

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5223786号
(P5223786)

(45) 発行日 平成25年6月26日 (2013. 6. 26)

(24) 登録日 平成25年3月22日 (2013. 3. 22)

(51) Int. Cl. F I
G 1 O L 21/0388 (2013. 01) G 1 O L 21/04 I 3 O A
G 1 O L 19/06 (2013. 01) G 1 O L 19/06 Z

請求項の数 5 (全 24 頁)

(21) 出願番号	特願2009-139390 (P2009-139390)	(73) 特許権者	000005223
(22) 出願日	平成21年6月10日 (2009. 6. 10)		富士通株式会社
(65) 公開番号	特開2010-286608 (P2010-286608A)		神奈川県川崎市中原区上小田中4丁目1番1号
(43) 公開日	平成22年12月24日 (2010. 12. 24)	(74) 代理人	100099759
審査請求日	平成24年2月14日 (2012. 2. 14)		弁理士 青木 篤
		(74) 代理人	100119987
			弁理士 伊坪 公一
		(74) 代理人	100081330
			弁理士 樋口 外治
		(74) 代理人	100141254
			弁理士 榎原 正巳
		(74) 代理人	100113826
			弁理士 倉地 保幸

最終頁に続く

(54) 【発明の名称】 音声帯域拡張装置、音声帯域拡張方法及び音声帯域拡張用コンピュータプログラムならびに電話機

(57) 【特許請求の範囲】

【請求項 1】

第1の周波数帯域を有する音声信号を、所定の時間長を有するフレームごとに時間周波数変換することにより、該音声信号の周波数スペクトルを算出する時間周波数変換部と、

前記周波数スペクトルから、該周波数スペクトルの包絡振幅スペクトルと、周波数に対して周期的にスペクトル強度が変動する周期性振幅スペクトルと、周波数に対してランダムにスペクトル強度が変動するランダム振幅スペクトルとを抽出する分離部と、

前記包絡振幅スペクトルが有する周波数帯域を前記第1の周波数帯域と異なる第2の周波数帯域まで拡張することにより、前記包絡振幅スペクトルを広帯域化する包絡振幅スペクトル広帯域化部と、

前記第1の周波数帯域における周期性振幅スペクトルの包絡線を算出し、前記第2の周波数帯域においても当該包絡線の傾きを維持し、かつ、前記第2の周波数帯域において周波数が高くなるほど前記周期性振幅スペクトルの周波数に対する周期性を弱くすることにより、前記周期性振幅スペクトルが有する周波数帯域を前記第2の周波数帯域まで拡張して前記周期性振幅スペクトルを広帯域化する周期性振幅スペクトル広帯域化部と、

前記第1の周波数帯域におけるランダム振幅スペクトルの包絡線を算出し、前記第2の周波数帯域においても当該包絡線の傾きを維持し、かつ、周波数が高くなるにつれて当該包絡線を大きくすることにより、前記ランダム振幅スペクトルが有する周波数帯域を前記第2の周波数帯域まで拡張して前記ランダム振幅スペクトルを広帯域化するランダム振幅スペクトル広帯域化部と、

前記広帯域化された包絡振幅スペクトルと、前記広帯域化された周期性振幅スペクトルと、前記広帯域化されたランダム振幅スペクトルを合成することにより、前記第1の周波数帯域及び前記第2の周波数帯域を有する広帯域周波数スペクトルを合成する広帯域スペクトル算出部と、

前記広帯域周波数スペクトルを周波数時間変換することにより、前記第1の周波数帯域及び前記第2の周波数帯域を有する音声信号を生成する周波数時間変換部と、
を有する音声帯域拡張装置。

【請求項2】

前記第1の周波数帯域に含まれる各周波数に対する前記周波数スペクトルの位相を表す位相スペクトルが有する周波数帯域を前記第2の周波数帯域にまで拡張することにより、
前記位相スペクトルを広帯域化する位相スペクトル広帯域化部をさらに有し、

10

前記広帯域スペクトル算出部は、前記広帯域化された包絡振幅スペクトルと、前記広帯域化された周期性振幅スペクトルと、前記広帯域化されたランダム振幅スペクトルと、前記広帯域化された位相スペクトルを合成することにより、前記広帯域周波数スペクトルを合成する、請求項1に記載の音声帯域拡張装置。

【請求項3】

第1の周波数帯域を有する音声信号を、所定の時間長を有するフレームごとに時間周波数変換することにより、該音声信号の周波数スペクトルを算出し、

前記周波数スペクトルから、該周波数スペクトルの包絡振幅スペクトルと、周波数に対して周期的にスペクトル強度が変動する周期性振幅スペクトルと、周波数に対してランダムにスペクトル強度が変動するランダム振幅スペクトルとを抽出し、

20

前記包絡振幅スペクトルが有する周波数帯域を前記第1の周波数帯域と異なる第2の周波数帯域まで拡張することにより、前記包絡振幅スペクトルを広帯域化し、

前記第1の周波数帯域における周期性振幅スペクトルの包絡線を算出し、前記第2の周波数帯域においても当該包絡線の傾きを維持し、かつ、前記第2の周波数帯域において周波数が高くなるほど前記周期性振幅スペクトルの周波数に対する周期性を弱くすることにより、前記周期性振幅スペクトルが有する周波数帯域を前記第2の周波数帯域まで拡張して前記周期性振幅スペクトルを広帯域化し、

前記第1の周波数帯域におけるランダム振幅スペクトルの包絡線を算出し、前記第2の周波数帯域においても当該包絡線の傾きを維持し、かつ、周波数が高くなるにつれて当該包絡線を大きくすることにより、前記ランダム振幅スペクトルが有する周波数帯域を前記第2の周波数帯域まで拡張して前記ランダム振幅スペクトルを広帯域化し、

30

前記広帯域化された包絡振幅スペクトルと、前記広帯域化された周期性振幅スペクトルと、前記広帯域化されたランダム振幅スペクトルを合成することにより、前記第1の周波数帯域及び前記第2の周波数帯域を有する広帯域周波数スペクトルを合成し、

前記広帯域周波数スペクトルを周波数時間変換することにより、前記第1の周波数帯域及び前記第2の周波数帯域を有する音声信号を生成する、
ことを含む音声帯域拡張方法。

【請求項4】

第1の周波数帯域を有する音声信号を、所定の時間長を有するフレームごとに時間周波数変換することにより、該音声信号の周波数スペクトルを算出し、

40

前記周波数スペクトルから、該周波数スペクトルの包絡振幅スペクトルと、周波数に対して周期的にスペクトル強度が変動する周期性振幅スペクトルと、周波数に対してランダムにスペクトル強度が変動するランダム振幅スペクトルとを抽出し、

前記包絡振幅スペクトルが有する周波数帯域を前記第1の周波数帯域と異なる第2の周波数帯域まで拡張することにより、前記包絡振幅スペクトルを広帯域化し、

前記第1の周波数帯域における周期性振幅スペクトルの包絡線を算出し、前記第2の周波数帯域においても当該包絡線の傾きを維持し、かつ、前記第2の周波数帯域において周波数が高くなるほど前記周期性振幅スペクトルの周波数に対する周期性を弱くすることにより、前記周期性振幅スペクトルが有する周波数帯域を前記第2の周波数帯域まで拡張し

50

て前記周期性振幅スペクトルを広帯域化し、

前記第 1 の周波数帯域におけるランダム振幅スペクトルの包絡線を算出し、前記第 2 の周波数帯域においても当該包絡線の傾きを維持し、かつ、周波数が高くなるにつれて当該包絡線を大きくすることにより、前記ランダム振幅スペクトルが有する周波数帯域を前記第 2 の周波数帯域まで拡張して前記ランダム振幅スペクトルを広帯域化し、

前記広帯域化された包絡振幅スペクトルと、前記広帯域化された周期性振幅スペクトルと、前記広帯域化されたランダム振幅スペクトルを合成することにより、前記第 1 の周波数帯域及び前記第 2 の周波数帯域を有する広帯域周波数スペクトルを合成し、

前記広帯域周波数スペクトルを周波数時間変換することにより、前記第 1 の周波数帯域及び前記第 2 の周波数帯域を有する音声信号を生成する、
ことをコンピュータに実行させる音声帯域拡張用コンピュータプログラム。

10

【請求項 5】

第 1 の周波数帯域を有する符号化された音声信号を受信する通信部と、

前記音声信号を復号するベースバンド処理部と、

前記音声信号を広帯域化する制御部であって、

前記音声信号を、所定の時間長を有するフレームごとに時間周波数変換することにより、該音声信号の周波数スペクトルを算出し、

前記周波数スペクトルから、該周波数スペクトルの包絡振幅スペクトルと、周波数に対して周期的にスペクトル強度が変動する周期性振幅スペクトルと、周波数に対してランダムにスペクトル強度が変動するランダム振幅スペクトルとを抽出し、

20

前記包絡振幅スペクトルが有する周波数帯域を前記第 1 の周波数帯域と異なる第 2 の周波数帯域まで拡張することにより、前記包絡振幅スペクトルを広帯域化し、

前記第 1 の周波数帯域における周期性振幅スペクトルの包絡線を算出し、前記第 2 の周波数帯域においても当該包絡線の傾きを維持し、かつ、前記第 2 の周波数帯域において周波数が高くなるほど前記周期性振幅スペクトルの周波数に対する周期性を弱くすることにより、前記周期性振幅スペクトルが有する周波数帯域を前記第 2 の周波数帯域まで拡張して前記周期性振幅スペクトルを広帯域化し、

前記第 1 の周波数帯域におけるランダム振幅スペクトルの包絡線を算出し、前記第 2 の周波数帯域においても当該包絡線の傾きを維持し、かつ、周波数が高くなるにつれて当該包絡線を大きくすることにより、前記ランダム振幅スペクトルが有する周波数帯域を前記第 2 の周波数帯域まで拡張して前記ランダム振幅スペクトルを広帯域化し、

30

前記広帯域化された包絡振幅スペクトルと、前記広帯域化された周期性振幅スペクトルと、前記広帯域化されたランダム振幅スペクトルを合成することにより、前記第 1 の周波数帯域及び前記第 2 の周波数帯域を有する広帯域周波数スペクトルを合成し、

前記広帯域周波数スペクトルを周波数時間変換することにより、前記第 1 の周波数帯域及び前記第 2 の周波数帯域を有する広帯域音声信号を生成する、
制御部と、

前記広帯域化音声信号を再生するスピーカと、
を有する電話機。

【発明の詳細な説明】

40

【技術分野】

【0001】

ここに開示される実施形態は、音声信号が有する周波数帯域を拡張する音声帯域拡張装置、音声帯域拡張方法及び音声帯域拡張用コンピュータプログラムならびに電話機に関する。

【背景技術】

【0002】

音声伝送システムでは、限られた伝送周波数帯域で音声信号を伝送するために、一般に、音声信号は狭帯域化され、その狭帯域化された音声信号が伝送される。そのため、音声信号を受信した受信機が再生する音声が含まれる周波数帯域も、元の音声が含まれる周波

50

数帯域よりも狭くなるので、受信機により再生された音声の品質が劣化する。そこで、音声信号が含まれる周波数帯域を擬似的に拡張することで、再生音声の品質を向上させる技術が開発されている（例えば、特許文献1を参照）。

【0003】

例えば、特許文献1に開示された技術は、入力信号からスペクトル包絡情報と残差信号を抽出する。そしてその技術は、スペクトル包絡情報の周波数帯域と残差信号の周波数帯域をそれぞれ拡大し、周波数帯域が拡大されたスペクトル包絡情報と残差信号とを用いて音声を合成する。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開平8-248997号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

音声信号には、人の声のように、音声信号の周波数スペクトルの振幅値が周波数の変化に応じて周期的に変化する周期性振幅スペクトルと、周波数スペクトルの振幅値が周波数の変化によらずランダムに変化するランダム振幅スペクトルが含まれることがある。

しかし、従来の技術は、入力された音声信号から、周期性振幅スペクトルとランダム振幅スペクトルとを分離せず、スペクトルの包絡情報と残差信号を広帯域化する。また従来の技術は、周波数ごとの位相を表す位相スペクトルを考慮していない。そのため、従来の技術は、周期性振幅スペクトル、ランダム振幅スペクトル及び位相スペクトルを、それぞれの特徴に応じて広帯域化することができない。

【0006】

自然な音質となるように音声信号を広帯域化するためには、広帯域化された周期性振幅スペクトル及びランダム振幅スペクトルも、元の音声信号に対応する周期性振幅スペクトル及びランダム振幅スペクトルと同じ特徴を持つことが望ましい。例えば、周期性振幅スペクトルの包絡線の周波数に対する傾きと、ランダム振幅スペクトルの包絡線の周波数に対する傾きが異なることがある。このような場合、従来の技術は、それぞれの振幅スペクトルの包絡線の傾きを保ちつつ、音声信号を広帯域化することができない。そのため、広帯域化された周期性振幅スペクトル及びランダム振幅スペクトルの特徴が、元の音声信号に対応する周期性振幅スペクトル及びランダム振幅スペクトルの特徴と異なってしまふ。したがって、広帯域化された音声信号の品質が劣化する。

【0007】

また、周期性振幅スペクトルについて、一般に、周波数が高くなるほど周期性が弱くなることが知られている。しかし、従来の技術は、周期性振幅スペクトルのみを分離して広帯域化できないので、このような周期性振幅スペクトルが有する性質を再現できない。そのため、再現された音声は、自然な音声にならないことがある。

さらに、従来の技術は、入力された音声信号を解析する単位であるフレーム間の位相の連続性を考慮していないので、音声の周波数と対応する角速度によって定まる位相がフレーム間で不連続となるおそれがある。そして、フレーム間で位相が不連続になると、再生された音声信号も不連続となるので、再生された音声信号の品質が劣化してしまう。

【0008】

そこで、本明細書は、再生される音声の品質を向上可能な音声帯域拡張装置、音声帯域拡張方法及び音声帯域拡張用コンピュータプログラムならびに電話機を提供することを目的とする。

【課題を解決するための手段】

【0009】

一つの実施形態によれば、音声帯域拡張装置が提供される。この音声帯域拡張装置は、第1の周波数帯域を有する音声信号を、所定の時間長を有するフレームごとに時間周波数

10

20

30

40

50

変換することにより、音声信号の周波数スペクトルを算出する時間周波数変換部と、周波数スペクトルから、周波数スペクトルの包絡振幅スペクトルと、周波数に対して周期的にスペクトル強度が変動する周期性振幅スペクトルと、周波数に対してランダムにスペクトル強度が変動するランダム振幅スペクトルとを抽出する分離部と、包絡振幅スペクトルが有する周波数帯域を第1の周波数帯域と異なる第2の周波数帯域まで拡張することにより、包絡振幅スペクトルを広帯域化する包絡振幅スペクトル広帯域化部と、周期性振幅スペクトルが有する周波数帯域を第2の周波数帯域まで拡張することにより、周期性振幅スペクトルを広帯域化する周期性振幅スペクトル広帯域化部と、ランダム振幅スペクトルが有する周波数帯域を第2の周波数帯域まで拡張することにより、ランダム振幅スペクトルを広帯域化するランダム振幅スペクトル広帯域化部と、広帯域化された包絡振幅スペクトルと、広帯域化された周期性振幅スペクトルと、広帯域化されたランダム振幅スペクトルを合成することにより、第1の周波数帯域及び第2の周波数帯域を有する広帯域周波数スペクトルを合成する広帯域スペクトル算出部と、広帯域周波数スペクトルを周波数時間変換することにより、第1の周波数帯域及び第2の周波数帯域を有する音声信号を生成する周波数時間変換部とを有する。

10

【0010】

また他の実施形態によれば、音声帯域拡張方法が提供される。この音声帯域拡張方法は、第1の周波数帯域を有する音声信号を、所定の時間長を有するフレームごとに時間周波数変換することにより、音声信号の周波数スペクトルを算出し、周波数スペクトルから、周波数スペクトルの包絡振幅スペクトルと、周波数に対して周期的にスペクトル強度が変動する周期性振幅スペクトルと、周波数に対してランダムにスペクトル強度が変動するランダム振幅スペクトルとを抽出し、包絡振幅スペクトルが有する周波数帯域を第1の周波数帯域と異なる第2の周波数帯域まで拡張することにより、包絡振幅スペクトルを広帯域化し、周期性振幅スペクトルが有する周波数帯域を第2の周波数帯域まで拡張することにより、周期性振幅スペクトルを広帯域化し、ランダム振幅スペクトルが有する周波数帯域を第2の周波数帯域まで拡張することにより、ランダム振幅スペクトルを広帯域化し、広帯域化された包絡振幅スペクトルと、広帯域化された周期性振幅スペクトルと、広帯域化されたランダム振幅スペクトルを合成することにより、第1の周波数帯域及び第2の周波数帯域を有する広帯域周波数スペクトルを合成し、広帯域周波数スペクトルを周波数時間変換することにより、第1の周波数帯域及び第2の周波数帯域を有する音声信号を生成することを含む。

20

30

【0011】

さらに他の実施形態によれば、コンピュータに第1の周波数帯域を有する音声信号の周波数帯域を拡張させる音声帯域拡張用コンピュータプログラムが提供される。このコンピュータプログラムは、第1の周波数帯域を有する音声信号を、所定の時間長を有するフレームごとに時間周波数変換することにより、音声信号の周波数スペクトルを算出し、周波数スペクトルから、周波数スペクトルの包絡振幅スペクトルと、周波数に対して周期的にスペクトル強度が変動する周期性振幅スペクトルと、周波数に対してランダムにスペクトル強度が変動するランダム振幅スペクトルとを抽出し、包絡振幅スペクトルが有する周波数帯域を第1の周波数帯域と異なる第2の周波数帯域まで拡張することにより、包絡振幅スペクトルを広帯域化し、周期性振幅スペクトルが有する周波数帯域を第2の周波数帯域まで拡張することにより、周期性振幅スペクトルを広帯域化し、ランダム振幅スペクトルが有する周波数帯域を第2の周波数帯域まで拡張することにより、ランダム振幅スペクトルを広帯域化し、広帯域化された包絡振幅スペクトルと、広帯域化された周期性振幅スペクトルと、広帯域化されたランダム振幅スペクトルを合成することにより、第1の周波数帯域及び第2の周波数帯域を有する広帯域周波数スペクトルを合成し、広帯域周波数スペクトルを周波数時間変換することにより、第1の周波数帯域及び第2の周波数帯域を有する音声信号を生成する、ことをコンピュータに実行させる命令を有する。

40

【0012】

さらに他の実施形態によれば、電話機が提供される。この電話機は、第1の周波数帯域

50

を有する符号化された音声信号を受信する通信部と、音声信号を復号するベースバンド処理部と、音声信号を広帯域化する制御部と、広帯域化された音声信号を再生するスピーカとを有する。そして制御部は、音声信号を、所定の時間長を有するフレームごとに時間周波数変換することにより、音声信号の周波数スペクトルを算出し、周波数スペクトルから、周波数スペクトルの包絡振幅スペクトルと、周波数に対して周期的にスペクトル強度が変動する周期性振幅スペクトルと、周波数に対してランダムにスペクトル強度が変動するランダム振幅スペクトルとを抽出し、包絡振幅スペクトルが有する周波数帯域を第1の周波数帯域と異なる第2の周波数帯域まで拡張することにより、包絡振幅スペクトルを広帯域化し、周期性振幅スペクトルが有する周波数帯域を第2の周波数帯域まで拡張することにより、周期性振幅スペクトルを広帯域化し、ランダム振幅スペクトルが有する周波数帯域を第2の周波数帯域まで拡張することにより、ランダム振幅スペクトルを広帯域化し、広帯域化された包絡振幅スペクトルと、広帯域化された周期性振幅スペクトルと、広帯域化されたランダム振幅スペクトルを合成することにより、第1の周波数帯域及び第2の周波数帯域を有する広帯域周波数スペクトルを合成し、広帯域周波数スペクトルを周波数時間変換することにより、第1の周波数帯域及び第2の周波数帯域を有する広帯域音声信号を生成する。

10

【0013】

本発明の目的及び利点は、請求項において特に指摘されたエレメント及び組み合わせにより実現され、かつ達成される。

上記の一般的な記述及び下記の詳細な記述の何れも、例示的かつ説明的なものであり、請求項のように、本発明を限定するものではないことを理解されたい。

20

【発明の効果】

【0014】

本明細書に開示された音声帯域拡張装置、音声帯域拡張方法及び音声帯域拡張用コンピュータプログラムならびに電話機は、再生される音声の品質を向上することが可能である。

【図面の簡単な説明】

【0015】

【図1】一つの実施形態に係る音声帯域拡張装置の概略構成図である。

【図2】(a)は周波数スペクトルに含まれる包絡振幅スペクトルの一例を示す図であり、(b)は周期性振幅スペクトルの一例を示す図であり、(c)はランダム振幅スペクトルの一例を示す図である。

30

【図3】周波数スペクトル分離処理の動作フローチャートである。

【図4】高周波数帯域包絡振幅スペクトル生成処理の動作フローチャートである。

【図5】高周波数帯域周期性振幅スペクトル生成処理の動作フローチャートである。

【図6】高周波数帯域ランダム振幅スペクトル生成処理の動作フローチャートである。

【図7】一つの実施形態に係る音声帯域拡張装置が実行する音声帯域拡張処理の動作フローチャートである。

【図8】音声帯域拡張装置が組み込まれた電話機の概略構成図である。

【発明を実施するための形態】

40

【0016】

以下、図を参照しつつ、一つの実施形態による、音声帯域拡張装置について説明する。

この音声帯域拡張装置は、入力された音声信号を、包絡振幅スペクトルと、周期性振幅スペクトルと、ランダム振幅スペクトルと、位相スペクトルとに分離する。そしてこの音声帯域拡張装置は、各スペクトルの特徴に応じて各スペクトルの周波数帯域を高周波側へ広げることにより、再生される音声の品質を向上する。

なお、本実施形態では、一例として、音声帯域拡張装置に入力された音声信号は、300Hz～4kHzの周波数帯域に含まれる。そして音声帯域拡張装置は、擬似的に4kHz～8kHzの周波数帯域に含まれる音声信号成分を生成することにより、音声信号を広帯域化する。ただし、入力音声信号の周波数帯域は、300Hz～4kHzに限られない。入力音声信号の周波数帯

50

域は、300Hz～3.4kHzであってもよい。また、音声帯域拡張装置が擬似的に生成する音声信号成分の周波数帯域も、4kHz～8kHzに限られない。例えば、音声帯域拡張装置は、4kHz～16kHzの周波数帯域に含まれる音声信号成分を生成してもよい。また音声帯域拡張装置は、入力音声信号の周波数帯域の下限よりも低周波数の可聴帯域、例えば、50Hz～300Hzの周波数帯域に含まれる音声信号成分を生成してもよい。

【0017】

図1は、一つの実施形態による音声帯域拡張装置1の概略構成図である。音声帯域拡張装置1は、バッファメモリ10と、時間周波数変換部11と、分離部12と、包絡振幅スペクトル広帯域化部13と、周期性振幅スペクトル広帯域化部14と、ランダム振幅スペクトル広帯域化部15と、位相スペクトル広帯域化部16と、スペクトル合成部17と、周波数時間変換部18とを有する。

10

【0018】

音声帯域拡張装置1が有するこれらの各部は、それぞれ別個の回路として形成される。あるいは音声帯域拡張装置1が有するこれらの各部は、その各部に対応する回路が集積された一つの集積回路として音声帯域拡張装置1に実装されてもよい。さらに、音声帯域拡張装置1が有するこれらの各部は、音声帯域拡張装置1が有するプロセッサ上で実行されるコンピュータプログラムにより実現される、機能モジュールであってもよい。

【0019】

バッファメモリ10は、入力音声信号を一時的に記憶する。そしてバッファメモリ10に記憶された入力音声信号は、入力された時間順に沿って所定のフレーム単位で、時間周波数変換部11により読み出される。

20

【0020】

時間周波数変換部11は、バッファメモリ10から読み出した入力音声信号を、所定のフレーム単位で時間周波数変換することにより、入力音声信号の周波数スペクトルを算出する。なお、時間周波数変換部11により実行される時間周波数変換は、例えば、高速フーリエ変換あるいは離散コサイン変換とすることができる。またフレーム長は、例えば、10msec～80msecの何れかの期間とすることができる。

時間周波数変換部11は、フレーム単位で周波数スペクトルを算出する度に、その算出された周波数スペクトルを分離部12及びスペクトル合成部17へ出力する。

【0021】

30

ここで、周波数スペクトルは、包絡振幅スペクトル、周期性振幅スペクトル、ランダム振幅スペクトル及び位相スペクトルが合成されたものとして表すことができる。このうち、振幅に関する、包絡振幅スペクトル、周期性振幅スペクトル及びランダム振幅スペクトルは、それぞれ、周波数の変化に対して異なる特徴を持つことがある。

図2(a)は周波数スペクトルに含まれる包絡振幅スペクトルの一例を示す図であり、図2(b)は周期性振幅スペクトルの一例を示す図であり、図2(c)はランダム振幅スペクトルの一例を示す図である。

図2(a)～図2(c)において、横軸は周波数を表し、縦軸はスペクトルの強度を表す。また、周波数 f_{nbu} は、入力音声信号が有する周波数帯域の上限値を表す。図2(a)に示されるように、包絡振幅スペクトル200は、例えば、特定の周波数で最大となり、その特定の周波数よりも周波数が高くなるにつれてなだらかに減少するスペクトル形状を有する。

40

また、図2(b)に示されるように、周期性振幅スペクトル210は、周期的に変動する。またこの例では、周期性振幅スペクトル210の包絡線211は、周波数が高くなるにつれて強度が減少する関数となる。

一方、図2(c)に示されるように、ランダム振幅スペクトル220は、例えば、周波数が高くなるにつれて全体的に増加する。そのため、ランダム振幅スペクトル220の包絡線221は、周波数が高くなるにつれて強度が増加するような関数となる。

【0022】

このように、包絡振幅スペクトル、周期性振幅スペクトル及びランダム振幅スペクトル

50

は、それぞれ、周波数の変化に対して異なる特徴を持つ。また、再生された音声信号が自然な音となるために、周波数 f_{nbu} よりも高い周波数帯域において擬似的に生成される各振幅スペクトルも、各振幅スペクトルが周波数 f_{nbu} 以下の周波数の変化に対して持つ特徴と同じ特徴を持つことが好ましい。

例えば、周波数 f_{nbu} よりも高い高周波数帯域において、擬似的に生成される周期性振幅スペクトル 2 1 2 の極大値も、周波数が高くなるにつれて、包絡線 2 1 1 に沿って減少することが好ましい。また、周波数 f_{nbu} よりも高い高周波数帯域において、擬似的に生成されるランダム振幅スペクトル 2 2 2 の極大値も、周波数が高くなるにつれて、包絡線 2 2 1 に沿って増加することが好ましい。

【 0 0 2 3 】

10

そこで分離部 1 2 は、時間周波数変換部 1 1 から周波数スペクトルを受け取る度に、その周波数スペクトルから、包絡振幅スペクトルと、周期性振幅スペクトルと、ランダム振幅スペクトルとを抽出する。さらに分離部 1 2 は、時間周波数変換部 1 1 から周波数スペクトルを受け取る度に、位相スペクトルを抽出する。

【 0 0 2 4 】

図 3 は、分離部 1 2 により実行される周波数スペクトル分離処理の動作フローチャートである。分離部 1 2 は、周波数スペクトルから、次式に従って位相スペクトルを算出する(ステップ S 1 0 1)。

【数 1】

20

$$ps[f] = \tan^{-1} \frac{im[f]}{re[f]} \quad (1)$$

(1) 式において、 f は周波数である。また、 $ps[f]$ は、周波数 f に対する位相を表す位相スペクトルである。さらに $re[f]$ は、周波数 f に対する周波数スペクトルの実部の成分を表す。さらに、 $im[f]$ は、周波数 f に対する周波数スペクトルの虚部の成分を表す。

【 0 0 2 5 】

30

また分離部 1 2 は、周波数スペクトルから、次式に従って対数パワースペクトルを算出する(ステップ S 1 0 2)。

【数 2】

$$lps[f] = 10 \log_{10} (re[f]^2 + im[f]^2) \quad (2)$$

(2) 式において、 f は周波数である。また、 $lps[f]$ は、周波数 f の関数として表される対数パワースペクトルである。さらに $re[f]$ は、周波数 f に対する周波数スペクトルの実部の成分を表す。さらに、 $im[f]$ は、周波数 f に対する周波数スペクトルの虚部の成分を表す。

40

対数パワースペクトルが算出された後、分離部 1 2 は、対数パワースペクトルを時間周波数変換することにより、ケプストラムを算出する(ステップ S 1 0 3)。なお、時間周波数変換として、例えば、高速フーリエ変換あるいは離散コサイン変換が用いられる。そして分離部 1 2 は、ケプストラムが最大値となるケフレンシー Q_{max} を求める(ステップ S 1 0 4)。なお、 Q_{max} は、周期性振幅スペクトルのピッチ周波数に相当する。

【 0 0 2 6 】

次に、分離部 1 2 は、周波数スペクトルから包絡振幅スペクトル、周期性振幅スペクトル及びランダム振幅スペクトルを抽出するために、周期性振幅スペクトルに対応するケフ

50

レンシーの上限値及び下限値を次式に従って決定する（ステップ S 1 0 5）。

【数 3】

$$TH_L=Q_{max}*COEF_L \quad (3)$$

$$TH_H=Q_{max}*COEF_H \quad (4)$$

10

ここで、THLは、周期性振幅スペクトルに対応するケフレンシーの下限値である。またTHHは、周期性振幅スペクトルに対応するケフレンシーの上限値である。さらに、COEFLは、周期性振幅スペクトルに対応するケフレンシーの下限値THLを算出するための係数である。係数COEFLは、以下の条件を満たす任意の数に設定される。

$$0 < COEFL < 1 \quad (5)$$

COEFHは、周期性振幅スペクトルに対応するケフレンシーの上限値THHを算出するための係数である。係数COEFHは、例えば、以下の条件を満たす任意の数に設定される。

$$1 < COEFH < 3 \quad (6)$$

【0 0 2 7】

周期性振幅スペクトルに対応するケフレンシーの上限値及び下限値が決定されると、分離部 1 2 は、ケプストラムから包絡振幅スペクトルを抽出する（ステップ S 1 0 6）。その際、分離部 1 2 は、下限値THL以上のケフレンシーに対するケプストラムの成分を 0 に置換する。そして分離部 1 2 は、置換後のケプストラムを周波数時間変換することにより、包絡振幅スペクトルを算出する。

20

また、分離部 1 2 は、ケプストラムから周期性振幅スペクトルを抽出する（ステップ S 1 0 7）。その際、分離部 1 2 は、下限値THL未満のケフレンシーに対するケプストラムの成分と、上限値THH以上のケフレンシーに対するケプストラムの成分を 0 に置換する。そして分離部 1 2 は、置換後のケプストラムを周波数時間変換することにより、周期性振幅スペクトルを算出する。なお、THLとTHHの差が小さくなるほど、周期性振幅スペクトルのピッチ周波数に相当するスペクトルのみが算出される。

30

さらに、分離部 1 2 は、ケプストラムからランダム振幅スペクトルを抽出する（ステップ S 1 0 8）。その際、分離部 1 2 は、上限値THH未満のケフレンシーに対するケプストラムの成分を 0 に置換する。そして分離部 1 2 は、置換後のケプストラムを周波数時間変換することにより、ランダム振幅スペクトルを算出する。

なお、ステップ S 1 0 6 ~ S 1 0 8 において実行される周波数時間変換は、ステップ S 1 0 3 において実行される時間周波数変換の逆変換である。

また、分離部 1 2 は、ステップ S 1 0 1 の処理を、ステップ S 1 0 2 ~ S 1 0 8 の処理と並列に実行してもよい。あるいは分離部 1 2 は、ステップ S 1 0 1 の処理とステップ S 1 0 2 ~ S 1 0 8 の処理を実行する順序を入れ替えてもよい。さらに分離部 1 2 は、ステップ S 1 0 6 ~ S 1 0 8 の処理を実行する順序を入れ替えてもよい。

40

【0 0 2 8】

分離部 1 2 は、包絡振幅スペクトルを包絡振幅スペクトル広帯域化部 1 3 へ渡す。また分離部 1 2 は、元の周波数スペクトル及び周期性振幅スペクトルと、ケプストラムの最大値及びその最大値に対応するケフレンシー Q_{max} を周期性振幅スペクトル広帯域化部 1 4 へ渡す。さらに分離部 1 2 は、ランダム振幅スペクトルをランダム振幅スペクトル広帯域化部 1 5 へ渡す。そして分離部 1 2 は、元の周波数スペクトル及び位相スペクトルを位相スペクトル広帯域化部 1 6 へ渡す。

【0 0 2 9】

包絡振幅スペクトル広帯域化部 1 3 は、分離部 1 2 から受け取った包絡振幅スペクトルの周波数帯域を広げる。そのために、包絡振幅スペクトル広帯域化部 1 3 は、分離部 1 2

50

から受け取った包絡振幅スペクトルに基づいて、入力音声信号の周波数帯域の上限よりも高い高周波数帯域の包絡振幅スペクトルを生成する。なお、高周波数帯域は、一例として、4kHz～8kHzである。

【 0 0 3 0 】

図 4 は、包絡振幅スペクトル広帯域化部 1 3 により実行される高周波数帯域包絡振幅スペクトル生成処理の動作フローチャートである。

包絡振幅スペクトル広帯域化部 1 3 は、分離部 1 2 から受け取った包絡振幅スペクトルを周波数方向に平滑化する（ステップ S 2 0 1）。例えば、包絡振幅スペクトル広帯域化部 1 3 は、次式に従って包絡振幅スペクトルを平滑化する。

【 数 4 】

$$PEsm(f) = \frac{1}{2w+1} \sum_{i=-w}^{i=w} PE(f+i) \quad (7)$$

ここで関数 $PE(f)$ は、周波数 f に対する包絡振幅スペクトルである。また関数 $PEsm(f)$ は、周波数 f に対して平滑化された包絡振幅スペクトルである。 w は、平滑化される周波数帯域の幅を表し、例えば、 w は 100Hz に設定される。

【 0 0 3 1 】

次に、包絡振幅スペクトル広帯域化部 1 3 は、平滑化された包絡振幅スペクトルに基づいて、高周波数帯域における包絡振幅スペクトルの振幅を決定する（ステップ S 2 0 2）。例えば、包絡振幅スペクトル広帯域化部 1 3 は、次式に従って高周波数帯域における包絡振幅スペクトルの振幅を決定する。

【 数 5 】

$$PE(f) = rate * PEsm(f - f_L) \quad (f \geq f_L + \Delta w) \quad (8)$$

ここで係数 $rate$ は、予め、音声帯域拡張装置 1 が出力する音声の有する周波数帯域と等しい周波数帯域を持つ、さまざまな話者または発声内容が含まれる音声を用いて求められた、低周波数帯域の音声に対する高周波数帯域の音声の平均的なパワー比を表す。この低周波数帯域は、入力音声信号が有する周波数帯域である。一方、高周波数帯域は、包絡振幅スペクトル広帯域化部 1 3 により生成される包絡振幅スペクトルが有する周波数帯域である。また f_L は、高周波数帯域の下限値を表す。本実施形態では、 f_L は 4kHz である。さらに w は、高周波数帯域と低周波数帯域の包絡線を滑らかに接続するための帯域幅に相当する。例えば、 w は、100Hz に設定される。

【 0 0 3 2 】

包絡振幅スペクトル広帯域化部 1 3 は、低周波数帯域における包絡振幅スペクトルと高周波数帯域における包絡振幅スペクトルが滑らかに接続されるように、高周波数帯域のうち、低周波数帯域に近接した帯域の包絡振幅スペクトルを補間する（ステップ S 2 0 3）。例えば、包絡振幅スペクトル広帯域化部 1 3 は、次式に従って高周波数帯域のうち、低周波数帯域に近接した帯域における包絡振幅スペクトルを決定する。

10

20

30

40

【数 6】

$$PE(f) = (1 - coef) * PEsm(f_L) + coef * rate * PEsm(f - f_L) \quad f_L < f < f_L + \Delta w \quad (9)$$

$$coef = \frac{f - f_L}{\Delta w}$$

【0033】

10

なお、包絡振幅スペクトル広帯域化部13は、他の方法によって高周波数帯域における包絡振幅スペクトルを生成してもよい。例えば、包絡振幅スペクトル広帯域化部13は、入力音声信号が有する周波数帯域の上限値における、包絡振幅スペクトルの強度を、高周波数帯域に含まれる各周波数に対する包絡振幅スペクトルの強度としてもよい。あるいは、包絡振幅スペクトル広帯域化部13は、入力音声信号が有する周波数帯域の上限値近傍における、包絡振幅スペクトルの接線、あるいは包絡振幅スペクトルを近似する3次スプライン関数を、高周波数帯域における包絡振幅スペクトルとして求めてもよい。

包絡振幅スペクトル広帯域化部13は、高周波数帯域における包絡振幅スペクトルをスペクトル合成部17へ出力する。

【0034】

20

周期性振幅スペクトル広帯域化部14は、分離部12から受け取った周期性振幅スペクトルの周波数帯域を広げる。そのために、周期性振幅スペクトル広帯域化部14は、分離部12から受け取った周期性振幅スペクトルに基づいて、入力音声信号の周波数帯域の上限よりも高い高周波数帯域における周期性振幅スペクトルを生成する。なお、高周波数帯域は、一例として、4kHz～8kHzである。

【0035】

図5は、周期性振幅スペクトル広帯域化部14により実行される高周波数帯域周期性振幅スペクトル生成処理の動作フローチャートである。

周期性振幅スペクトル広帯域化部14は、分離部12から受け取った周期性振幅スペクトルの包絡線を算出する(ステップS301)。包絡線を算出するために、周期性振幅スペクトル広帯域化部14は、周期性振幅スペクトルの極大点を求める。極大点は、例えば、周波数 f_j におけるスペクトルの強度を I_j ($j=1, 2, \dots, n$ 、ただし n は一つのフレームに含まれるスペクトル点数)としたとき、以下の条件を満たす点である。

$$I_{j-1} < I_j \quad \text{かつ} \quad I_{j+1} < I_j$$

周期性振幅スペクトル広帯域化部14は、極大点(f_j, I_j)の集合に対して、例えば、最小二乗法を適用することにより、近似的に各極大点(f_j, I_j)を結ぶ直線 $I = af + b$ を包絡線として算出する。

あるいは、周期性振幅スペクトル広帯域化部14は、各極大点(f_j, I_j)を結ぶ3次スプライン関数を求め、最も周波数が高い極大点における、その3次スプライン関数を包絡線を表す関数として算出してもよい。

40

さらに、周期性振幅スペクトル広帯域化部14は、周期性振幅スペクトルの極大点の代わりに、下記の条件を満たす極小点を求めてもよい。

$$I_{j-1} > I_j \quad \text{かつ} \quad I_{j+1} > I_j$$

そして周期性振幅スペクトル広帯域化部14は、極小点(f_j, I_j)の集合に対して、上記のように、最小二乗法または3次スプライン関数を適用することにより、包絡線を算出してもよい。

【0036】

また、周期性振幅スペクトル広帯域化部14は、周期性振幅スペクトルの初期位相を次式に従って算出する(ステップS302)。

【数7】

$$\theta_0 = \tan^{-1} \frac{im_p}{re_p} \quad (10)$$

ここで θ_0 は、周期性振幅スペクトルの初期位相である。また re_p は、周期性振幅スペクトルに相当する、すなわち、閾値THL以上かつ閾値THH未満のケフレンシーにおけるケプストラムの最大値の実部であり、 im_p は、周期性振幅スペクトルに相当するケプストラムの最大値の虚部である。

10

【0037】

次に、周期性振幅スペクトル広帯域化部14は、高周波数帯域における周期性振幅スペクトルを、入力音声信号が有する周波数帯域における周期性振幅スペクトルの包絡線の傾きを維持するように生成する(ステップS303)。その際、周期性振幅スペクトル広帯域化部14は、再生される音声が自然な音声となるように、周波数が高くなるほど、周期性振幅スペクトルの周期性を弱くすることが好ましい。周期性振幅スペクトル広帯域化部14は、例えば、次式に従って高周波数帯域における周期性振幅スペクトルを生成できる。

【数8】

20

$$PP(f) = (1 - c(f))s(f) \sin\left(\frac{2\pi f}{T} + \theta_{fL}\right) + c(f)r(f) \quad (11)$$

ここで $PP(f)$ は、周波数 f における周期性振幅スペクトルの強度を表す。また、 $c(f)$ は、周波数が高くなるほど大きくなる関数であり、 $c(f)$ の値は0~1の範囲に含まれる。例えば、関数 $c(f)$ として、以下の関数を用いることができる。

$$c(f) = (f - f_L) / (f_H - f_L)$$

30

ただし、 f_H 、 f_L は、それぞれ、高周波数帯域の上限値及び下限値である。また、関数 $c(f)$ は、非線形な関数でもよい。例えば、関数 $c(f)$ として、以下の関数を用いてもよい。

$$c(f) = 1 / (1 + e^{-((f - (f_L + f_H) / 2))})$$

係数 θ_{fL} は、高周波数帯域の下限値 f_L において関数 $c(f)$ が略0となり、かつ高周波数帯域の上限値 f_H において関数 $c(f)$ が略1となるように定められる。

【0038】

また(11)式において、関数 $s(f)$ は包絡線を表す。この関数 $s(f)$ は、ステップS301で算出された包絡線の関数である。さらに、 θ_{fL} は、周波数スペクトルの周波数 f_L における位相であり、次式により求められる。

$$\theta_{fL} = \theta_0 + f_L * 2\pi / f$$

40

さらに、 $r(f)$ はランダム関数であり、例えば、 $r(f)$ の値は0~1の範囲に含まれる。また T は、周期性振幅スペクトルの周期である。周期性振幅スペクトルの周期 T は、例えば、周波数のずれ量 f (ただし、 $f > 0$) を初期値から徐々に大きくなるように変化させたときに、周期性振幅スペクトルの自己相関関数 $ACF(f)$ が最初に極大値となる f の値である。また f の初期値は、例えば、経験的に周期 T よりも小さいと推定される任意の正数に設定される。例えば、自己相関関数 $ACF(f)$ は、次式により表される。

【数 9】

$$ACF(j) = \frac{\sum_{i=1}^N NP(i)NP(i+j)}{\sqrt{\sum_{i=1}^N NP(i)^2} \sqrt{\sum_{i=1}^N NP(i+j)^2}} \quad (12)$$

ただし、 $NP(i)$ ($i=1,2,\dots,N$) は、時間周波数変換部 11 により算出された周波数スペクトルを表すベクトルである。そのベクトルの各要素の値は、入力音声信号が有する周波数帯域を N 等分したサブ周波数帯域の振幅値を表す。また N は、その周波数スペクトルを表すベクトルの要素の数である。そして j は、周波数のずれ量 f に対応する。周波数のずれ量 f は、 j にサブ周波数帯域の幅を乗じることにより算出される。

【0039】

また、周期性振幅スペクトル広帯域化部 14 は、(11) 式の代わりに、次式に従って高周波数帯域における周期性振幅スペクトルを生成してもよい。

【数 10】

$$PP(f) = s(f) \sin\left(\frac{2\pi f}{T + c(f)dT(f)} + \theta_L\right) \quad (13)$$

ここで $PP(f)$ は、周波数 f における周期性振幅スペクトルの強度を表す。また、 $c(f)$ は、周波数が高くなるほど大きくなる関数である。関数 $s(f)$ は包絡線を表す。 θ_L は、周波数スペクトルの周波数 f_L における位相である。また T は、周期性振幅スペクトルの周期である。そして $dT(f)$ はランダム関数であり、 $dT(f)$ の絶対値は、例えば、周期性振幅スペクトルの周期 T の10%~20%の範囲に含まれる。

【0040】

(13) 式では、周期性振幅スペクトルの周期 T に対して、周波数が高くなるほどランダム関数の寄与を大きくすることにより、周波数が高いほど、周期性振幅スペクトルの周期性が弱くなる。また、他の方法として、周期性振幅スペクトル広帯域化部 14 は、関数 $s(f)$ にランダム関数を加算することにより、周波数が高いほど、周期性振幅スペクトルの周期性が弱くなるようにしてもよい。例えば、周期性振幅スペクトル広帯域化部 14 は、(13) 式において、関数 $s(f)$ の代わりに $(s(f)+c(f)dT(f))$ を用い、 \sin 関数内の周波数 f の係数を $(2/T)$ としてもよい。さらに、周期性振幅スペクトル広帯域化部 14 は、周波数が高くなるほど周期性振幅スペクトルの周期性が弱まる他の方法を用いることができる。また、例えば、周期性振幅スペクトルがランダム振幅スペクトルよりも小さい場合、周期性振幅スペクトル広帯域化部 14 は、高周波数帯域における周期性振幅スペクトルを、周波数によらず、周期 T が維持されるように生成してもよい。

【0041】

最後に、周期性振幅スペクトル広帯域化部 14 は、高周波数帯域における周期性振幅スペクトルをスペクトル合成部 17 へ出力する。

【0042】

ランダム振幅スペクトル広帯域化部 15 は、分離部 12 から受け取ったランダム振幅スペクトルの周波数帯域を広げる。そのために、ランダム振幅スペクトル広帯域化部 15 は、分離部 12 から受け取ったランダム振幅スペクトルに基づいて、入力音声信号の周波数帯域の上限よりも高い高周波数帯域におけるランダム振幅スペクトルを生成する。なお、

高周波数帯域は、周期性振幅スペクトル広帯域化部 1 4 によって生成される周期性振幅スペクトルの高周波数帯域と等しく、その高周波数帯域は、一例として、4kHz ~ 8kHzである。

【 0 0 4 3 】

図 6 は、ランダム振幅スペクトル広帯域化部 1 5 により実行される高周波数帯域ランダム振幅スペクトル生成処理の動作フローチャートである。

ランダム振幅スペクトル広帯域化部 1 5 は、ランダム振幅スペクトルの包絡線を算出する (ステップ S 4 0 1)。なお、包絡線の算出の具体的な方法は、例えば、周期性振幅スペクトル広帯域化部 1 4 が周期性振幅スペクトルの包絡線を算出する方法と同じとすることができる。具体的には、ランダム振幅スペクトル広帯域化部 1 5 は、ランダム振幅スペクトルの極大点または極小点を求め、それら極大点または極小点の集合に最小二乗法を適用することにより、包絡線を算出できる。

10

【 0 0 4 4 】

次に、ランダム振幅スペクトル広帯域化部 1 5 は、高周波数帯域におけるランダム振幅スペクトルを、入力音声信号が有する周波数帯域におけるランダム振幅スペクトルの包絡線の傾きを維持するように生成する (ステップ S 4 0 2)。ランダム振幅スペクトル広帯域化部 1 5 は、例えば、次式に従って高周波数帯域におけるランダム振幅スペクトルを生成できる。

【 数 1 1 】

20

$$PR(f) = sr(f)rr(f) \quad (14)$$

ここでPR(f)は、周波数fにおけるランダム振幅スペクトルの強度を表す。また関数sr(f)は、ステップS 4 0 1で算出された、ランダム振幅スペクトルの包絡線の関数である。さらに、関数rr(f)はランダム関数である。ランダム関数rr(f)は、再生される音声は自然な音声になるように、高周波数帯域におけるランダム振幅スペクトルの絶対値が包絡線sr(f)の値を超えないランダムな値となるように設定される。例えば、rr(f)の値は-1~1の範囲に含まれる。

30

【 0 0 4 5 】

ランダム振幅スペクトル広帯域化部 1 5 は、高周波数帯域におけるランダム振幅スペクトルをスペクトル合成部 1 7 へ出力する。

【 0 0 4 6 】

位相スペクトル広帯域化部 1 6 は、高周波数帯域における周波数スペクトルの位相を決定する。

例えば、位相スペクトル広帯域化部 1 6 は、高周波数帯域に含まれる周波数fに対する位相を、周波数fよりも所定周波数だけ低い周波数に対する位相と同じ値に設定する。所定周波数は、例えば、4kHzとすることができる。あるいは、位相スペクトル広帯域化部 1 6 は、高周波数帯域に含まれる周波数fに対する位相を、入力音声信号が有する周波数帯域に含まれる何れかの周波数に対する位相としてもよい。

40

【 0 0 4 7 】

ただし、位相スペクトル広帯域化部 1 6 は、時間的に連続するフレーム間で、各周波数に対する位相が連続するように、各周波数に対する位相を決定する。そこで、位相スペクトル広帯域化部 1 6 は、着目フレームの前のフレームについて決定された各周波数に対する位相、周波数及びフレーム長から、着目フレーム開始時の各周波数に対する位相を推定位相として算出する。そして位相スペクトル広帯域化部 1 6 は、推定位相と、着目フレームについて、上記のように決定された各周波数に対する位相との位相差を求める。位相スペクトル広帯域化部 1 6 は、その位相差が所定の範囲を超えていれば、位相差が所定の範

50

帯に含まれるように位相を補正する。

【0048】

例えば、位相スペクトル広帯域化部16は、次式に従って、フレームtにおける高周波数帯域に含まれる周波数fに対する位相 $\phi(f, t)$ を決定する。

【数12】

$$\phi(f, t) = \phi(f - 4000, t) \quad (15)$$

$$\Delta\phi(f, t) = \phi(f, t) - (\phi(f, t-1) + \frac{2\pi f}{\Delta t}) \quad (16) \quad 10$$

$$-\pi < \phi(f, t) \leq \pi$$

(15)式では、原則として、周波数fよりも4kHz低い周波数の位相が、周波数fの位相とされる。なお、周波数fよりも4kHz低い周波数が、入力音声信号に存在しない周波数帯域に含まれる場合、 $\phi(f, t)$ は任意の値、例えば0に設定される。

【0049】

また、位相スペクトル広帯域化部16は、(16)式に従って、(15)式により算出された周波数fの位相 $\phi(f, t)$ と、一つ前のフレーム(t-1)の位相 $\phi(f, t-1)$ 、周波数f及びフレーム長 Δt から算出される推定位相との位相差 $\Delta\phi(f, t)$ を算出する。そして位相スペクトル広帯域化部16は、位相差 $\Delta\phi(f, t)$ が(-)よりも大きい場合、位相 $\phi(f, t)$ に、オフセット値である $\pi/2$ を減算する。一方、位相スペクトル広帯域化部16は、位相差 $\Delta\phi(f, t)$ が(- +)よりも小さい場合、位相 $\phi(f, t)$ に、オフセット値である $\pi/2$ を加える。なお、 $\pi/2$ は、許容可能な位相差の最大値に相当する値であり、例えば、ユーザが位相のずれによる再生音の不連続性に気づかない位相差の最大値とすることができる。例えば、 $\pi/2$ は、 $\pi/2$ に設定される。

【0050】

なお、位相スペクトル広帯域化部16は、最初のフレームについてのみ、高周波数帯域に含まれる周波数fに対する位相を、周波数fよりも所定周波数だけ低い周波数に対する位相と同じ値に設定してもよい。そして位相スペクトル広帯域化部16は、2番目以降のフレームについて、高周波数帯域に含まれる周波数fに対する位相を、上記の推定位相としてもよい。

位相スペクトル広帯域化部16は、高周波数帯域における位相スペクトルをスペクトル合成部17へ出力する。また位相スペクトル広帯域化部16は、高周波数帯域における位相スペクトルを、次のフレームの位相スペクトルの算出に利用できるように、音声帯域拡張装置1が有するメモリに記憶する。

【0051】

スペクトル合成部17は、高周波数帯域における包絡振幅スペクトル、周期性振幅スペクトル、ランダム振幅スペクトル及び位相スペクトルを合成することにより、高周波数帯域における周波数スペクトルを生成する。そしてスペクトル合成部17は、高周波数帯域における周波数スペクトルと、時間周波数変換部11から受け取った、入力音声信号が有する周波数帯域の周波数スペクトルを結合することにより、広帯域周波数スペクトルを生成する。

【0052】

スペクトル合成部17は、高周波数帯域における周波数スペクトルを、次式に従って合成する。

【数 1 3】

$$BR(f) = (PE(f) \cdot (PP(f) + PR(f))) \cdot \cos(\phi(f)) \quad (17)$$

$$BI(f) = (PE(f) \cdot (PP(f) + PR(f))) \cdot \sin(\phi(f))$$

ただし、関数BR(f)は、合成された周波数スペクトルの実部を表し、関数BI(f)は、合成された周波数スペクトルの虚部を表す。また関数PE(f)は、高周波数帯域における包絡振幅スペクトルを表す。また関数PP(f)は、周期性振幅スペクトル広帯域化部 1 4 により生成された、高周波数帯域における周期性振幅スペクトルを表す。さらにPR(f)は、ランダム振幅スペクトル広帯域化部 1 5 により生成された、高周波数帯域におけるランダム振幅スペクトルを表す。さらに $\phi(f)$ は、位相スペクトル広帯域化部 1 6 により生成された、高周波数帯域における位相スペクトルを表す。

10

スペクトル合成部 1 7 は、生成した広帯域周波数スペクトルを周波数時間変換部 1 8 へ出力する。

【0 0 5 3】

周波数時間変換部 1 8 は、スペクトル合成部 1 7 から受け取った広帯域周波数スペクトルを周波数時間変換することにより、擬似的に周波数帯域が拡張された音声信号を生成する。なお、周波数時間変換部 1 8 が実行する周波数時間変換は、時間周波数変換部 1 1 が実行する時間周波数変換の逆変換である。

20

そして周波数時間変換部 1 8 は、生成した音声信号を出力する。

【0 0 5 4】

図 7 は、1 フレーム長の音声信号に対して音声帯域拡張装置 1 により実行される音声帯域拡張処理の動作フローチャートである。なお、音声帯域拡張装置 1 は、図 7 に示される音声帯域拡張処理を、入力音声信号に含まれるフレームの数だけ繰り返し実行する。

まず、時間周波数変換部 1 1 は、バッファメモリ 1 0 から読み出した入力音声信号を、所定のフレーム単位で時間周波数変換することにより、入力音声信号の周波数スペクトルを算出する(ステップ S 5 0 1)。そして時間周波数変換部 1 1 は、フレーム単位で周波数スペクトルを算出する度に、その算出された周波数スペクトルを分離部 1 2 及びスペクトル合成部 1 7 へ出力する。

30

【0 0 5 5】

分離部 1 2 は、時間周波数変換部 1 1 から周波数スペクトルを受け取る度に、その周波数スペクトルから、包絡振幅スペクトルと、周期性振幅スペクトルと、ランダム振幅スペクトルと、位相スペクトルとを抽出する(ステップ S 5 0 2)。分離部 1 2 は、包絡振幅スペクトルを包絡振幅スペクトル広帯域化部 1 3 へ渡す。また分離部 1 2 は、元の周波数スペクトル及び周期性振幅スペクトルと、ケプストラムの最大値及びその最大値に対応するケフレンシー Q_{max} を周期性振幅スペクトル広帯域化部 1 4 へ渡す。さらに分離部 1 2 は、ランダム振幅スペクトルをランダム振幅スペクトル広帯域化部 1 5 へ渡す。そして分離部 1 2 は、元の周波数スペクトル及び位相スペクトルを位相スペクトル広帯域化部 1 6 へ渡す。

40

【0 0 5 6】

ステップ S 5 0 2 の後、包絡振幅スペクトル広帯域化部 1 3 は、分離部 1 2 から受け取った包絡振幅スペクトルに基づいて、入力音声信号が含まれる周波数帯域の上限値よりも高い高周波数帯域における包絡振幅スペクトルを生成する(ステップ S 5 0 3)。そして包絡振幅スペクトル広帯域化部 1 3 は、高周波数帯域における包絡振幅スペクトルをスペクトル合成部 1 7 へ出力する。

また、周期性振幅スペクトル広帯域化部 1 4 は、分離部 1 2 から受け取った周期性振幅

50

スペクトルに基づいて、高周波数帯域における周期性振幅スペクトルを生成する（ステップS504）。そして周期性振幅スペクトル広帯域化部14は、高周波数帯域における周期性振幅スペクトルをスペクトル合成部17へ出力する。

【0057】

さらに、ランダム振幅スペクトル広帯域化部15は、分離部12から受け取ったランダム振幅スペクトルに基づいて、高周波数帯域におけるランダム振幅スペクトルを生成する（ステップS505）。そしてランダム振幅スペクトル広帯域化部15は、高周波数帯域におけるランダム振幅スペクトルをスペクトル合成部17へ出力する。

また、位相スペクトル広帯域化部16は、分離部12から受け取った位相スペクトルに基づいて、高周波数帯域の位相スペクトルを生成する（ステップS506）。そしてランダム振幅スペクトル広帯域化部15は、生成した高周波数帯域の位相スペクトルをスペクトル合成部17へ出力する。

10

【0058】

ステップS506の後、スペクトル合成部17は、高周波数帯域の包絡振幅スペクトル、周期性振幅スペクトル、ランダム振幅スペクトル及び位相スペクトルを合成することにより、高周波数帯域における周波数スペクトルを合成する（ステップS507）。そしてスペクトル合成部17は、入力音声信号が有する周波数帯域における周波数スペクトルと高周波数帯域における周波数スペクトルを結合することにより、広帯域周波数スペクトルを生成する（ステップS508）。スペクトル合成部17は、広帯域周波数スペクトルを周波数時間変換部18へ出力する。

20

【0059】

最後に、周波数時間変換部18は、スペクトル合成部17から受け取った広帯域周波数スペクトルを周波数時間変換することにより、擬似的に周波数帯域が拡大された音声信号を生成する（ステップS509）。

なお、音声帯域拡張装置1は、上記のステップS503～506の処理の実行順序を入れ替えてもよい。あるいは音声帯域拡張装置1は、上記のステップS503～506の処理を並列に実行してもよい。

【0060】

以上に説明してきたように、この音声帯域拡張装置は、入力音声信号の周波数スペクトルから、包絡振幅スペクトル、周期性振幅スペクトル、ランダム振幅スペクトル及び位相スペクトルをそれぞれ抽出し、各スペクトルをそれらの特徴に応じて別個に広帯域化する。そのため、この音声帯域拡張装置は、入力音声信号が有する周波数帯域における各スペクトルの特徴を保ったまま、各振幅スペクトルを広帯域化できる。さらにこの音声帯域拡張装置は、連続するフレーム間で高周波数帯域に含まれる各周波数に対する周波数スペクトルの位相の不連続性を抑制するので、再生音声の不連続性を防止できる。したがって、この音声帯域拡張装置は、再生音声の音質を向上することができる。

30

【0061】

なお、他の実施形態によれば、再生音声の不連続性がユーザにとって許容可能な範囲に収まると想定される場合、音声帯域拡張装置は、位相スペクトル広帯域化部を有さなくてもよい。この場合、音声帯域拡張装置の分離部は、周波数スペクトルから位相スペクトルを算出しない。その代わりに、音声帯域拡張装置のスペクトル合成部は、例えば、高周波数帯域に含まれる各周波数に対する周波数スペクトルの位相を所定の設定値としてもよい。

40

【0062】

図8は、上述した音声帯域拡張装置が組み込まれた電話機の概略構成図である。電話機100は、制御部101と、ベースバンド処理部102と、呼制御部103と、通信部104と、アンテナ105と、マイクロフォン106と、スピーカ107を有する。このうち、制御部101、ベースバンド処理部102、呼制御部103及び通信部104は、それぞれ、別個の回路であってもよく、あるいは、これらの各部は、それら回路が集積された一つの集積回路であってもよい。

【0063】

50

制御部101は、電話機100全体を制御する。そして制御部101は、電話機100で動作する各種のアプリケーションプログラムを実行する。そのために、制御部101は、プロセッサと不揮発性メモリ及び揮発性メモリを有する。そして制御部101は、電話機100が有するキーパッドなどの操作部(図示せず)を介したユーザの操作により、電話などの通信を行うアプリケーションが起動されると、そのアプリケーションにしたがって呼制御部103を動作させる。

【0064】

また制御部101は、マイクロホン106から取得した音声信号に対して情報源符号化処理を実行する。そして制御部101は、得られた信号をアップリンク信号としてベースバンド処理部102に渡す。また制御部101は、ベースバンド処理部102からダウンリンク信号を受け取ると、情報源符号化された音声信号を復号する。

10

さらに、制御部101は、上述した音声帯域拡張装置1を有する。そして制御部101は、復号された音声信号に対して、周波数帯域を拡張する処理を実行する。そして制御部101は、スピーカ107に、周波数帯域が拡張された音声信号を再生させる。

【0065】

ベースバンド処理部102は、制御部101からアップリンク信号を受け取り、そのアップリンク信号に対して畳込み符号化あるいはターボ符号化などの誤り訂正用符号化処理及び拡散処理などの送信処理を実行し、その符号化されたアップリンク信号を通信部104へ出力する。

またベースバンド処理部102は、通信部104から受け取ったダウンリンク信号に対して、逆拡散処理及び誤り訂正復号処理などの受信処理を実行する。そしてベースバンド処理部102は、受信処理が施されたダウンリンク信号を制御部101へ出力する。

20

【0066】

呼制御部103は、電話機100と基地局装置との間における、呼び出し、応答、切断などの呼制御処理を実行する。そして呼制御部103は、その呼制御処理の結果に応じて、ベースバンド処理部102に対して動作の開始または終了を指示する。

【0067】

通信部104は、ベースバンド処理部102から受信した、符号化されたアップリンク信号に対してDifferential Quadrature Phase Shift Keying (DQPSK)などの直交変調処理を行う。通信部104は、直交変調されたアップリンク信号を、無線周波数を持つ搬送波に重畳する。そして通信部104は、搬送波に重畳されたアップリンク信号を増幅し、その増幅されたアップリンク信号をアンテナ105を介して送信する。

30

また通信部104は、基地局から発信されたダウンリンク信号をアンテナ105を介して受信する。そして通信部104は、受信したダウンリンク信号を増幅する。そして、通信部104は、増幅されたダウンリンク信号を復調する。そして通信部104は、復調されたダウンリンク信号をベースバンド処理部102に渡す。

【0068】

このように、本実施形態に係る音声帯域拡張装置が組み込まれた電話機は、受信した音声信号の周波数帯域を擬似的に拡張するので、再生される音声の品質を向上することができる。特に、この電話機は、受信した音声信号の周波数スペクトルから、包絡振幅スペクトル、周期性振幅スペクトル、ランダム振幅スペクトル及び位相スペクトルをそれぞれ抽出し、各スペクトルをそれらの特徴に応じて別個に広帯域化する。そのため、この電話機は、音声信号が有する周波数帯域における各スペクトルの特徴を保ったまま、各振幅スペクトルを広帯域化できる。さらにこの電話機は、連続するフレーム間で高周波数帯域に含まれる各周波数に対する周波数スペクトルの位相の不連続性を抑制するので、再生音声の不連続になることを防止できる。したがって、この電話機は、再生音声の音質を向上することができる。

40

【0069】

ここに挙げられた全ての例及び特定の用語は、読者が、本発明及び当該技術の促進に対する本発明者により寄与された概念を理解することを助ける、教示的な目的において意図

50

されたものであり、本発明の優位性及び劣等性を示すことに関する、本明細書の如何なる例の構成、そのような特定の挙げられた例及び条件に限定しないように解釈されるべきものである。本発明の実施形態は詳細に説明されているが、本発明の精神及び範囲から外れることなく、様々な変更、置換及び修正をこれに加えることが可能であることを理解されたい。

【 0 0 7 0 】

以上説明した実施形態及びその変形例に関し、更に以下の付記を開示する。

(付記 1)

第 1 の周波数帯域を有する音声信号を、所定の時間長を有するフレームごとに時間周波数変換することにより、該音声信号の周波数スペクトルを算出する時間周波数変換部と、

前記周波数スペクトルから、該周波数スペクトルの包絡振幅スペクトルと、周波数に対して周期的にスペクトル強度が変動する周期性振幅スペクトルと、周波数に対してランダムにスペクトル強度が変動するランダム振幅スペクトルとを抽出する分離部と、

前記包絡振幅スペクトルが有する周波数帯域を前記第 1 の周波数帯域と異なる第 2 の周波数帯域まで拡張することにより、前記包絡振幅スペクトルを広帯域化する包絡振幅スペクトル広帯域化部と、

前記周期性振幅スペクトルが有する周波数帯域を前記第 2 の周波数帯域まで拡張することにより、前記周期性振幅スペクトルを広帯域化する周期性振幅スペクトル広帯域化部と、

前記ランダム振幅スペクトルが有する周波数帯域を前記第 2 の周波数帯域まで拡張することにより、前記ランダム振幅スペクトルを広帯域化するランダム振幅スペクトル広帯域化部と、

前記広帯域化された包絡振幅スペクトルと、前記広帯域化された周期性振幅スペクトルと、前記広帯域化されたランダム振幅スペクトルを合成することにより、前記第 1 の周波数帯域及び前記第 2 の周波数帯域を有する広帯域周波数スペクトルを合成する広帯域スペクトル算出部と、

前記広帯域周波数スペクトルを周波数時間変換することにより、前記第 1 の周波数帯域及び前記第 2 の周波数帯域を有する音声信号を生成する周波数時間変換部と、

を有する音声帯域拡張装置。

(付記 2)

前記周期性振幅スペクトル広帯域化部は、前記第 1 の周波数帯域における周期性振幅スペクトルの包絡線を算出し、前記第 2 の周波数帯域においても当該包絡線を維持するように前記周期性振幅スペクトルを広帯域化する、付記 1 に記載の音声帯域拡張装置。

(付記 3)

前記周期性振幅スペクトル広帯域化部は、前記第 2 の周波数帯域において周波数が高くなるほど前記広帯域化された周期性振幅スペクトルの周波数に対する周期性を弱くする、付記 1 または 2 に記載の音声帯域拡張装置。

(付記 4)

前記ランダム振幅スペクトル広帯域化部は、前記第 1 の周波数帯域におけるランダム振幅スペクトルの包絡線を算出し、前記第 2 の周波数帯域においても当該包絡線を維持するように前記ランダム振幅スペクトルを広帯域化する、付記 1 ~ 3 の何れか一項に記載の音声帯域拡張装置。

(付記 5)

前記第 1 の周波数帯域に含まれる各周波数に対する前記周波数スペクトルの位相を表す位相スペクトルが有する周波数帯域を前記第 2 の周波数帯域にまで拡張することにより、前記位相スペクトルを広帯域化する位相スペクトル広帯域化部をさらに有し、

前記広帯域スペクトル算出部は、前記広帯域化された包絡振幅スペクトルと、前記広帯域化された周期性振幅スペクトルと、前記広帯域化されたランダム振幅スペクトルと、前記広帯域化された位相スペクトルを合成することにより、前記広帯域周波数スペクトルを合成する、付記 1 ~ 4 の何れか一項に記載の音声帯域拡張装置。

(付記6)

前記位相スペクトル広帯域化部は、第1のフレームにおける前記第2の周波数帯域に含まれる所定の周波数に対する前記周波数スペクトルの位相を、前記第1のフレームよりも前の第2のフレームにおいて決定された当該所定の周波数に対する前記周波数スペクトルの位相、当該所定の周波数及びフレーム長から算出される、前記第1のフレームの開始時点における位相と連続するように決定する、付記5に記載の音声帯域拡張装置。

(付記7)

第1の周波数帯域を有する音声信号を、所定の時間長を有するフレームごとに時間周波数変換することにより、該音声信号の周波数スペクトルを算出し、

前記周波数スペクトルから、該周波数スペクトルの包絡振幅スペクトルと、周波数に対して周期的にスペクトル強度が変動する周期性振幅スペクトルと、周波数に対してランダムにスペクトル強度が変動するランダム振幅スペクトルとを抽出し、

前記包絡振幅スペクトルが有する周波数帯域を前記第1の周波数帯域と異なる第2の周波数帯域まで拡張することにより、前記包絡振幅スペクトルを広帯域化し、

前記周期性振幅スペクトルが有する周波数帯域を前記第2の周波数帯域まで拡張することにより、前記周期性振幅スペクトルを広帯域化し、

前記ランダム振幅スペクトルが有する周波数帯域を前記第2の周波数帯域まで拡張することにより、前記ランダム振幅スペクトルを広帯域化し、

前記広帯域化された包絡振幅スペクトルと、前記広帯域化された周期性振幅スペクトルと、前記広帯域化されたランダム振幅スペクトルを合成することにより、前記第1の周波数帯域及び前記第2の周波数帯域を有する広帯域周波数スペクトルを合成し、

前記広帯域周波数スペクトルを周波数時間変換することにより、前記第1の周波数帯域及び前記第2の周波数帯域を有する音声信号を生成する、
ことを含む音声帯域拡張方法。

(付記8)

第1の周波数帯域を有する音声信号を、所定の時間長を有するフレームごとに時間周波数変換することにより、該音声信号の周波数スペクトルを算出し、

前記周波数スペクトルから、該周波数スペクトルの包絡振幅スペクトルと、周波数に対して周期的にスペクトル強度が変動する周期性振幅スペクトルと、周波数に対してランダムにスペクトル強度が変動するランダム振幅スペクトルとを抽出し、

前記包絡振幅スペクトルが有する周波数帯域を前記第1の周波数帯域と異なる第2の周波数帯域まで拡張することにより、前記包絡振幅スペクトルを広帯域化し、

前記周期性振幅スペクトルが有する周波数帯域を前記第2の周波数帯域まで拡張することにより、前記周期性振幅スペクトルを広帯域化し、

前記ランダム振幅スペクトルが有する周波数帯域を前記第2の周波数帯域まで拡張することにより、前記ランダム振幅スペクトルを広帯域化し、

前記広帯域化された包絡振幅スペクトルと、前記広帯域化された周期性振幅スペクトルと、前記広帯域化されたランダム振幅スペクトルを合成することにより、前記第1の周波数帯域及び前記第2の周波数帯域を有する広帯域周波数スペクトルを合成し、

前記広帯域周波数スペクトルを周波数時間変換することにより、前記第1の周波数帯域及び前記第2の周波数帯域を有する音声信号を生成する、
ことをコンピュータに実行させる音声帯域拡張用コンピュータプログラム。

(付記9)

第1の周波数帯域を有する符号化された音声信号を受信する通信部と、

前記音声信号を復号するベースバンド処理部と、

前記音声信号を広帯域化する制御部であって、

前記音声信号を、所定の時間長を有するフレームごとに時間周波数変換することにより、該音声信号の周波数スペクトルを算出し、

前記周波数スペクトルから、該周波数スペクトルの包絡振幅スペクトルと、周波数に対して周期的にスペクトル強度が変動する周期性振幅スペクトルと、周波数に対してラン

10

20

30

40

50

ダムにスペクトル強度が変動するランダム振幅スペクトルとを抽出し、

前記包絡振幅スペクトルが有する周波数帯域を前記第 1 の周波数帯域と異なる第 2 の周波数帯域まで拡張することにより、前記包絡振幅スペクトルを広帯域化し、

前記周期性振幅スペクトルが有する周波数帯域を前記第 2 の周波数帯域まで拡張することにより、前記周期性振幅スペクトルを広帯域化し、

前記ランダム振幅スペクトルが有する周波数帯域を前記第 2 の周波数帯域まで拡張することにより、前記ランダム振幅スペクトルを広帯域化し、

前記広帯域化された包絡振幅スペクトルと、前記広帯域化された周期性振幅スペクトルと、前記広帯域化されたランダム振幅スペクトルを合成することにより、前記第 1 の周波数帯域及び前記第 2 の周波数帯域を有する広帯域周波数スペクトルを合成し、

前記広帯域周波数スペクトルを周波数時間変換することにより、前記第 1 の周波数帯域及び前記第 2 の周波数帯域を有する広帯域音声信号を生成する、
制御部と、

前記広帯域化音声信号を再生するスピーカと、
を有する電話機。

【符号の説明】

【 0 0 7 1 】

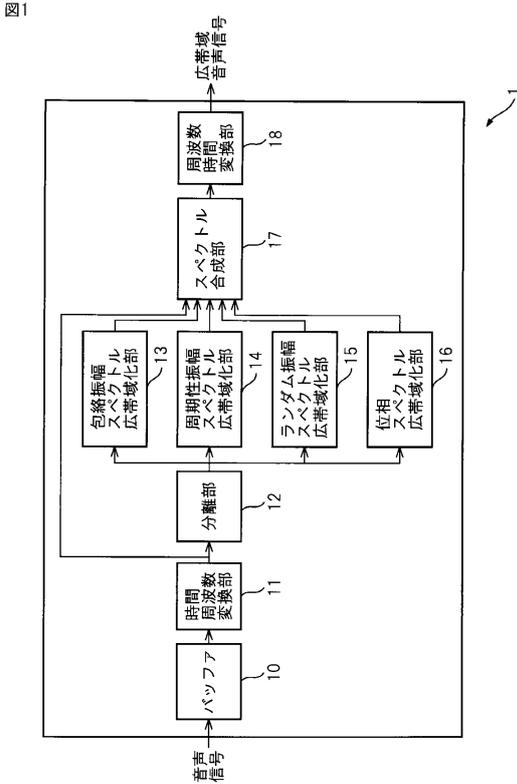
- 1 音声帯域拡張装置
- 1 0 バッファメモリ
- 1 1 時間周波数変換部
- 1 2 分離部
- 1 3 包絡振幅スペクトル広帯域化部
- 1 4 周期性振幅スペクトル広帯域化部
- 1 5 ランダム振幅スペクトル広帯域化部
- 1 6 位相スペクトル広帯域化部
- 1 7 スペクトル合成部
- 1 8 周波数時間変換部
- 1 0 0 電話機
- 1 0 1 制御部
- 1 0 2 ベースバンド処理部
- 1 0 3 呼制御部
- 1 0 4 通信部
- 1 0 5 アンテナ
- 1 0 6 マイクロフォン
- 1 0 7 スピーカ

10

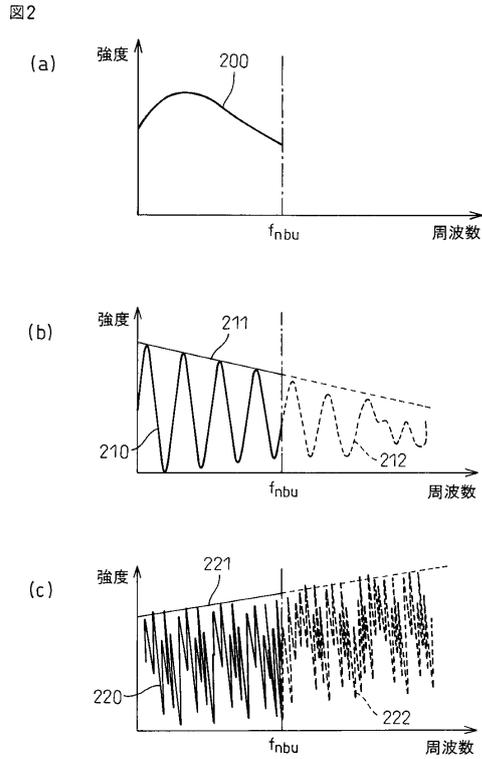
20

30

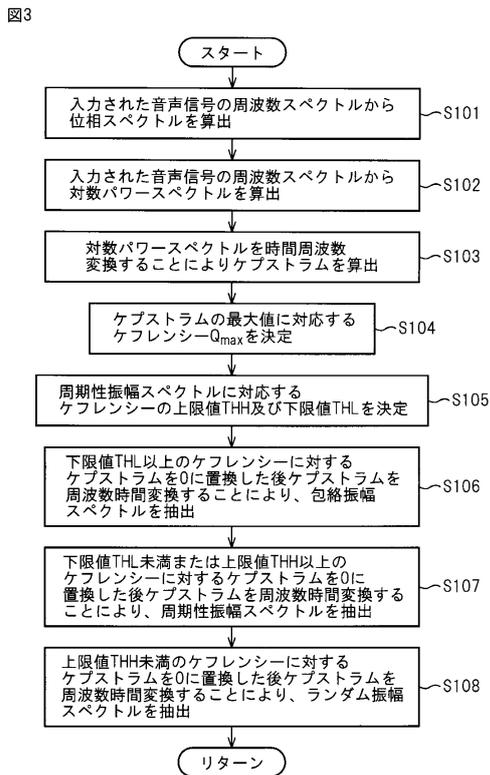
【図1】



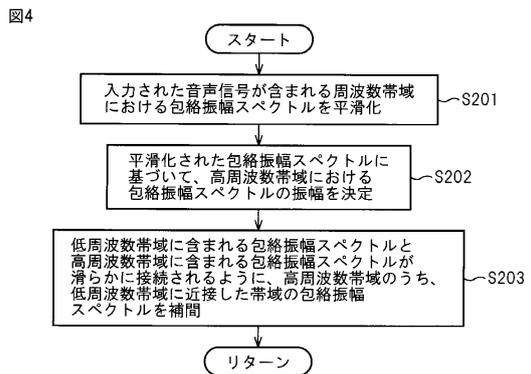
【図2】



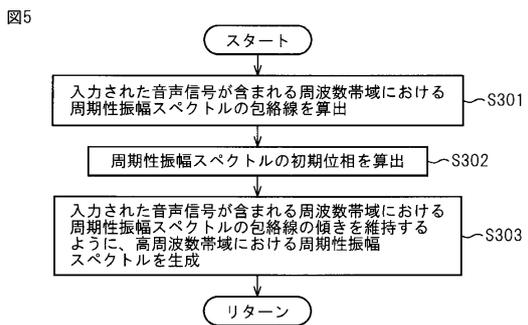
【図3】



【図4】

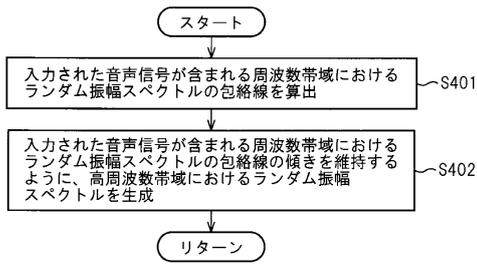


【図5】



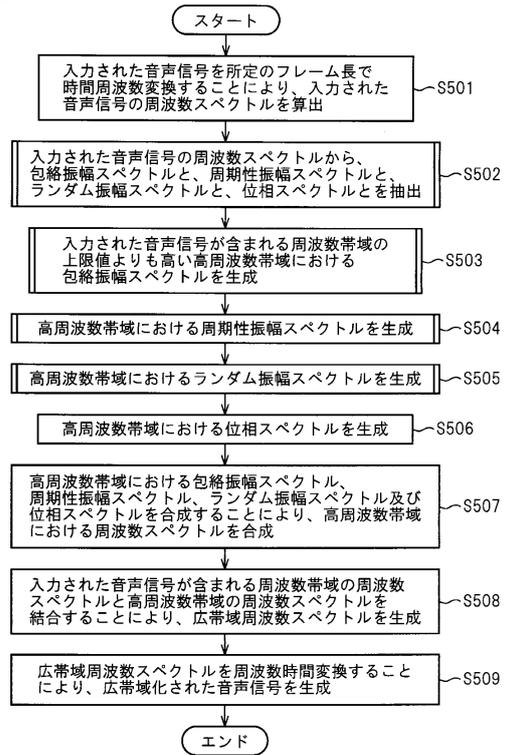
【図6】

図6



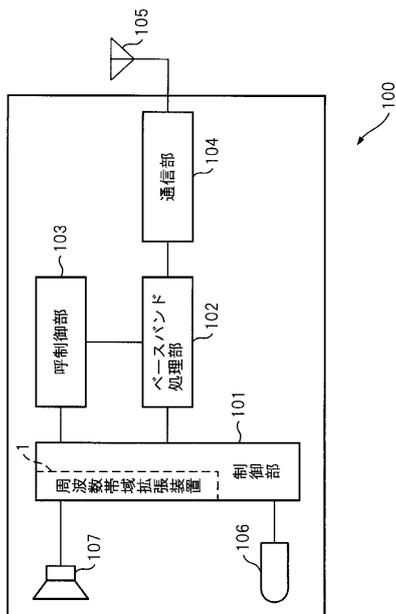
【図7】

図7



【図8】

図8



フロントページの続き

- (74)代理人 100114177
弁理士 小林 龍
- (72)発明者 遠藤 香緒里
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
- (72)発明者 大谷 猛
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
- (72)発明者 外川 太郎
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
- (72)発明者 大田 恭士
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

審査官 菊地 陽一

- (56)参考文献 特開2001-100773(JP,A)
特開平09-101798(JP,A)
特開平08-123495(JP,A)
特表2004-521394(JP,A)

- (58)調査した分野(Int.Cl., DB名)
G10L 21/0388
G10L 19/06