

(19)日本国特許庁(JP)

(12)特許公報 ( B 2 )

(11)特許番号

特許第 3 3 4 8 7 5 9 号

( P 3 3 4 8 7 5 9 )

(45)発行日 平成14年11月20日(2002.11.20)

(24)登録日 平成14年9月13日(2002.9.13)

(51)Int. Cl. <sup>7</sup>	識別記号	F I		
G 1 0 L	19/02	G 1 0 L	7/04	G
	11/00			B
	11/04		9/16	
	19/00		9/18	E

請求項の数 8

(全 8 頁)

(21)出願番号	特願平7-248145	(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
(22)出願日	平成7年9月26日(1995.9.26)	(72)発明者	岩上 直樹 東京都千代田区内幸町1丁目1番6号 日本 電信電話株式会社内
(65)公開番号	特開平9-90989	(72)発明者	守谷 健弘 東京都千代田区内幸町1丁目1番6号 日本 電信電話株式会社内
(43)公開日	平成9年4月4日(1997.4.4)	(72)発明者	三樹 聡 東京都千代田区内幸町1丁目1番6号 日本 電信電話株式会社内
審査請求日	平成10年11月17日(1998.11.17)	(74)代理人	100064908 弁理士 志賀 正武
		審査官	山下 剛史

最終頁に続く

(54)【発明の名称】変換符号化方法および変換復号化方法

1

(57)【特許請求の範囲】

【請求項 1】 楽音信号あるいは音声信号を一定時間間隔のフレームに分割し、各フレームに時間 - 周波数変換を施して周波数領域信号を生成する時間 - 周波数変換段階と、  
周波数領域信号の概形信号を生成し、当該概形信号を量子化して量子化概形インデックスを出力する概形計算・量子化段階と、  
前記周波数領域信号を量子化された前記概形信号により除算して平坦化信号を生成する平坦化段階と、  
前記平坦化信号からピッチ成分を検出して量子化し量子化ピッチ成分インデックスを出力するピッチ符号化段階と、  
前記平坦化信号から量子化した前記ピッチ成分を除去した平坦化信号の量子化平坦化信号インデックスを出力す

2

る平坦化信号量子化段階とを有することを特徴とする変換符号化方法。

【請求項 2】 ピッチ符号化段階では、  
ピッチ成分の基本周波数を求めて量子化し、  
周波数領域信号から前記基本周波数の自然数倍の周波数またはこの周波数に最も近い周波数のサンプルをピッチ成分サンプルとして抽出して量子化し、  
このようにして得られた量子化ピッチ基本周波数インデックスと量子化ピッチ成分サンプルインデックスとを量子化ピッチ成分インデックスとして出力することを特徴とする請求項 1 記載の変換符号化方法。

10

【請求項 3】 ピッチ符号化段階では、  
周波数領域信号から基本周波数の自然数倍の周波数あるいはこの周波数に最も近い周波数のサンプル及びこれを含めた連続する複数のサンプルを 1 単位としてピッチ成

分を抽出して量子化することを特徴とする請求項 2 記載の変換符号化方法。

【請求項 4】 ピッチ符号化段階では、ピッチ成分サンプルを一括または各単位毎にベクトル量子化することを特徴とする請求項 2 または 3 記載の変換符号化方法。

【請求項 5】 量子化平坦化信号インデックスから平坦化信号を再生する平坦化信号再生段階と、

量子化ピッチ成分インデックスからピッチ成分を再生するピッチ再生段階と、

量子化概形インデックスから概形信号を再生する概形信号再生段階と、

前記平坦化信号に前記ピッチ成分を加えた信号を前記概形信号で逆平坦化して周波数領域信号を再生する逆平坦化段階と、

前記周波数領域信号に時間 - 周波数逆変換を施して時間領域の楽音信号あるいは音声信号を生成する時間 - 周波数逆変換段階とを有することを特徴とする変換復号化方法。

【請求項 6】 ピッチ再生段階では、ピッチ成分の基本周波数を復号化し、

ピッチ成分として量子化ピッチ成分サンプルインデックスから復号したピッチ成分サンプルを前記基本周波数の自然数倍の周波数またはこの周波数に最も近い周波数に周波数領域信号として配置することを特徴とする請求項 5 記載の変換復号化方法。

【請求項 7】 ピッチ再生段階では、ピッチ成分として量子化ピッチ成分サンプルインデックスから復号したピッチ成分サンプルを基本周波数の自然数倍の周波数またはこの周波数に最も近い周波数のサンプル及びこれを含めた連続する複数のサンプルを 1 単位として周波数領域信号として配置することを特徴とする請求項 6 記載の変換復号化方法。

【請求項 8】 ピッチ再生段階では、ピッチ成分サンプルを一括または各単位毎に復号することを特徴とする請求項 6 または 7 記載の変換復号化方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、楽音信号あるいは音声信号等、ピッチ成分を含む信号の変換符号化方法および変換復号化方法に関する。

【0002】

【従来の技術】現在、楽音信号あるいは音声信号等のオーディオ信号を高効率に符号化する方法として、該オーディオ信号をフレームと呼ばれる 5 ~ 50 ms 程度の一定間隔の区間に分割し、この 1 フレームの信号に時間 - 周波数変換を施して得られた周波数領域信号を、周波数特性の包絡形状（周波数特性の概形）と、周波数領域信号を周波数特性概形で平坦化して得られる残差信号の 2 つの情報に分離し、それぞれを符号化することが提案されている。

【0003】また、このような符号化法の具体的な方法として、適応スペクトル聴感制御エントロピー符号化法（ASPEC, Adaptive Spectral Perceptual Entropy Coding）、重み付きベクトル量子化による変換符号化法（TCWVQ, Transform Coding with Weighted Vector Quantization）、およびエムペグ - オーディオ・レイヤ 3 方式（MP3 - Audio Layer 3）等が提案されている。

【0004】なお、これらの技術については、K.Brandenburg, J.Herre, J.D.Johnston et al: "ASPEC: Adaptive spectral entropy coding of high quality music signals", Proc. AES'91、T.Moriya, H.Suda: "An 8 Kbit/s transform coder for noisy channels", Proc. ICASSP'89 pp.196--199、および ISO/IEC 標準 IS-11172-3 に詳しく述べられている。

【0005】ここで、これらの符号化法によって高効率な符号化を実現するためには、残差信号は、できるだけ周波数特性が平坦であることが望ましい。このため、上述の適応スペクトル聴感制御エントロピー符号化法（ASPEC）あるいはエムペグ - オーディオ・レイヤ 3 方式（MP3 - Audio Layer 3）では、図 5 に示すように、周波数領域信号をいくつかの小帯域に分割し、各小帯域内の信号を帯域の強さを表すスケーリングファクタと呼ばれる値で除算して正規化することにより、残差信号の周波数特性の平坦化を図っている。

【0006】一方、これらの方法よりも高効率な周波数領域信号の平坦化方法として、図 6 に示すような線形予測分析を用いる方法がある。この方法では、入力信号を線形予測して得られた線形予測係数で線形予測分析フィルタを駆動することにより周波数特性の平坦化を行う。この方法は、上記重み付きベクトル量子化による変換符号化法（TCWVQ）で用いられている手法である。

【0007】なお、線形予測分析、離散コサイン変換（DCT）、変形離散コサイン変換（MDCT）等の各関連各技術については、斉藤、中田 "音声情報処理の基礎"（オーム社）の第 6 章、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"（オーム社）の第 2 章、H.S.Malvar, "Signal Processing with Lapped Transforms", Artech House、および ISO/IEC 標準 IS-11172-3 に記載されている。

【0008】

【発明が解決しようとする課題】しかし、これらの符号化方法では、周波数特性の大局的な概形を正規化するととどまり、楽音や音声のピッチ成分による微視的な周波数特性の凹凸を能率良く除去することができない。したがって、このことが障害となり、上記従来の符号化方法は、ピッチ成分の強いオーディオ信号を符号化する場合に高効率化することが困難であった。

【0009】本発明は、上述する問題点を鑑みてなされたもので、ピッチ成分が含まれたオーディオ信号を能率良く符号化することが可能な変換符号化方法および変換

復号化方法を提供することを目的としている。

【0010】

【課題を解決するための手段】請求項1記載の発明は、楽音信号あるいは音声信号を一定時間間隔のフレームに分割し、各フレームに時間-周波数変換を施して周波数領域信号を生成する時間-周波数変換段階と、周波数領域信号の概形信号を生成し、当該概形信号を量子化して量子化概形インデックスを出力する概形計算・量子化段階と、前記周波数領域信号を量子化された前記概形信号により除算して平坦化信号を生成する平坦化段階と、前記平坦化信号からピッチ成分を検出して量子化し量子化ピッチ成分インデックスを出力するピッチ符号化段階と、前記平坦化信号から量子化した前記ピッチ成分を除去した平坦化信号の量子化平坦化信号インデックスを出力する平坦化信号量子化段階とを有することを特徴としている。

【0011】請求項2記載の発明は、請求項1記載の発明において、ピッチ符号化段階では、ピッチ成分の基本周波数を求めて量子化し、周波数領域信号から前記基本周波数の自然数倍の周波数またはこの周波数に最も近い周波数のサンプルをピッチ成分サンプルとして抽出して量子化し、このようにして得られた量子化ピッチ基本周波数インデックスと量子化ピッチ成分サンプルインデックスとを量子化ピッチ成分インデックスとして出力することを特徴としている。

【0012】請求項3記載の発明は、請求項2記載の発明において、ピッチ符号化段階では、周波数領域信号から基本周波数の自然数倍の周波数あるいはこの周波数に最も近い周波数のサンプル及びこれを含めた連続する複数のサンプルを1単位としてピッチ成分を抽出して量子化することを特徴としている。

【0013】請求項4記載の発明は、請求項2または3記載の発明において、ピッチ符号化段階では、ピッチ成分サンプルを一括または各単位毎にベクトル量子化することを特徴としている。

【0014】請求項5記載の発明は、量子化平坦化信号インデックスから平坦化信号を再生する平坦化信号再生段階と、量子化ピッチ成分インデックスからピッチ成分を再生するピッチ再生段階と、量子化概形インデックスから概形信号を再生する概形信号再生段階と、前記平坦化信号に前記ピッチ成分を加えた信号を前記概形信号で逆平坦化して周波数領域信号を再生する逆平坦化段階と、前記周波数領域信号に時間-周波数逆変換を施して時間領域の楽音信号あるいは音声信号を生成する時間-周波数逆変換段階とを有することを特徴としている。

【0015】請求項6記載の発明は、請求項5記載の発明において、ピッチ再生段階では、ピッチ成分の基本周波数を復号化し、ピッチ成分として量子化ピッチ成分サンプルインデックスから復号したピッチ成分サンプルを前記基本周波数の自然数倍の周波数またはこの周波数に

最も近い周波数に周波数領域信号として配置することを特徴としている。

【0016】請求項7記載の発明は、請求項6記載の発明において、ピッチ再生段階では、ピッチ成分として量子化ピッチ成分サンプルインデックスから復号したピッチ成分サンプルを基本周波数の自然数倍の周波数またはこの周波数に最も近い周波数のサンプル及びこれを含めた連続する複数のサンプルを1単位として周波数領域信号として配置することを特徴としている。

【0017】請求項8記載の発明は、請求項6または7記載の発明において、ピッチ再生段階では、ピッチ成分サンプルを一括または各単位毎に復号することを特徴としている。

【0018】

【作用】楽音あるいは音声は、ピッチすなわち音程の高/低を有する。この楽音あるいは音声を周波数変換して得られる周波数領域信号には、一定の周波数間隔で並ぶピッチ成分が含まれる。したがって、該周波数領域信号を自らの周波数特性の概形で正規化して得られる残差信号にも、上記ピッチ成分が含まれている。このピッチ成分は、全体のパワーに対してエネルギーの大きいスパイクとなって現れるので、残差信号の平坦度を落として量子化能率を悪化させる。しかし、本発明は、ピッチ成分が周波数軸上で等間隔に並んでいる点に着目し、ピッチ成分を残差信号から差し引くことにより、少ない付加情報量で残差係数の平坦度を高める。

【0019】

【発明の実施の形態】以下、図面を参照して本発明の一実施形態について説明する。図1は、本実施形態による変換符号化方法および変換復号化方法を説明する図であり、符号Aは符号器、またBは復号器である。図示するように、符号器Aは、時間-周波数変換器1、大局的概形計算・量子化器2、第1平坦化器3、ピッチ符号化器4、加算器5、微細スペクトル概形計算・量子化器6、第2平坦化器7、および量子化器8によって構成されている。

【0020】時間-周波数変換器1は、時間領域の入力信号（楽音信号あるいは音声信号等のオーディオ信号）を一定時間間隔のフレームに分割し、各々のフレームに時間-周波数変換を施して周波数領域信号を生成する。

【0021】図2は、この周波数領域信号の周波数特性を示したものである。この図に示すように、楽音信号あるいは音声信号の周波数領域信号は、一定周波数間隔pで配列するピッチ成分が含まれている。なお、変換手法としては、離散コサイン変換（Discrete Cosine Transformation, DCT）や変形離散コサイン変換（Modified Discrete Cosine Transformation, MDCT）を用いることができる。

【0022】大局的概形計算・量子化器2は、上記時間-周波数変換器1から出力された周波数領域信号の大局

的な概形を示す信号を生成し、また量子化する。そして、この信号を上記第1平坦化器3に出力すると共に、量子化大局的概形インデックスとして外部に出力する。該大局的概形の算出手法としては、線形予測スペクトル、あるいは周波数領域信号を複数のサブバンドに分割し、各バンドの代表値によって周波数領域信号全体の概形を表現するスケールファクタを用いてもよい。

【0023】なお、線形予測スペクトルを量子化する場合は、線形予測パラメータをLSPパラメータに変換して量子化する。またはKパラメータに変換して量子化する。

【0024】第1平坦化器3は、上記時間-周波数変換器1から出力された周波数領域信号を大局的概形計算・量子化器2から出力された上記大局的概形信号によって除算することにより平坦化し、第1の平坦化信号を出力する。

【0025】次に、ピッチ符号化器4は、上記第1の平坦化信号からピッチ成分を検出して符号化する。また、図3はピッチ符号化器4の詳細を示す図であり、上記第1の平坦化信号は、図示するピッチ基本周波数抽出器4aおよびピッチサンプル抽出器4bに入力される。

【0026】このピッチ基本周波数抽出器4aは、第1の平坦化信号を分析することによりピッチ成分の基本周波数(ピッチ基本周波数)を求める。すなわち、ピッチ基本周波数抽出器4aは、第1の平坦化係数のケプストラムを計算し、その最大値をピッチ成分の基本周期とする。そして、該基本周期の逆数を演算することによりピッチ基本周波数を求め、ピッチ基本周波数量子化器4cに出力する。

【0027】なお、ピッチ基本周波数をより正確にするために、求められたピッチ基本周波数の前後で、ピッチ基本周波数ごとの第1の平坦化信号のサンプルのパワーの総和が最大になる基本周波数を検索し、新たにこれをピッチ基本周波数としてもよい。

【0028】ピッチ基本周波数量子化器4cは、このようにして求められたピッチ基本周波数を量子化する。すなわち、このピッチ基本周波数量子化器4cは、ピッチ基本周波数の対数値をスカラ量子化し、量子化ピッチ基本周波数インデックスとして外部に出力すると共に、このスカラ量子化された信号を上記ピッチサンプル抽出器4bに出力する。

【0029】このピッチサンプル抽出器4bは、第1平坦化器3から入力された第1の平坦化信号に対して、ピッチ基本周波数量子化器4cから入力された量子化ピッチ基本周波数の自然数倍の周波数に最も近いサンプルを中心として前後1サンプルを抽出し、この3サンプル組を一本のピッチ成分のサンプル群としてピッチサンプル量子化器4dに出力する。なお、このピッチ成分のサンプル群の数は、固定値でも良いし、可変としても良い。

【0030】ピッチサンプル量子化器4dは、上記ピッチ成分のサンプル群を量子化して量子化ピッチ成分インデックスとして外部に出力すると共に、この量子化ピッチ成分インデックスを復号した量子化ピッチ成分を上記加算器5に出力する。なお、該サンプル群の量子化は、スカラ量子化であっても良いし、3サンプルからなるサンプル群ごとにベクトル量子化してもよい。また、全サンプル群を一括でベクトル量子化しても良い。以上がピッチ符号化器4において行われる処理である。

【0031】次に、加算器5は、該ピッチ符号化器4から入力されたピッチ成分の量子化信号を用いて、第1平坦化器3から入力された第1の平坦化信号からピッチ成分のみを差し引いて第2の平坦化信号を生成し、微細スペクトル概形計算・量子化器6および第2平坦化器7に出力する。

【0032】ここで、図4は、この第2の平坦化信号の周波数特性を示す図である。上記図2との比較でわかるように、第2の平坦化信号は、時間-周波数変換器1から出力された周波数領域信号からピッチ成分を除去したものとなる。

【0033】微細スペクトル概形計算・量子化器6は、該第2の平坦化信号から微細なスペクトルの概形(微細スペクトル概形)を計算し、これを量子化する。そして、この量子化した信号を量子化微細スペクトル概形インデックスとして外部に出力すると共に、第2平坦化器7に出力する。

【0034】この微細スペクトル概形は、微細スペクトル概形を直接量子化して求めてもよいし、過去のフレームの微細スペクトル概形を線形合成して求めてもよい。また、過去および現在のフレームの量子化された微細スペクトル概形の情報を線形合成して求めてもよい。さらに、この微細スペクトル概形は、例えば、第2の平坦化信号の絶対値に3から5程度の幅の窓関数を畳み込んだものを用いてもよいし、サブバンド分割した第2の平坦化信号の振幅の代表値を各バンドごとに用意し、これを概形としてもよい。

【0035】第2平坦化器7は、加算器5から入力された第2の平坦化信号を微細スペクトル概形計算・量子化器6で得られた微細スペクトル概形で除算して平坦化し、第3の平坦化信号として量子化器8に出力する。この量子化器8は、該第3の平坦化信号をスカラ量子化あるいはベクトル量子化し、量子化インデックスとして外部に出力する。

【0036】なお、ベクトル量子化する場合は、フレームの全サンプルを一括で量子化してもよいが、フレームのサンプル列を複数のサブベクトルに分割して、このサブベクトルごとに量子化する方が演算量の面で現実的である。また、分割の方法は、単純なサブバンド分割でもよいし、サンプルをインタリーブしてから分割するインタリーブ分割でもよい。また、量子化の際必要な情報量

にあわせて適応的ビット割り当てをしてもよい。

【0037】次に、復号器Bについて説明する。図1に示すように、復号器Bは、再生器9、微細スペクトル概形再生器10、第1逆平坦化器11、ピッチ再生器12、加算器13、大局的概形再生器14、第2逆平坦化器15、および時間-周波数逆変換器16によって構成されている。

【0038】このうち、再生器9は、上記符号器Aから伝送されてきた量子化インデックスから上記第3の平坦化信号を再生する。この再生器9は、上記量子化器8の逆処理を行うことにより第3の平坦化信号を再生し、第1逆平坦化器11に出力する。微細スペクトル概形再生器10は、符号器Aから伝送されてきた微細スペクトル概形量子化インデックスから微細スペクトル概形を再生する。

【0039】第1逆平坦化器11は、再生器9から入力された第3の平坦化信号に微細スペクトル概形を付加して、上記第2の平坦化信号を再生して加算器13に出力する。また、ピッチ再生器12は、符号器Aから伝送されてきた量子化ピッチ成分インデックスおよび量子化ピッチ基本周波数インデックスから上記ピッチ成分を再生し、加算器13に出力する。

【0040】加算器13は、第1逆平坦化器11から入力された第2の平坦化信号に、ピッチ再生器12から入力されたピッチ成分を加えて上記第1の平坦化信号を再生し、第2逆平坦化器15に出力する。また、大局的概形再生器14は、符号器Aから伝送されてきた量子化大局的概形インデックスから上記大局的概形を再生し、第2逆平坦化器15に出力する。

【0041】第2逆平坦化器15は、加算器13から入力された第1の平坦化信号に、大局的概形再生器14から入力された大局的概形を付加し、上記周波数領域信号を生成する。そして、時間-周波数逆変換器16は、該第2逆平坦化器15から入力された周波数領域信号に時間-周波数逆変換を施して復号し、時間領域の音声信号あるいは楽音信号を出力する。

【0042】

【発明の効果】以上説明したように、本発明によれば、ピッチ成分を有する楽音信号あるいは音声信号を符号化する際に、該信号を周波数領域に変換した周波数領域信号に現れるスパイク状のピッチ成分の規則性を利用して、これを高能率に符号化する。したがって、より平坦化された残差係数を得ることができ、符号化器全体の能率を高めることが可能である。

【図面の簡単な説明】

【図1】本発明の一実施形態を示す符号器および復号器を説明する図である。

【図2】本発明において時間-周波数変換器の出力信号の周波数特性を示す図である。

【図3】本発明においてピッチ符号化器の詳細構成を示す図である。

【図4】本発明において第2平坦化信号の周波数特性を示す図である。

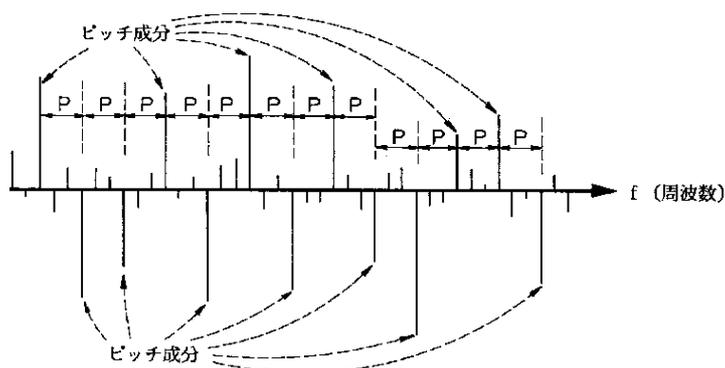
【図5】従来の変換符号化方法を説明する第1の図である。

【図6】従来の変換符号化方法を説明する第2の図である。

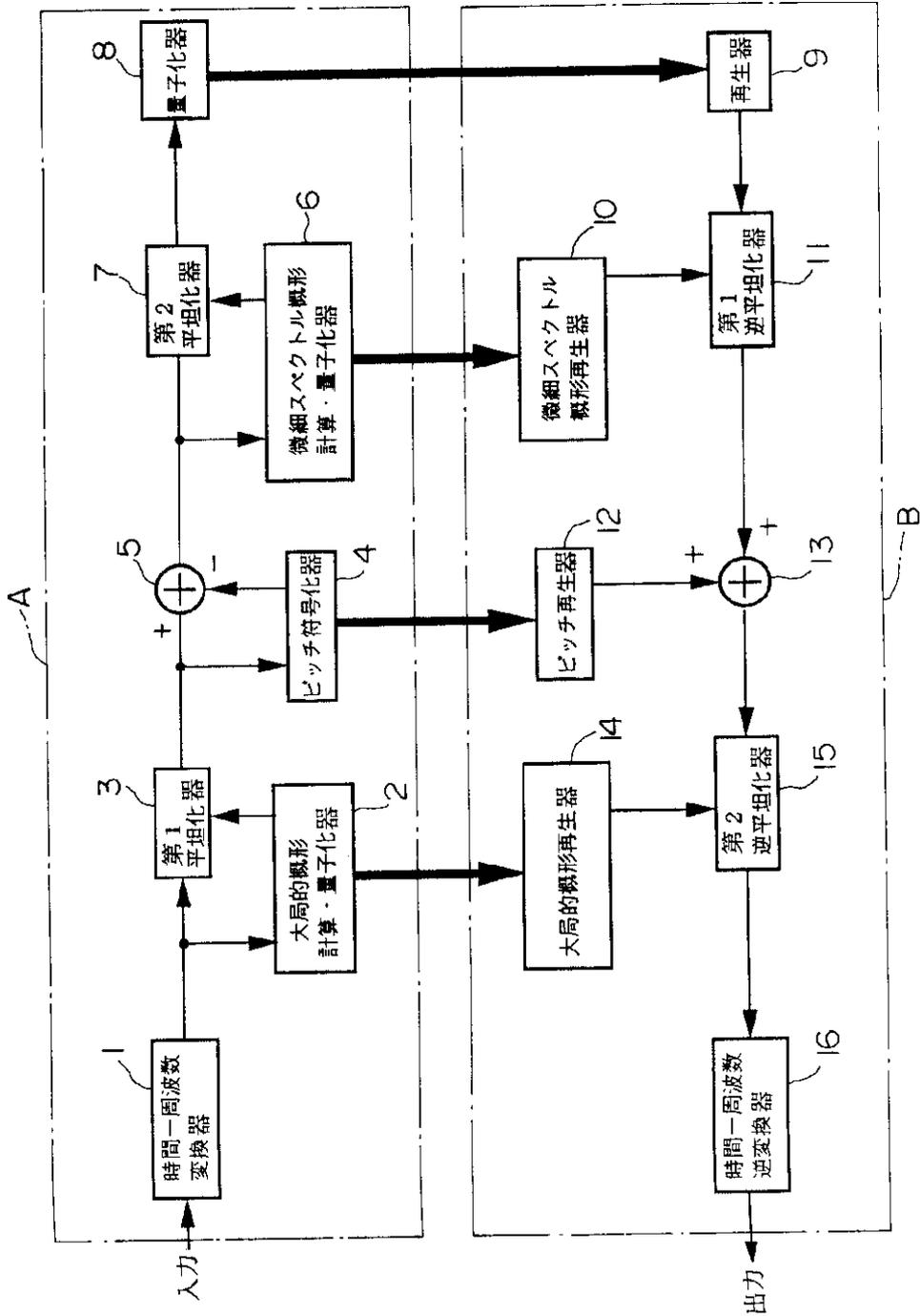
【符号の説明】

- 1 時間-周波数変換器
- 2 大局的概形計算・量子化器
- 3 第1平坦化器
- 4 ピッチ符号化器
- 5、13 加算器
- 6 微細スペクトル概形計算・量子化器
- 7 第2平坦化器
- 8 量子化器
- 9 再生器
- 10 微細スペクトル概形再生器
- 11 第1逆平坦化器
- 12 ピッチ再生器
- 14 大局的概形再生器
- 15 第2逆平坦化器
- 16 時間-周波数逆変換器

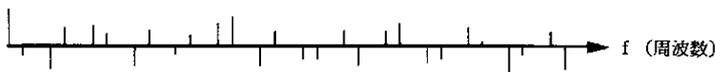
【図2】



【図 1】

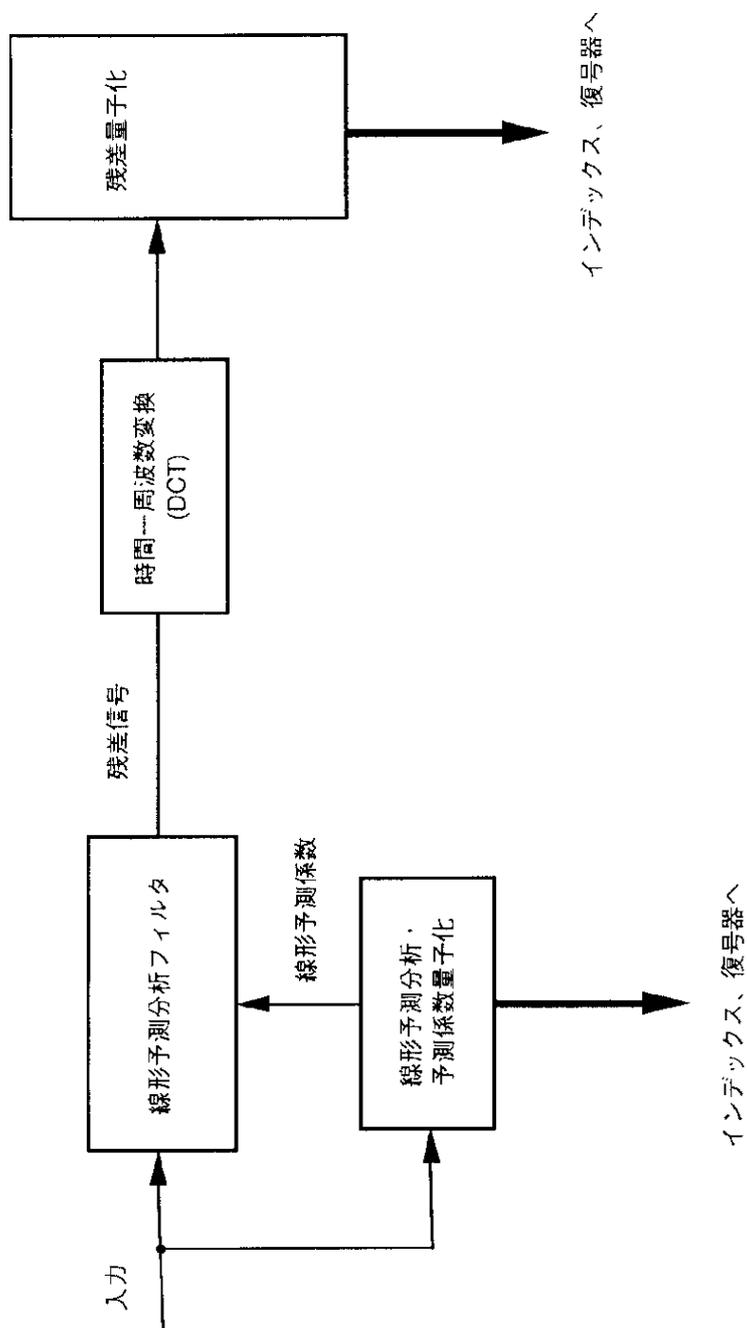


【図 4】





【図 6】



フロントページの続き

(56) 参考文献  
 特開 平 1 - 239597 ( J P , A )  
 特開 昭 57 - 62096 ( J P , A )  
 特開 昭 57 - 161795 ( J P , A )  
 特開 昭 63 - 37400 ( J P , A )  
 特開 平 7 - 261800 ( J P , A )

(58) 調査した分野 (Int. Cl. 7, D B 名)  
 G10L 11/00 - 11/04  
 G10L 19/00 - 19/02  
 G11B 20/10  
 H03M 7/30