換部 58 で時間領域信号に逆変換され、窓掛け部 59 で窓掛け演算が施され、オーバーラップ加算部 61 で 2 フレーム分ずつ出力される時間領域信号の前の 2 フレームの後半の 1 フレームと後の 2 フレームの前半の 1 フレームとを重ね合わせて加算し、その重ね合わせたフレームの音声波形を再生する。

#### 【0015】

##### 【発明が解決しようとする課題】

以上の説明から分かるように、上述したエコーキャンセラ 301 及び符号器 10、復号器 50 は音声信号を一旦周波数領域係数に変換し、周波数領域で反響抑圧処理及び符号化復号化処理を施すから、窓掛け部 15 及びオーバーラップ加算部 61 で遅延が発生する欠点がある。

図 20 に周波数領域で反響抑圧処理を行うエコーキャンセラ 301 と、周波数領域で符号化処理する符号器 10 と、符号処理を行う復号器 50 を縦続接続した音声処理装置の一例を示す。尚、図 20 に示す周波数分析部 401 は同一符号を付して示しているが、一般にエコーキャンセラ 301 の前段に用いられるものと、符号器 10 の前段に用いられるものはそれぞれ異なる仕様で動作する周波数分析部が用いられており、共通の仕様の周波数分析部が用いられる例はない。この音声処理装置動作状況を図 21 に示す。図 21 の右側程新しく左が過去になる。一番右のフレーム F0 を現在のフレームとすると、この現在の

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の時間の中ですでにサンプルが蓄えられている過去のフレームF 1、F 2、F 3...の信号に対する処理を逐次すべて行うことを前提とする。尚、伝送遅延はこの図では0としている。

【0016】

まず、1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Bに対して、エコーキャンセラの処理を行う。この処理が終わった時点で前回エコーキャンセル処理が終了している1フレームの信号Aの後半の1フレームF 2と現在エコーキャンセル処理が終わった2フレームの信号Bの後半の1フレーム（これもフレームF 2になる）との重ね合わせ加算が終了してフレームF 2の反響抑圧処理された波形が確定する。この波形を使って窓掛けから始まる符号化を行う。すなわち符号化は2つ前のフレームF 2と3つ前のフレームF 3にまたがる信号Aに対応する圧縮ビット列を作成する。

10

【0017】

このビット列が瞬時に相手の復号器に渡り、十分高速に復号できたとすると、復号器ではフレームF 2とF 3をまたぐ2フレームの信号とフレームF 3とF 4をまたぐ2フレームの信号から3つ前のフレームF 3の信号を確定することができるのでその信号を再生することができる。すなわち、伝送速度を除くとフレーム長の4倍の時間の遅れが生じているということになる。

図22に図21を用いて説明した動作の処理の流れを示す。ステップSP1～SP5はエコーキャンセラで実行される前処理ステップを示す。ステップSP1では2フレーム分の音声信号に対してオーバーラップ窓掛けを実行する。

20

【0018】

ステップSP2では周波数領域係数への変換処理を実行する。

ステップSP3では反響抑圧のための処理と、データの変形処理を実行する。

ステップSP4では周波数領域係数を時間軸データに逆変換し、ステップSP5で窓掛け演算とオーバーラップ加算を実行し、音声信号を再生する。

ステップSP6～SP10は符号器及び復号器で実行される後処理ステップを示す。ステップSP6でオーバーラップ窓掛け処理を実行し、ステップSP7で周波数領域係数に変換し、ステップSP8で符号化のための量子化処理と、伝送、復号化のための逆量子化処理が実行される。ステップSP9で時間軸データに逆変換され、ステップSP10で窓掛けオーバーラップ加算処理により音声信号を再生する。

30

【0019】

図20に示した音声処理装置では符号化及び復号化を行う前に反響抑圧処理を実行する形式の場合を説明したが、符号化及び復号化処理後に例えばノイズ抑圧のような後処理を行う場合もある。

図23はその形式の音声処理装置を示す。この形式の音声処理装置の場合も符号器10と後処理となるノイズキャンセラのような後処理手段404のそれぞれで、図24に示すステップSP11とSP16に示すように別々にオーバーラップ窓掛けを施し、更に符号器50とエコーキャンセラ301でステップSP15とSP20に示すように別々に窓掛けオーバーラップ加算を実行しているから、図20で説明したと同様の遅延が発生する。

【0020】

40

更に、他の例として周波数領域係数を演算処理してノイズ抑圧処理を実行する後処理手段404は図23にも示しているように周波数分析部401を利用している。このために周波数分析部401の数が多く必要となるため、装置の規模が大きくなる欠点もある。

この発明の目的は周波数領域で反響抑圧処理のような前処理と、周波数領域で符号化、復号化処理を継続して実行する音声処理装置或はその逆に符号化、復号化処理の後に後処理として反響抑圧処理を施す音声処理装置において、処理に要する遅延量を小さくし、更に、装置規模も小さくできる音声処理装置を提供しようとするものである。

【0021】

【課題を解決するための手段】

この発明では音声信号をNサンプルからなるフレーム時刻ごと（Nは2以上の整数）に

50

2 N サンプルの切り出し窓を乗算してサンプルの取り込みが完了している 1 フレーム前と 2 フレーム前の 2 N サンプルの窓掛け音声信号を得る第 1 オーバーラップ窓掛け演算処理部と、

この第 1 オーバーラップ窓掛け演算処理部が演算処理した 2 N サンプルの窓掛け音声信号を周波数領域係数に変換する第 1 変換部と、

この第 1 変換部で変換した周波数領域係数を演算処理して反響抑圧処理を施す前処理手段と、

この前処理手段の演算処理結果を時間領域信号に変換する第 1 逆変換部と、

この第 1 逆変換部で逆変換した時間領域信号を直接、周波数領域係数に変換する第 2 変換部と、

この第 2 変換部で変換した周波数領域係数を量子化し、符号化・復号化する手段と、

この手段の処理結果を時間領域信号に変換する第 2 逆変換部と、

この第 2 逆変換部が逆変換した時間領域信号を 1 フレーム分ずつオーバーラップさせて加算し、オーバーラップしたフレームの音声信号を確定するオーバーラップ加算部と、

によって構成した音声処理装置を提案する。

#### 【 0 0 2 2 】

この発明では更に音声信号を N サンプルからなるフレーム時刻ごと ( N は 2 以上の整数 ) に 2 N サンプルの切り出し窓を乗算してサンプルの取り込みが完了している 1 フレーム前と 2 フレーム前の 2 N サンプルとの窓掛け音声信号を得るオーバーラップ窓掛け演算処理部と、

このオーバーラップ窓掛け演算処理部が演算処理した 2 N サンプルの窓掛け音声信号を周波数領域係数に変換する変換部と、

この変換部で変換した周波数領域係数を演算処理して反響抑圧処理を施す前処理手段と、

この前処理手段の処理結果に直接、符号化・復号化処理を施す手段と、

この手段で処理した処理結果を時間領域信号に変換する逆変換部と、

この逆変換部が変換した時間領域信号を 1 フレーム分ずつオーバーラップさせて加算し、オーバーラップされたフレームの音声信号を確定するオーバーラップ加算部と、

によって構成した音声処理装置を提案する。

#### 【 0 0 2 3 】

この発明では更に音声信号を N サンプルからなるフレーム時刻ごと ( N は 2 以上の整数 ) に 2 N サンプルの切り出し窓を乗算してサンプルの取り込みが完了している 1 フレーム前と 2 フレーム前の 2 N サンプルの窓掛け音声信号を得る第 1 オーバーラップ窓掛け演算処理部と、

この第 1 オーバーラップ窓掛け演算処理部が演算処理した 2 N サンプルの窓掛け音声信号を周波数領域係数に変換する第 1 変換部と、

この第 1 変換部が変換した周波数領域係数を圧縮符号化及び復号化処理する符号器及び復号器と、

復号器が復号した復号信号を時間領域信号に逆変換する逆変換部と、

この逆変換部で逆変換した時間領域信号を直接、周波数領域係数に変換する第 2 変換部と、

この第 2 変換部で変換した周波数領域係数を参照信号として利用するエコーキャンセラと、

によって構成した音声処理装置を提案する。

#### 【 0 0 2 4 】

この発明では更に音声信号を N サンプルからなるフレーム時刻ごと ( N は 2 以上の整数 ) に 2 N サンプルの切り出し窓を乗算してサンプルの取り込みが完了している 1 フレーム前と 2 フレーム前の 2 N サンプルの窓掛け音声信号を得る第 1 オーバーラップ窓掛け演算部と、

この第 1 オーバーラップ窓掛け演算部が演算処理した 2 N サンプルの窓掛け音声信号を

周波数領域係数に変換する第 1 変換部と、

この第 1 変換部が変換した周波数領域係数を圧縮符号化及び復号化処理する符号器及び復号器と、

この復号器の復号結果を直接、参照信号として利用するエコーキャンセラと、  
 によって構成した音声処理装置を提案する。

#### 【0025】

この発明では更に音声信号を N サンプルからなるフレーム時刻ごと (N は 2 以上の整数) に 2N サンプルの切り出し窓を乗算してサンプルの取り込みが完了している 1 フレーム前と 2 フレーム前の 2N サンプルの窓掛け音声信号を得る第 1 オーバーラップ窓掛け演算処理部と、

この第 1 オーバーラップ窓掛け演算処理部が演算処理した 2N サンプルの窓掛け音声信号を周波数領域係数に変換する第 1 変換部と、

この第 1 変換部で変換した周波数領域係数を演算処理して符号化・復号化処理を施す手段と、

この手段の演算処理結果を時間領域信号に変換する第 1 逆変換部と、

この第 1 逆変換部で逆変換した時間領域信号を直接、周波数領域係数に変換する第 2 変換部と、

この第 2 変換部で変換した周波数領域係数に雑音抑圧処理を施す後処理手段と、

この後処理手段の処理結果を時間領域信号に変換する第 2 逆変換部と、

この第 2 逆変換部が逆変換した時間領域信号を 1 フレーム分ずつオーバーラップさせて加算し、オーバーラップしたフレームの音声信号を確定するオーバーラップ加算部と、

によって構成した音声処理装置を提案する。

#### 【0026】

この発明では更に音声信号を N サンプルからなるフレーム時刻ごと (N は 2 以上の整数) に 2N サンプルの切り出し窓を乗算してサンプルの取り込みが完了している 1 フレーム前と 2 フレーム前の 2N サンプルとの窓掛け音声信号を得るオーバーラップ窓掛け演算処理部と、

このオーバーラップ窓掛け演算処理部が演算処理した 2N サンプルの窓掛け音声信号を周波数領域係数に変換する変換部と、

この変換部で変換した周波数領域係数を演算処理して符号化・復号化処理を施す手段と

この手段の処理結果に直接、雑音抑圧処理を施す後処理手段と、

この後処理手段で処理した処理結果を時間領域信号に変換する逆変換部と、

この逆変換部が変換した時間領域信号を 1 フレーム分ずつオーバーラップさせて加算し、オーバーラップされたフレームの音声信号を確定するオーバーラップ加算部と、

によって構成した音声処理装置を提案する。

#### 【0027】

##### 作用

この発明によれば前処理及び後処理の何れでも少なくとも窓を共通に用いることにより、前処理と後処理を同一のフレームで処理を実現することができる。この結果、遅延量を縮小することができる。

更に、前処理と後処理の双方において、窓と変換及び逆変換を共通とすることにより、前処理から後処理に至る処理ステップを減縮することができる。この結果、装置の演算量を少なくできるため、装置の規模も縮小することができる。この結果装置のコストダウンも期待することができる。

#### 【0028】

##### 【発明の実施の形態】

図 1 にこの発明の請求項 1 で提案する音声処理装置の一実施例を示す。この実施例では前処理手段としてエコーキャンセラ 301 を用いた場合を示す。エコーキャンセラ 301 は図 15 に示した構成のものとし、その入力端子側には周波数分析部 401 が挿入され、



また出力端子側には周波数合成部 4 0 3 が接続され、エコーキャンセラ 3 0 1 の内部ではフレーム単位で時間領域でオーバーラップのある窓を使って周波数領域で反響抑圧処理が行われる。

この発明では、この例のように時間領域でオーバーラップのある窓を使って周波数領域で前処理を行い、その前処理された音声信号を圧縮符号化して伝送し、復号化して再生する音声処理装置において、周波数分析部 4 0 1 で施した窓掛け処理を符号器 1 0 で共用することにより処理ステップを省略して遅延フレーム数を低減しようとするものである。

#### 【 0 0 2 9 】

図 1 に示す実施例ではエコーキャンセラ 3 0 1 の出力側に設けた周波数合成部 4 0 3 の逆変換部 5 8 の後段から、時間軸に逆変換された音声信号を取り出し、この音声信号を符号器 1 0 に入力する構成とした場合を示す。符号器 1 0 には図 1 9 で説明したように変換部 1 6 と量子化部 4 1 1 を有し、変換部 1 6 で周波数領域係数に変換し、量子化部 4 1 1 で量子化して符号化が達せられる。

10

図 1 ではこの発明の効果を主張するために、エコーキャンセラ 3 0 1 の出力側に設けた周波数合成部 4 0 3 に窓掛け部 5 9 とフレームオーバーラップ加算部 6 1 を設けている状況を示しているが、この発明を適用する場合はこれらの窓掛け部 5 9 とオーバーラップ加算部 6 1 は削除することができる。

逆変換部 5 8 の出力を直接符号器 1 0 に接続した構成とした場合には符号器 1 0 は周波数分析部 4 0 1 で窓掛け部 1 5 が施した窓をそのまま利用することができる。この結果として、符号器 1 0 の符号化処理は前処理である反響抑圧処理と同一のフレーム内で処理することができる。

20

#### 【 0 0 3 0 】

図 2 にその様子を示す。フレーム F 0 は現在音声信号を蓄積中のフレームである。フレーム F 1 と F 2 にまたがって前処理が施され、更に、その同一の窓を使って符号化処理を実行することができる。つまり、周波数合成部 4 0 3 のフレームオーバーラップ加算処理を省略した結果、1 フレームの遅延を削除することができる。

復号器 5 0 ではフレーム F 2 と F 3 にまたがる信号に関して復号化処理を実行する。つまり、復号器 5 0 の周波数合成部 4 0 3 に備えられているオーバーラップ加算部 6 1 は前回フレーム F 2 と F 3 にまたがって逆量子化した復号信号と今回フレーム F 1 と F 2 にまたがって複合処理した復号信号をフレーム F 2 に関してオーバーラップして加算することによりフレーム F 2 の音声波形を確定する。従って、図 2 1 で説明した従来の復号処理に至るまでの遅延と比較して 1 フレーム短縮することができる。

30

#### 【 0 0 3 1 】

図 3 にこの発明による音声処理装置の処理の流れを示す。図 2 2 に示した処理ステップと対応するステップには同一のステップ番号 S P 1 ~ S P 1 0 を付して示す。図 3 から分かるように前処理と符号化、復号化処理を共通の窓を使うことによりステップ S P 5 の窓掛け及びオーバーラップ加算処理と、ステップ S P 6 のオーバーラップ窓掛け処理を削除することができる。

尚、図 1 に示した実施例では前処理側では一旦逆変換部 5 8 により周波数領域係数を時間領域の信号の逆変換し、符号器 5 0 ではその逆変換した信号を再び周波数領域係数に変換しているから、前処理側の変換方式と、符号器 1 0 及び符号器 5 0 側の変換方式を揃える必要はない。

40

#### 【 0 0 3 2 】

つまり、前処理側の変換方式が M D C T であっても、符号器 1 0 の変換方式は D C T であっても F F T であってもよく、その選択は自由である。

図 4 はこの発明の請求項 2 で提案する音声処理装置の実施例を示す。この発明の請求項 2 では前処理と符号化、復号化処理を共通の窓と、共通の変換及び逆変換を行うことを特徴とするものである。

つまり、前処理の処理直後から直接符号器 1 0 の特に量子化部 4 1 1 に信号を取り出す構成とする点を特徴とするものである。この結果、この実施例ではエコーキャンセラ 3 0

50

1の出力側に設けられる周波数合成部403と符号器10に格納されている周波数分析部401は全く存在しなくてよく、エコーキャンセラ301の出力側に直接符号器10を構成する量子化部411を接続した構成とすることができる。

#### 【0033】

この構成とすることにより、エコーキャンセラ301から取り出される信号は周波数領域係数の形態であるため、符号器10には周波数分析部401を設ける必要がない。つまり、エコーキャンセラ301の出力(周波数領域係数の形態にある信号)をそのまま量子化部411に入力し、量子化することにより、直ちに符号化処理を達することができる。符号化された信号は符号器50に送られ、図1の場合と同様に復号処理され、オーバーラップ加算部61から音声信号が出力される。尚、この場合には前処理の前段に設けられる変換部16と復号器50に設ける逆変換部は同一の変換形式に揃える必要がある。

#### 【0034】

図5に図4に示した実施例の処理の流れを示す。この実施例ではステップSP4の逆変換処理からステップSP7の変換処理までを削除することができる。

図4に示した実施例によれば遅延量の短縮に関しては図1に示した実施例の場合と同じで従来より1フレーム分の遅延を短縮することができる。然し乍ら、この図4の実施例ではエコーキャンセラ301の出力側に存在するはずの周波数合成部403を全く必要としないことと、符号器10でも周波数分析部401を省略できることから、装置の規模を図1の実施例の場合より縮小することができコストの低減を期待することができる。

#### 【0035】

上述では前処理と符号化・復号化処理を縦続的に実行する場合に関してこの発明を適用した実施例を説明したが、符号化・復号化処理の後に後処理を実行する場合にも、この発明を適用することができる。つまり、図6に示すように例えばノイズキャンセラのような後処理手段404を復号器50の出力側に設けた場合にも、図23及び図24を用いて説明したように遅延量が4フレームになる問題点が発生する。

この発明の請求項3はこの欠点を解消するためにオーバーラップのある窓を使う周波数領域の符号化と復号化を行う手段と、時間領域でオーバーラップのある窓を使う周波数領域の後処理を行う手段とを縦続して音声信号の処理を行う音声処理装置において、共通の窓を使い、符号化及び復号化と後処理とを同一のフレーム内で実行する音声処理装置を提案するものである。

#### 【0036】

図6に請求項3の実施例を示す。図6に示す実施例では送信端側で符号器10で2フレームに渡って窓関数を乗算し、時間軸領域の音声信号を周波数領域係数に変換し、その変換された周波数領域係数を量子化部で量子化し、符号化処理を施す。

符号化処理された音声信号は復号器50に渡され、復号器50で逆量子化され、周波数合成部403で時間軸領域の音声信号に変換される。

この発明の請求項3では符号器、復号器及び後処理としてのノイズキャンセラのような後処理手段404の処理において、窓を共通に用いることを特徴とするものである。つまり、符号器10の窓掛け部で施された窓掛け関数をそのまま後処理側の後処理手段404でも共通に使用するものである。

#### 【0037】

このために、復号器50に設けられる周波数合成部403は逆変換部58のみが存在すればよく、この逆変換部58で時間領域の信号に逆変換した信号を後処理手段404の入力側に設けた周波数分析部401の変換部16に入力し、この変換部16で再度周波数領域係数に変換した信号で雑音抑圧処理を施し、その雑音抑圧処理後の信号を周波数合成部403で時間領域の信号に変換し、更にオーバーラップ加算部61でオーバーラップ加算して、1フレーム毎に音声信号を再生する。

#### 【0038】

図7に図6に示した実施例の処理の流れを示す。図6に示した実施例からも明らかなように、復号器50における窓掛け及びオーバーラップ加算処理を行うステップSP15と

、後処理側で実行するはずのオーバーラップ窓掛け処理を行うステップSP16を削除することができる。

従って、図6に示す実施例によれば復号器50において、オーバーラップ加算を削除し、符号化処理と復号化処理及び後処理を同一フレーム内で実行させることができるから遅延時間を1フレーム分短縮することができ、図2に示したと同様に全体で3フレームの遅延に改善することができる。

#### 【0039】

図8はこの発明の請求項4で提案する音声処理装置の実施例を示す。この発明の請求項4ではフレーム単位で時間領域でオーバーラップのある窓を使う周波数領域の符号化と復号化を行う手段と、時間領域でオーバーラップのある窓を使う周波数領域の後処理を行う手段とを縦続接続して音声信号の処理を行う音声処理装置において、

共通の窓と共通の変換及び逆変換を用いて符号化と復号処理とを同一のフレーム内で行う音声処理装置を提案する。

#### 【0040】

このために、図8に示す実施例では復号器50において、逆量子化部412の出力を直接後処理手段404に入力し、後処理手段404で後処理した後に周波数合成部403で時間領域の音声信号に再生するように構成し、符号器10で行われたオーバーラップ窓掛けと、変換処理を後処理でも共用した例を示す。

図8に示した実施例の処理の流れを図9に示す。つまり、図8に示す実施例では復号器50で行われていた逆変換処理ステップSP14と、窓掛けオーバーラップ加算処理ステップSP15を削除し、更に、後処理で行われていたオーバーラップ窓掛け処理ステップSP16と、変換処理ステップSP17とを削除することができる。

#### 【0041】

この処理ステップの削除により図2を用いて説明したのと同様に全体で3フレームの遅延に改善することができる。然も、この実施例によれば削除するステップ量が多いことから、装置の規模を大きく縮小することができる。この結果、コストの低減を期待することができる効果が得られる。

図10はこの発明の請求項5で提案する音声処理装置の実施例を示す。この発明の請求項5は参照信号 $X_r(k)$ を作成する周波数分析部401の構成を簡素化し、装置の規模を縮小することを目的とした発明である。

#### 【0042】

つまり、時間軸でオーバーラップを持たせる窓の形状が、符号化・復号化と、エコーキャンセラ301で共通であれば参照信号 $X_r(k)$ を復号化処理の途中から抜き出すことができる。図10に示す例では復号器50に使われている逆変換部58の出力信号を抜き出して周波数分析部401を構成する変換部16に入力し、変換部16で周波数領域係数に変換して参照信号 $X_r(k)$ を得るように構成した場合を示す。

図11にその様子を説明する処理の流れを示す。図10に示した実施例によれば復号化処理ステップSP14の逆変換処理が終了した信号を変換部16に入力し、この変換部16で周波数領域係数への変換ステップSP22を実行すれば周波数領域で動作するエコーキャンセラ301に入力する参照信号を作成することができる。

#### 【0043】

図10からも明らかなようにエコーキャンセラ301に参照信号 $X_r(k)$ を入力するために設けられている周波数分析部401は変換部16のみによって構成され、窓掛け部を省略できることから装置の規模を縮小することができる。但し、この実施例でも、参照信号 $X_r(k)$ を得るまでの遅延量を短縮することができるが、エコーキャンセラ301の参照信号の性格から遅延時間の短縮効果は大きな効果とは言えない。

図12はこの発明の請求項6で提案する音声処理装置の実施例を示す。請求項6では参照信号作成のための周波数分析部401を全く必要としない音声処理装置を提案するものである。

つまり、時間軸でオーバーラップを持たせる窓の形状及び周波数領域係数への変換方式

10

20

30

40

50

がエコーキャンセラ 301 と符号器 10 及び復号器 50 で共通であれば復号器 50 に設けられている逆量子化部 412 の出力をそのまま参照信号  $X_r(k)$  として利用することができる。従って、参照信号  $X_r(k)$  を作成するために必要としていた周波数分析部 401 は全く不要のものとなる。

#### 【0044】

図 13 は図 12 に示した実施例の処理の流れを示す。この実施例によれば逆量子化を実行するステップ SP13 を終了した信号を参照信号  $X_r(k)$  として利用し、その後、逆変換ステップ SP14 と窓掛けオーバーラップ加算ステップ SP15 を実行して出力音を得る。従って、ここでは従来行われていたオーバーラップ窓掛けステップ SP21 と変換ステップ SP22 は削除される。

図 12 に示す実施例によれば参照信号作成のために周波数分析部 401 は全く不要となり、装置の規模を更に一層縮小することができる。

#### 【0045】

#### 【発明の効果】

以上説明したように、この発明によれば時間軸でオーバーラップを持たせる窓と周波数領域係数を用いて符号化処理及び復号化処理と反響抑圧処理とを実行する音声処理装置の遅延を削減することができる。また処理ステップの軽減により演算量及び装置の規模を減縮することができる利点を得られる。

#### 【図面の簡単な説明】

【図 1】 この発明の請求項 1 で提案す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】 図 1 の動作を説明するための処理の流れを示すフローチャート。

【図 4】 この発明の請求項 2 で提案す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 5】 図 4 の動作を説明するための処理の流れを示すフローチャート。

【図 6】 この発明の請求項 3 で提案す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 7】 図 6 に示した装置の操作を説明するための処理の流れを示すフローチャート。

【図 8】 この発明の請求項 4 で提案す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 9】 図 8 に示した装置の動作を説明するための処理の流れを示すフローチャート。

【図 10】 この発明の請求項 5 で提案す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 11】 図 10 の動作を説明するための処理の流れを示すフローチャート。

【図 12】 この発明の請求項 6 で提案す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 13】 図 12 に示した装置の動作を説明するための処理の流れを示すフローチャート。

【図 14】 従来技術を説明するためのブロック図。

【図 15】 従来より使われているフレーム単位の窓を使用して反響抑圧処理を実行する形式のエコーキャンセラの一例を説明するためのブロック図。

【図 16】 図 15 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 17】 図 15 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 18】 従来から使用されているフレーム単位の窓を使用して符号化・復号化を実行する形式の符号器及び復号器の一例を説明するためのブロック図。

【図 19】 図 18 に示した符号器及び復号器を簡素化して説明するためのブロック図。

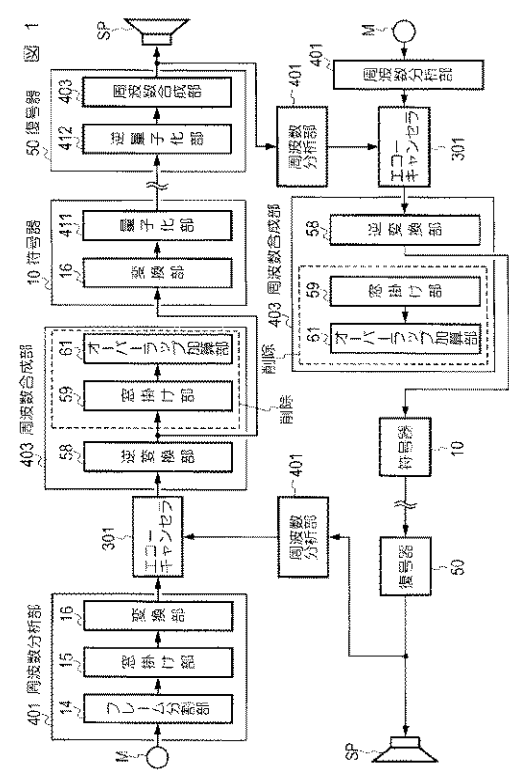
【図 20】 図 15 に示したエコーキャンセラと、図 19 に示した符号器及び復号器を組み合わせて音声処理を構成した一例を説明するためのブロック図。

- 【図 2 1】 図 2 0 に示した音声処理装置の動作を説明するためのタイミングチャート。
- 【図 2 2】 図 2 0 に示した音声処理装置の不都合を説明するための処理の流れを説明するためのフローチャート。
- 【図 2 3】 従来の音声処理装置の変形を説明するためのブロック図。
- 【図 2 4】 図 2 3 に示した音声処理装置の動作を説明するためのフローチャート。

【符号の説明】

1 0	符号器	5 8	逆変換部
1 4	フレーム分割部	5 0	復号器
1 5	窓掛け部	5 9	窓掛け部
1 6	変換部	6 1	オーバーラップ加算部
3 0 1	エコーキャンセラ	4 0 1	周波数分析部
4 0 3	周波数合成部	4 0 4	後処理手段

【図 1】



【図 2】

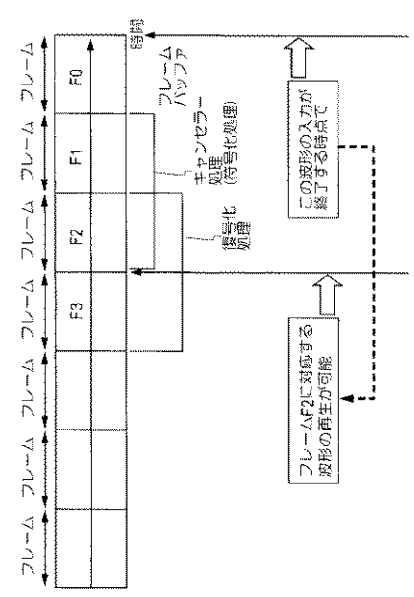


図 2

【図3】

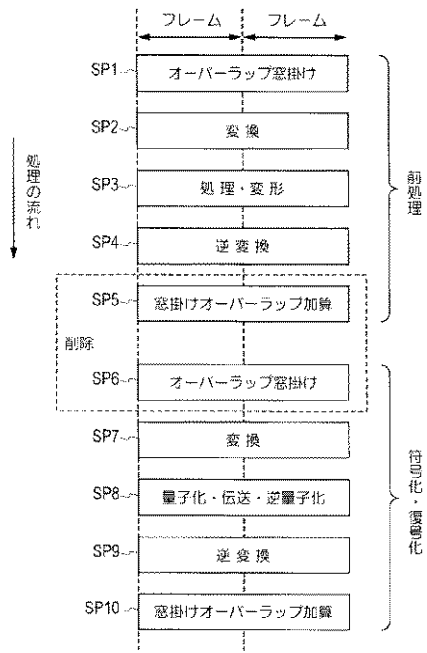
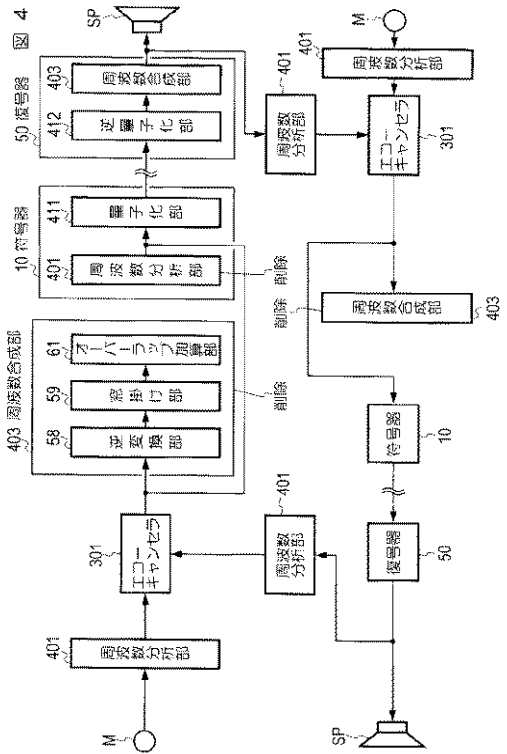


図 3

【図4】



【図5】

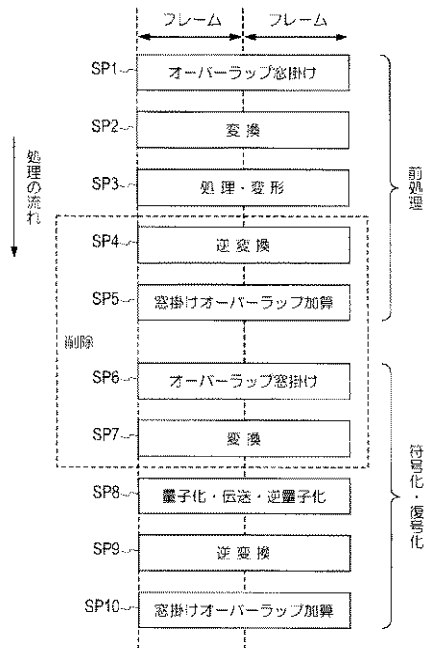
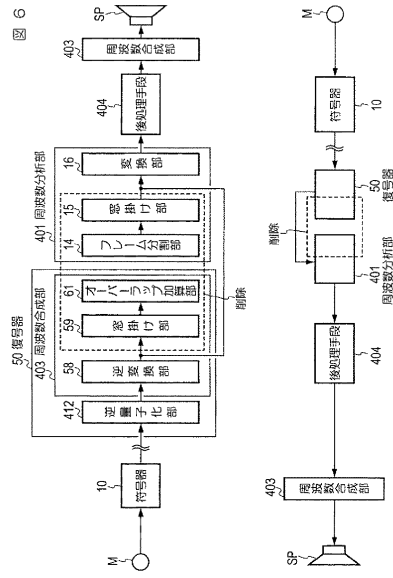


図 5

【図6】



【図7】

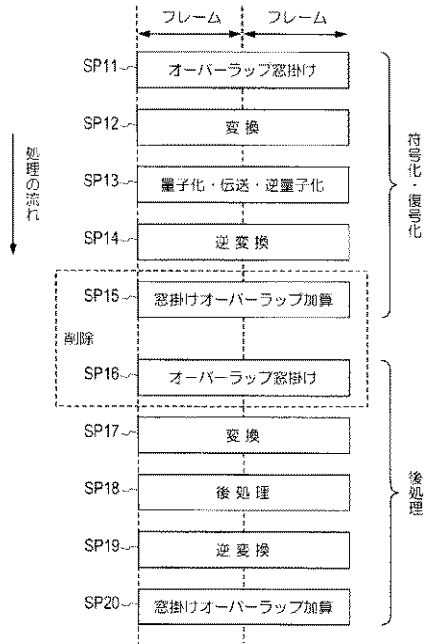
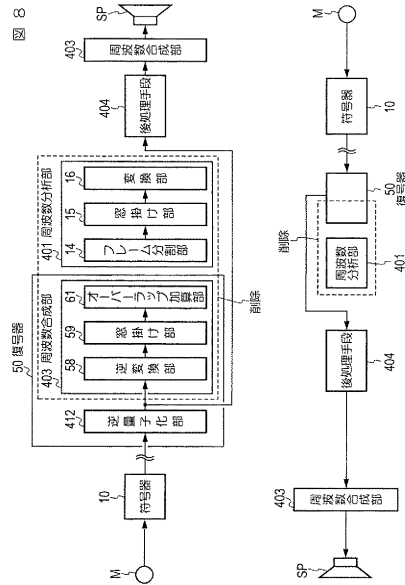


図 7

【図8】



【図9】

【図10】

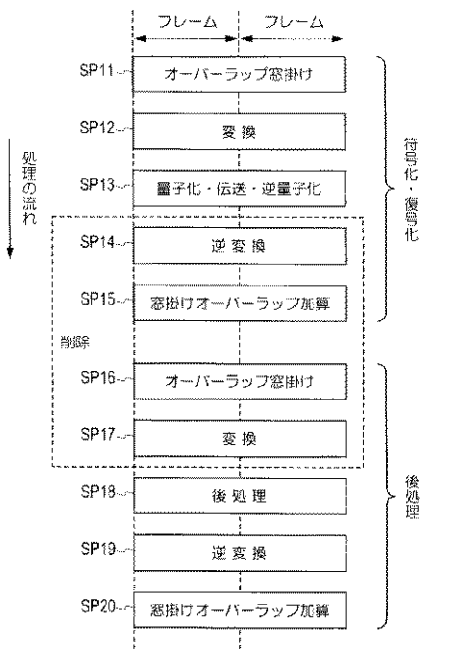
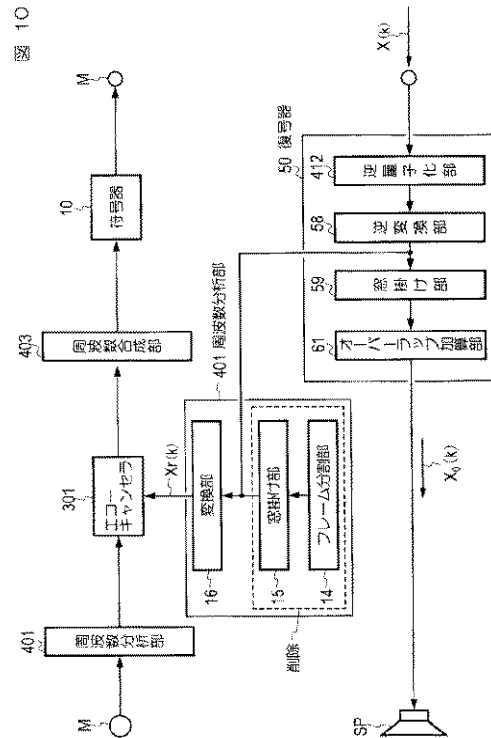
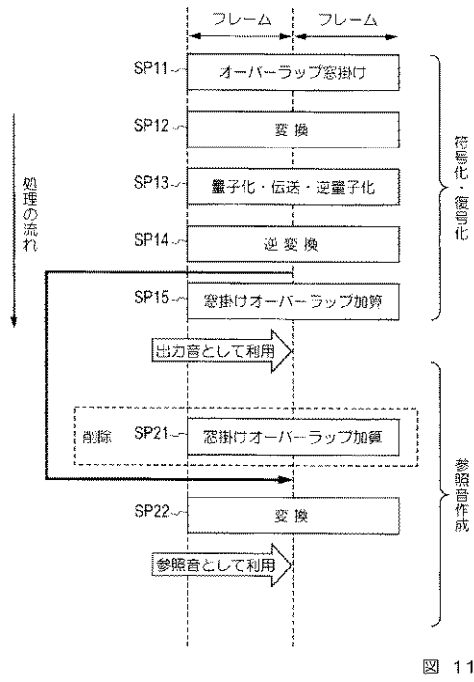


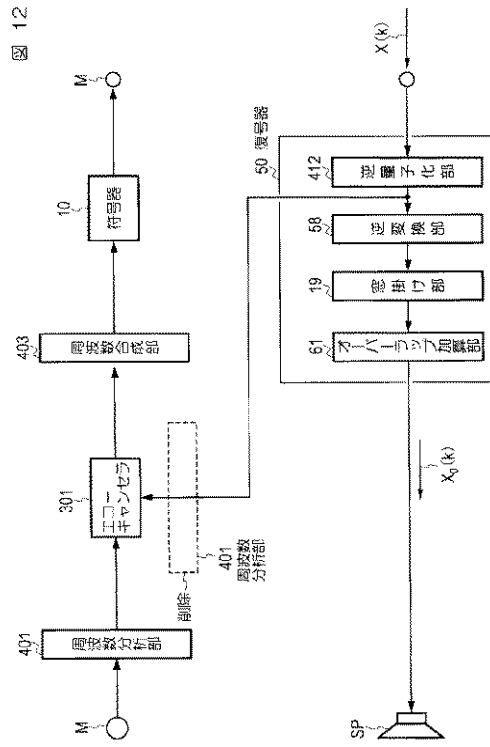
図 9



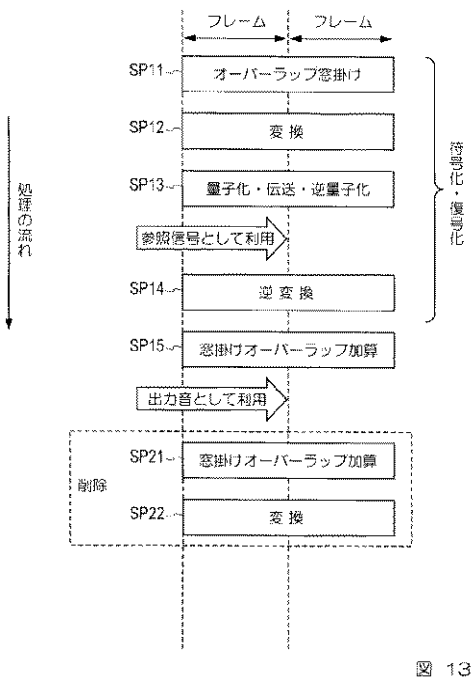
【図 1 1】



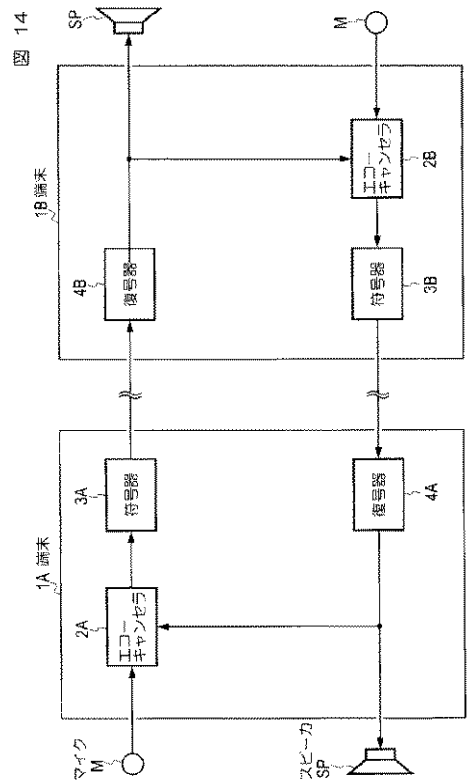
【図 1 2】



【図 1 3】

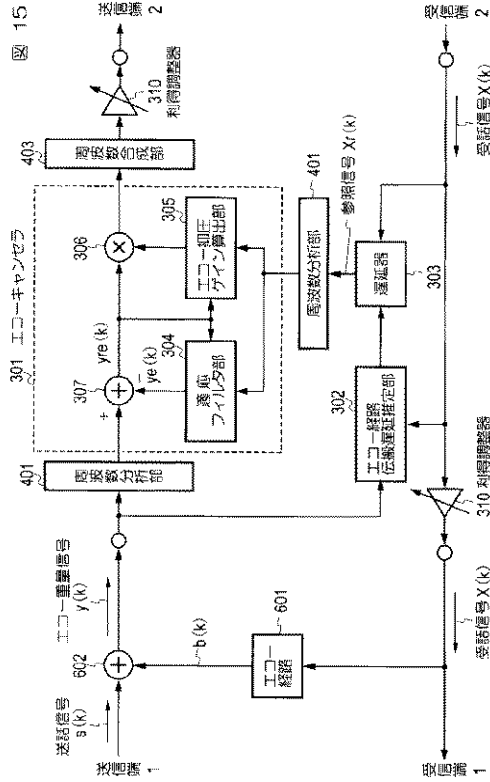


【図 1 4】





【図 15】



【図 16】

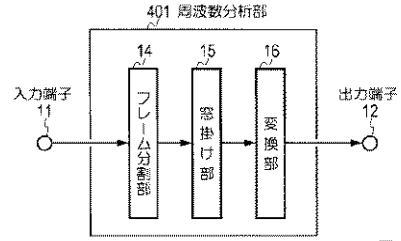


図 16

【図 17】

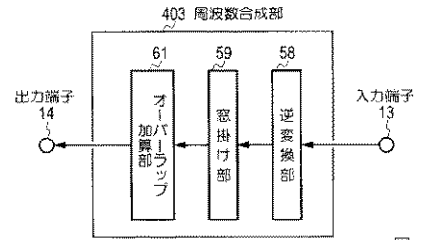


図 17

【図 18】

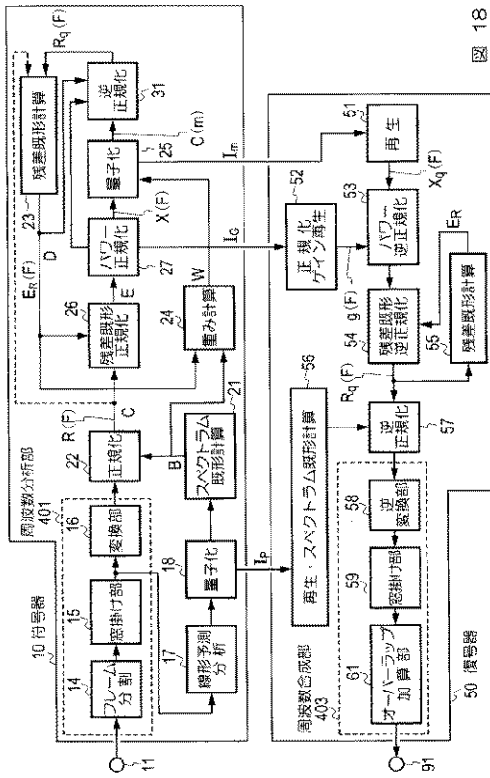
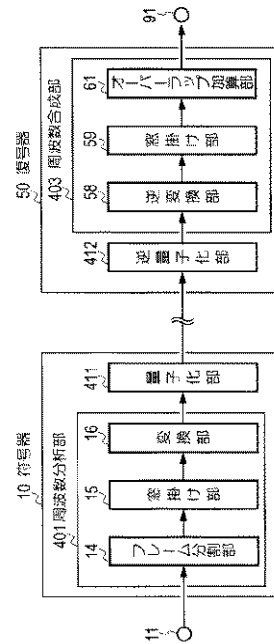


図 18

【図 19】

図 19



【図20】

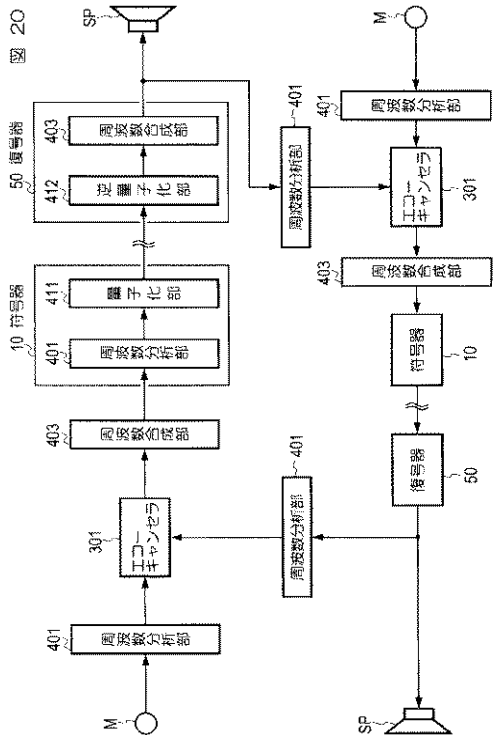


図 20

【図21】

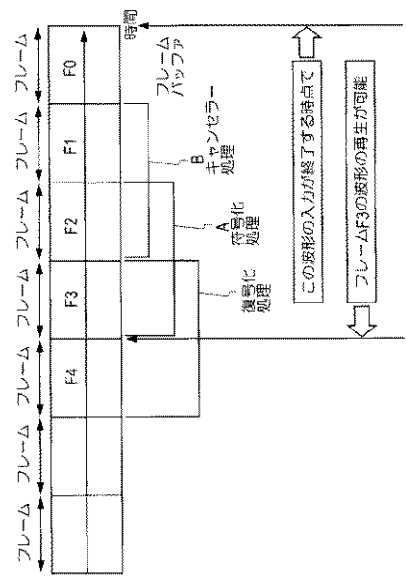


図 21

【図22】

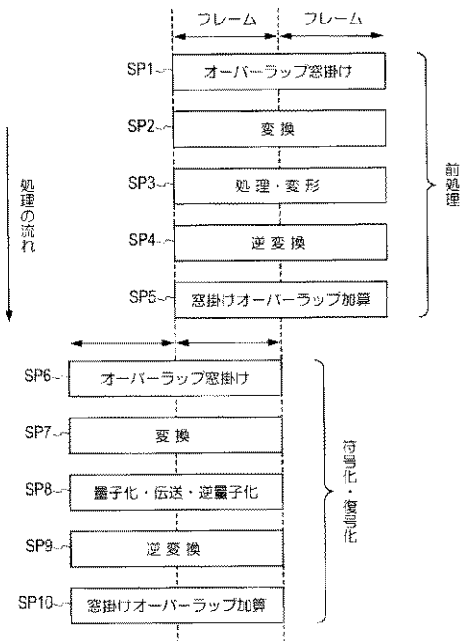


図 22

【図23】

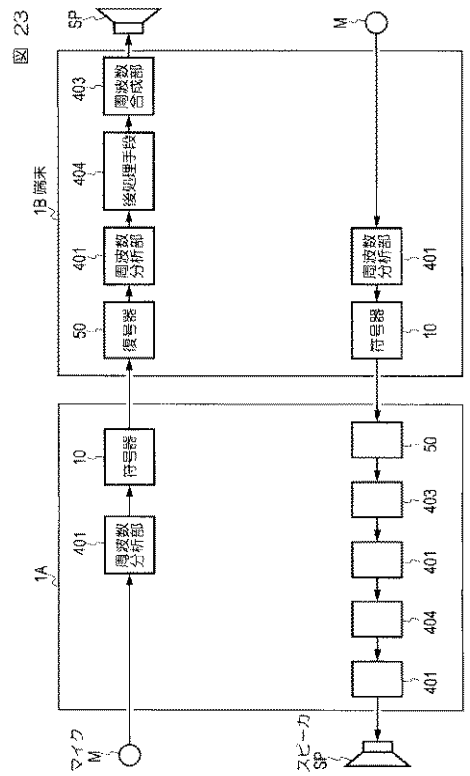


図 23

【図 24】

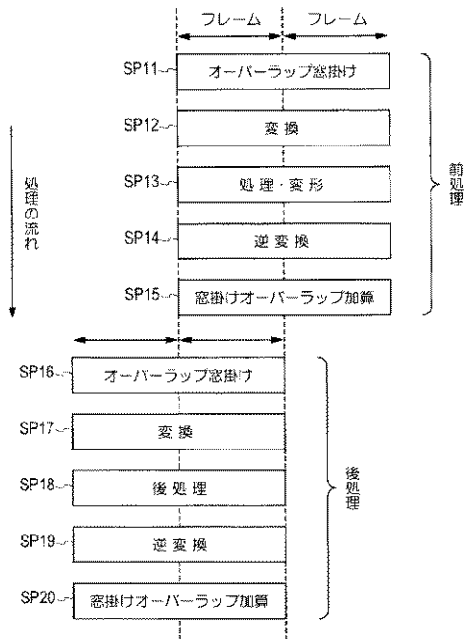


図 24

-----  
フロントページの続き

(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

審査官 山下 剛史

(56)参考文献 国際公開第01/003316(WO, A1)

特開平09-204196(JP, A)

(58)調査した分野(Int.Cl., DB名)

G10L 19/00-21/02

H04R 3/02

H04B 3/23