

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4358221号
(P4358221)

(45) 発行日 平成21年11月4日(2009.11.4)

(24) 登録日 平成21年8月14日(2009.8.14)

(51) Int.Cl.		F I		
G 1 0 L	19/00	(2006.01)	G 1 0 L	19/00 4 0 0 A
G 1 0 L	19/14	(2006.01)	G 1 0 L	19/14 5 2 0 B
G 1 0 L	21/02	(2006.01)	G 1 0 L	21/02 1 0 2 B

請求項の数 2 (全 29 頁)

(21) 出願番号	特願2006-316566 (P2006-316566)	(73) 特許権者	000006013 三菱電機株式会社 東京都千代田区丸の内二丁目7番3号
(22) 出願日	平成18年11月24日(2006.11.24)	(74) 代理人	100099461 弁理士 溝井 章司
(62) 分割の表示	特願2004-158788 (P2004-158788) の分割	(72) 発明者	田崎 裕久 東京都千代田区丸の内二丁目7番3号 三 菱電機株式会社内
原出願日	平成10年12月7日(1998.12.7)	審査官	菊池 智紀
(65) 公開番号	特開2007-79606 (P2007-79606A)		
(43) 公開日	平成19年3月29日(2007.3.29)		
審査請求日	平成18年11月24日(2006.11.24)		
(31) 優先権主張番号	特願平9-336803		
(32) 優先日	平成9年12月8日(1997.12.8)		
(33) 優先権主張国	日本国(JP)		

最終頁に続く

(54) 【発明の名称】 音信号加工方法及び音信号加工装置

(57) 【特許請求の範囲】

【請求項1】

音声符号から複数のパラメータを生成し、当該複数のパラメータを用いて前記音声符号に対応した復号音声生成ステップと、

前記復号音声生成ステップにおいて生成された前記復号音声の振幅の時間軸方向の変動が小さくされた第一の加工音声を、前記復号音声の所定時点における振幅に、前記所定時点から所定時間だけ前の時点における前記復号音声の振幅と、前記所定時点から前記所定時間だけ後の時点における前記復号音声の振幅とを重み付け加算する重み付け加算処理を行って生成する第一加工音声生成ステップと、

前記復号音声と前記第一の加工音声とを混合して第二の加工音声を生成する第二加工音声生成ステップであって、前記復号音声生成ステップによって生成された複数のパラメータの少なくとも一つに基づく前記復号音声の雑音らしさが高くなるにつれて大きくなる第一の重み付け係数と、周波数成分毎に異なる第二の重み付け係数とで重み付けられた前記第一の加工音声と、前記復号音声とを加算して第二の加工音声を生成する第二加工音声生成ステップとを備えたことを特徴とする音信号加工方法。

【請求項2】

音声符号から複数のパラメータを生成し、当該複数のパラメータを用いて前記音声符号に対応した復号音声生成ステップと、

前記復号音声生成ステップによって生成された前記復号音声の振幅の時間軸方向の変動が小さくされた第一の加工音声を、前記復号音声の所定時点における振幅に、前記所定時点か

ら所定時間だけ前の時点における前記復号音声の振幅と、前記所定時点から前記所定時間だけ後の時点における前記復号音声の振幅とを重み付け加算する重み付け加算処理を行って生成する第一加工音声生成手段と、

前記復号音声と前記第一の加工音声とを混合して第二の加工音声を生成する第二加工音声生成手段であって、前記復号音声生成手段によって生成された複数のパラメータの少なくとも一つに基づく前記復号音声の雑音らしさが高くなるにつれて大きくなる第一の重み付け係数と、周波数成分毎に異なる第二の重み付け係数とで重み付けられた前記第一の加工音声と、前記復号音声とを加算して第二の加工音声を生成する第二加工音声生成手段とを備えたことを特徴とする音信号加工装置。

【発明の詳細な説明】

10

【技術分野】

【0001】

本発明は、音声や楽音などの符号化復号化処理によって発生する量子化雑音や、雑音抑圧処理などのさまざまな信号加工処理によって生じる歪み、などの主観的に好ましくない成分を主観的に感じにくいように加工する音信号加工方法および音信号加工装置に関する。

【背景技術】

【0002】

音声や楽音などの情報源符号化の圧縮率を高めていくと、次第に符号化時の歪みである量子化雑音が増えてくるし、量子化雑音が変形してきて主観的に耐えられないものになってくる。一例を挙げて説明すると、PCM(Pulse Code Modulation)やADPCM(Adaptive Differential Pulse Code Modulation)のような信号自体を忠実に表現しようとする音声符号化方式の場合には、量子化雑音は乱数状であり、主観的にもあまり気にならないが、圧縮率が高まり、符号化方式が複雑になるにつれて、量子化雑音に符号化方式固有のスペクトル特性が表れ、主観的に大きな劣化となる場合がでてくる。特に背景雑音が支配的な信号区間においては、高圧縮率の音声符号化方式が利用している音声モデルが合わないため、非常に聞き苦しい音になってしまう。

20

【0003】

また、スペクトルサブトラクション法などの雑音抑圧処理を行った場合、雑音の推定誤差が処理後の信号上に歪みとして残り、これが処理前の信号と大きく異なる特性をもっているために、主観評価を大きく劣化させることがある。

30

【0004】

上記のような量子化雑音や歪みによる主観評価の低下を抑制する従来の方法としては、特開平8-130513号、特開平8-146998号、特開平7-160296号、特開平6-326670号、特開平7-248793号、およびS.F.Boll著 *raction SSP-27, No. 2, pp. 113-120, April 1979* (以降文献1と呼ぶ)に開示されているものがある。

【0005】

特開平8-130513号は、背景雑音区間の品質改善を目的としたもので、背景雑音のみの区間であるか否かを判定して、背景雑音のみの区間に専用の符号化処理または復号化処理を行うようにし、背景雑音のみの区間の復号化を行う場合に合成フィルタの特性を抑制することで、聴感的に自然な再生音を得るようにしたものである。

40

【0006】

特開平8-146998号は、白色雑音が符号化復号化によって耳障りな音色になることを抑制することを狙って、復号音声に対して白色雑音や予め格納しておいた背景雑音を加えるようにしたものである。

【0007】

特開平7-160296号は、量子化雑音を聴感的に低減することを狙って、復号音声または音声復号化部が受信したスペクトルパラメータに関するインデックスを基に、聴覚

50

マスキング閾値を求め、これを反映したフィルタ係数を求めて、この係数をポストフィルタに使用するようにしたものである。

【0008】

特開平6-326670号は、通信電力制御などのために音声を含まない区間で符号伝送を停止するシステムでは、符号伝送の無い時には復号側で疑似背景雑音を生成して出力するが、この時に発生する、音声区間に含まれる実際の背景雑音と無音区間の疑似背景雑音の間の違和感を軽減することを狙ったもので、音声を含まない区間だけでなく音声区間にも疑似背景雑音を重畳するようにしたものである。

【0009】

特開平7-248793号は、雑音抑圧処理によって発生する歪み音を聴感的に軽減することを目的としたもので、符号化側では、まず雑音区間か音声区間か判定し、雑音区間では雑音スペクトルを伝送し、音声区間では雑音抑圧処理後のスペクトルを伝送し、復号化側では、雑音区間では受信した雑音スペクトルを用いて合成音を生成して出力し、音声区間では受信した雑音抑圧処理後のスペクトルを用いて生成した合成音に、雑音区間で受信した雑音スペクトルを用いて生成した合成音に重畳倍率を乗じて加算して出力するようにしたものである。

10

【0010】

文献1は、雑音抑圧処理によって発生する歪み音を聴感的に軽減することを狙い、雑音抑圧処理後の出力音声に対して、時間的に前後の区間と振幅スペクトル上の平滑化を行い、更に背景雑音区間に限って振幅抑圧処理を行っている。

20

【特許文献1】特開平8-130513号

【特許文献2】特開平8-146998号

【特許文献3】特開平7-160296号

【特許文献4】特開平6-326670号

【特許文献5】特開平7-248793号

【非特許文献1】S.F. Boll著 reaction SSP-27, No. 2, p. 113-120, April 1979)

【発明の開示】

【発明が解決しようとする課題】

【0011】

30

上記の従来法には、以下に述べる課題がある。

【0012】

特開平8-130513号には、符号化処理や復号化処理を区間判定結果に従って大きく切り替えているために、雑音区間と音声区間の境界で特性の急変が起こる課題がある。特に雑音区間を音声区間と誤判定することが頻繁に起こった場合、本来比較的定常である雑音区間が不安定に変動してしまい、かえって雑音区間の劣化を起こす場合がある。雑音区間判定結果を伝送する場合、伝送するための情報の追加が必要で、更にその情報が伝送路上で誤った場合に、不必要な劣化を引き起こす課題がある。また、合成フィルタの特性を抑制するだけでは、音源符号化の際に生じる量子化雑音は軽減されないため、雑音種によっては改善効果がほとんど得られない課題がある。

40

【0013】

特開平8-146998号には、予め用意してある雑音を加えてしてしまうために、符号化された現在の背景雑音の特性が失われてしまう課題がある。劣化音を聞こえにくくするためには劣化音を上回るレベルの雑音を加える必要があり、再生される背景雑音が大きくなってしまふ課題がある。

【0014】

特開平7-160296号では、スペクトルパラメータに基づいて聴覚マスキング閾値を求めて、これに基づいてスペクトルポストフィルタを行うだけであるので、スペクトルが比較的平坦な背景雑音などでは、マスキングされる成分もほとんどなく、全く改善効果が得られない課題がある。また、マスキングされない主要成分については、大きな変化を

50

与えることができないので、主要成分に含まれている歪みについては何らの改善効果も得られない課題がある。

【0015】

特開平6-326670号では、実際の背景雑音に関係なく疑似背景雑音を生成しているので、実際の背景雑音の特性が失われてしまう課題がある。

【0016】

特開平7-248793号には、符号化处理や復号化处理を区間判定結果に従って大きく切り替えているために、雑音区間か音声区間かの判定を誤ると大きな劣化を引き起こす課題がある。雑音区間の一部を音声区間と誤った場合には、雑音区間内の音質が不連続に変動して聞き苦しくなる。逆に音声区間を雑音区間と誤った場合には、平均雑音スペクトルを用いた雑音区間の合成音と、音声区間で重畳される雑音スペクトルを用いた合成音に音声成分が混入し、全体的に音質劣化が起こる課題がある。更に、音声区間における劣化音を聞こえなくするためには、決して小さくない雑音を重畳することが必要である。

【0017】

文献1には、平滑化のために半区間分(10ms~20ms程度)の処理遅延が発生する課題がある。また、雑音区間内の一部を音声区間と誤判定してしまった場合、雑音区間内の音質が不連続に変動して聞き苦しくなる課題がある。

【0018】

この発明は、かかる課題を解決するためになされたものであり、区間判定誤りによる劣化が少なく、雑音種やスペクトル形状への依存度が少なく、大きな遅延時間を必要としない、実際の背景雑音の特性を残すことができ、背景雑音レベルを過度に大きくすることがなく、新たな伝送情報の追加が不要で、音源符号化などによる劣化成分についても良好な抑圧効果を与えることのできる音信号加工方法および音信号加工装置を提供することを目的としている。

【課題を解決するための手段】

【0019】

この発明の音信号加工方法は、音声符号から複数のパラメータを生成し、当該複数のパラメータを用いて前記音声符号に対応した復号音声生成ステップと、前記復号音声生成ステップにおいて生成された前記復号音声の振幅の時間軸方向の変動が小さくされた第一の加工音声を、前記復号音声の所定時点における振幅に、前記所定時点から所定時間だけ前の時点における前記復号音声の振幅と、前記所定時点から前記所定時間だけ後の時点における前記復号音声の振幅とを重み付け加算する重み付け加算処理を行って生成する第一加工音声生成ステップと、前記復号音声と前記第一の加工音声とを混合して第二の加工音声を生成する第二加工音声生成ステップであって、前記復号音声生成ステップによって生成された複数のパラメータの少なくとも1つに基づく前記復号音声の雑音らしさが高くなるにつれて大きくなる第一の重み付け係数と、周波数成分毎に異なる第二の重み付け係数とで重み付けられた前記第一の加工音声と、前記復号音声とを加算して第二の加工音声を生成する第二加工音声生成ステップとを備えたことを特徴とする。

【0020】

この発明の音信号加工装置は、音声符号から複数のパラメータを生成し、当該複数のパラメータを用いて前記音声符号に対応した復号音声を生成する復号音声生成手段と、前記復号音声生成手段によって生成された前記復号音声の振幅の時間軸方向の変動が小さくされた第一の加工音声を、前記復号音声の所定時点における振幅に、前記所定時点から所定時間だけ前の時点における前記復号音声の振幅と、前記所定時点から前記所定時間だけ後の時点における前記復号音声の振幅とを重み付け加算する重み付け加算処理を行って生成する第一加工音声生成手段と、前記復号音声と前記第一の加工音声とを混合して第二の加工音声を生成する第二加工音声生成手段であって、前記復号音声生成手段によって生成された複数のパラメータの少なくとも1つに基づく前記復号音声の雑音らしさが高くなるにつれて大きくなる第一の重み付け係数と、周波数成分毎に異なる第二の重み付け係数とで重み付けられた前記第一の加工音声と、前記復号音声とを加算して第二の加工音声を生成

10

20

30

40

50

する第二加工音声生成手段とを備えたことを特徴とする。

【発明の効果】

【0021】

以上説明したように本発明の音信号加工方法は、音声符号から複数のパラメータを生成し、当該複数のパラメータを用いて前記音声符号に対応した復号音声生成ステップと、前記復号音声生成ステップにおいて生成された前記復号音声の振幅の時間軸方向の変動が小さくされた第一の加工音声を、前記復号音声の所定時点における振幅に、前記所定時点から所定時間だけ前の時点における前記復号音声の振幅と、前記所定時点から前記所定時間だけ後の時点における前記復号音声の振幅とを重み付け加算する重み付け加算処理を行って生成する第一加工音声生成ステップと、前記復号音声と前記第一の加工音声とを混合して第二の加工音声を生成する第二加工音声生成ステップであって、前記復号音声生成ステップによって生成された複数のパラメータの少なくとも1つに基づく前記復号音声の雑音らしさが高くなるにつれて大きくなる第一の重み付け係数と、周波数成分毎に異なる第二の重み付け係数とで重み付けられた前記第一の加工音声と、前記復号音声とを加算して第二の加工音声を生成する第二加工音声生成ステップとを備えたことにより、主観品質を改善できる効果がある。

10

【0022】

本発明の音信号加工装置は、音声符号から複数のパラメータを生成し、当該複数のパラメータを用いて前記音声符号に対応した復号音声生成手段と、前記復号音声生成手段によって生成された前記復号音声の振幅の時間軸方向の変動が小さくされた第一の加工音声を、前記復号音声の所定時点における振幅に、前記所定時点から所定時間だけ前の時点における前記復号音声の振幅と、前記所定時点から前記所定時間だけ後の時点における前記復号音声の振幅とを重み付け加算する重み付け加算処理を行って生成する第一加工音声生成手段と、前記復号音声と前記第一の加工音声とを混合して第二の加工音声を生成する第二加工音声生成手段であって、前記復号音声生成手段によって生成された複数のパラメータの少なくとも1つに基づく前記復号音声の雑音らしさが高くなるにつれて大きくなる第一の重み付け係数と、周波数成分毎に異なる第二の重み付け係数とで重み付けられた前記第一の加工音声と、前記復号音声とを加算して第二の加工音声を生成する第二加工音声生成手段とを備えたことにより、主観品質を改善できる効果がある。

20

【発明を実施するための最良の形態】

30

【0023】

以下図面を参照しながら、この発明の実施の形態について説明する。

【0024】

実施の形態1.

図1は、本実施の形態による音信号加工方法を適用した音声復号方法の全体構成を示し、図中1は音声復号装置、2はこの発明による信号加工方法を実行する信号加工部、3は音声符号、4は音声復号部、5は復号音声、6は出力音声である。信号加工部2は、信号変形部7、信号評価部12、重み付き加算部18より構成されている。信号変形部7は、フーリエ変換部8、振幅平滑化部9、位相擾乱部10、逆フーリエ部11より構成されている。信号評価部12は、逆フィルタ部13、パワー算出部14、背景雑音らしさ算出部15、推定背景雑音パワー更新部16、推定雑音スペクトル更新部17より構成されている。

40

【0025】

以下、図に基づいて動作を説明する。

【0026】

まず音声符号3が音声復号装置1内の音声復号部4に入力される。なお、この音声符号3は、別途音声符号化部が音声信号を符号化した結果として出力され、通信路や記憶デバイスを介してこの音声復号部4に入力される。

【0027】

音声復号部4は、音声符号3に対して、前記音声符号化部と対を成す復号処理を行い、

50

得られた所定の長さ（１フレーム長）の信号を復号音声５として出力する。そして、この復号音声５は、信号加工部２内の信号変形部７、信号評価部１２、重み付き加算部１８に入力される。

【００２８】

信号変形部７内のフーリエ変換部８は、入力された現フレームの復号音声５と必要に応じ前フレームの復号音声５の最新部分を合わせた信号に対して、窓がけを行い、窓がけ後の信号に対してフーリエ変換処理を行うことで周波数毎のスペクトル成分を算出し、これを振幅平滑化部９に出力する。なお、フーリエ変換処理としては、離散フーリエ変換（ＤＦＴ）、高速フーリエ変換（ＦＦＴ）などが代表的である。窓がけ処理としては、台形窓、方形窓、ハニング窓など様々なものが適用可能であるが、ここでは、台形窓の両端の傾斜部分をそれぞれハニング窓の半分ずつに置換した変形台形窓を使用する。実際の形状例、復号音声５や出力音声６との時間関係については、図面を用いて後述説明する。

10

【００２９】

振幅平滑化部９は、フーリエ変換部８から入力された周波数毎のスペクトルの振幅成分に対して平滑化処理を行い、平滑化後のスペクトルを位相擾乱部１０に出力する。ここで用いる平滑化処理としては、周波数軸方向、時間軸方向の何れを用いても、量子化雑音などの劣化音の抑制効果が得られる。しかし、周波数軸方向の平滑化をあまり強くすると、スペクトルの怠けが生じ、本来の背景雑音の特性を損なってしまうことが多い。一方、時間軸方向の平滑化についても、あまり強くしていくと、長時間にわたって同じ音が残ることになり、反響感が発生してしまう。色々な背景雑音に対して調整を進めた結果、周波数軸方向の平滑化はなし、時間軸方向は振幅を対数領域で平滑化する、とした場合が出力音声６の品質が良かった。その時の平滑化方法は、次式で表わされる。

20

【００３０】

$$y_i = y_{i-1} (1 - \alpha) + x_i \quad \dots \text{式 1}$$

ここで、 x_i が現在のフレーム（第 i フレーム）の平滑化前の対数振幅スペクトル値、 y_{i-1} が前フレーム（第 $i-1$ フレーム）の平滑化後の対数振幅スペクトル値、 y_i が現在のフレーム（第 i フレーム）の平滑化後の対数振幅スペクトル値、 α が 0 ~ 1 の値を持つ平滑化係数である、平滑化係数 α はフレーム長、解消したい劣化音のレベルなどによって最適値が異なるが、概ね 0.5 程度の値となる。

【００３１】

位相擾乱部１０は、振幅平滑化部９から入力された平滑化後のスペクトルの位相成分に擾乱を与え、擾乱後のスペクトルを逆フーリエ変換部１１に出力する。各位相成分に擾乱を与える方法としては、乱数を用いて所定範囲の位相角を生成し、これを元々の位相角に加算すれば良い。位相角生成の範囲の制限を設けない場合には、各位相成分を乱数で生成した位相角に単に置換すればよい。符号化などによる劣化が大きい場合には、位相角生成の範囲は制限しない。

30

【００３２】

逆フーリエ変換部１１は、位相擾乱部１０から入力された擾乱後のスペクトルに対して逆フーリエ変換処理を行うことで、信号領域に戻し、前後のフレームとの滑らかな接続のための窓がけを行いつつ接続していき、得られた信号を変形復号音声３４として重み付き加算部１８に出力する。

40

【００３３】

信号評価部１２内の逆フィルタ部１３は、後述する推定雑音スペクトル更新部１７内に格納されている推定雑音スペクトルパラメータを用いて、前記音声復号部４から入力された復号音声５に対する逆フィルタ処理を行い、逆フィルタされた復号音声をパワー算出部１４に出力する。この逆フィルタ処理によって、背景雑音の振幅が大きい、つまり音声と背景雑音が拮抗している可能性が高い成分の振幅抑圧を行っており、逆フィルタ処理を行わない場合に比べて、音声区間と背景雑音区間の信号パワー比が大きくとれるようになっている。

【００３４】

50

なお、推定雑音スペクトルパラメータは、音声符号化処理や音声復号処理との親和性、ソフトウェアの共有化といった観点で選択する。現状では多くの場合、線スペクトル対(LSP)を使用する。LSPの他にも、線形予測係数(LPC)、ケプストラムなどのスペクトル包絡パラメータ、または振幅スペクトルそのものを用いても類似の効果を得ることができる。後述する推定雑音スペクトル更新部17における更新処理としては線形補間や平均処理などを用いる構成が簡単であり、スペクトル包絡パラメータの中では線形補間や平均処理を行ってもフィルタが安定であることが保証できるLSPとケプストラムが適している。雑音成分のスペクトルに対する表現力としてはケプストラムが優れているが、逆フィルタ部の構成の容易さという点ではLSPが勝る。振幅スペクトルを用いる場合には、この振幅スペクトル特性をもつLPCを算出して逆フィルタに使用するか、復号音声5をフーリエ変換した結果(フーリエ変換部8の出力に等しい)に対して振幅変形処理を行って逆フィルタと同様の効果を実現すればよい。

10

【0035】

パワー算出部14は、逆フィルタ部13から入力された逆フィルタされた復号音声のパワーを求め、算出されたパワー値を背景雑音らしさ算出部15に出力する。

【0036】

背景雑音らしさ算出部15は、パワー算出部14から入力されたパワーと、後述する推定雑音パワー更新部16内に格納されている推定雑音パワーを用いて、現在の復号音声5の背景雑音らしさを算出し、これを加算制御値35として重み付き加算部18に出力する。また、算出した背景雑音らしさを後述する推定雑音パワー更新部16と推定雑音スペクトル更新部17に対して出力し、パワー算出部14から入力されたパワーを後述する推定雑音パワー更新部16に対して出力する。ここで、背景雑音らしさについては、最も単純には、次式によって算出できる。

20

【0037】

$$v = \log(p_N) - \log(p) \cdots \text{式2}$$

ここで、 p がパワー算出部14から入力されたパワー、 p_N が推定雑音パワー更新部16内に格納されている推定雑音パワー、 v が算出された背景雑音らしさである。

【0038】

この場合、 v の値が大きい程(負値であればその絶対値が小さい程)背景雑音らしい、ということになる。この他にも、 p_N/p を計算して v とするなど、様々な算出方法が考えられる。

30

【0039】

推定雑音パワー更新部16は、背景雑音らしさ算出部15から入力された背景雑音らしさとパワーを用いて、その内部に格納してある推定雑音パワーの更新を行う。例えば、入力された背景雑音らしさが高い(v の値が大きい)時に、次式に従い、入力されたパワーを推定雑音パワーに反映させることで更新を行う。

【0040】

$$\log(p_N) = (1 - \alpha) \log(p_N) + \log(p) \cdots \text{式3}$$

ここで、 α は0~1の値を取る更新速度定数で、比較的0に近い値に設定するとよい。この式の右辺の値を求めて、左辺の p_N を新しい推定雑音パワーとすることで更新を行う。

40

【0041】

なお、この推定雑音パワーの更新方法については、更に推定精度を向上させるためにフレーム間での変動性を参照したり、入力された過去のパワーを複数格納しておいて、統計分析によって雑音パワーの推定を行ったり、 p の最低値をそのまま推定雑音パワーとしたりするなど様々な変形、改良が可能である。

【0042】

推定雑音スペクトル更新部17は、まず入力された復号音声5を分析して、現在のフレームのスペクトルパラメータを算出する。算出するスペクトルパラメータについては逆フ

50

ィルタ部 13 にて説明した通りで、多くの場合 LSP を使用する。そして、背景雑音らしさ算出部 15 から入力され背景雑音らしさとここで算出したスペクトルパラメータを用いて、内部に格納してある推定雑音スペクトルを更新する。例えば、入力された背景雑音らしさが高い (v の値が大きい) 時に、次式に従い、算出したスペクトルパラメータを推定雑音スペクトルに反映させることで更新を行う。

【0043】

$$x_N = (1 - \alpha) x_N + x \quad \text{式 4}$$

ここで、 x が現在のフレームのスペクトルパラメータ、 x_N が推定雑音スペクトル (パラメータ) である。 α は 0 ~ 1 の値を取る更新速度定数で、比較的 0 に近い値に設定するとよい。この式の右辺の値を求めて、左辺の x_N を新しい推定雑音スペクトル (パラメータ) とすることで更新を行う。

10

【0044】

なお、この推定雑音スペクトルの更新方法についても、上記推定雑音パワーの更新方法と同様に様々な改良が可能である。

【0045】

そして、最後の処理として、重み付き加算部 18 は、信号評価部 12 から入力された加算制御値 35 に基づいて、音声復号部 4 から入力された復号音声 5 と信号変形部 7 から入力された変形復号音声 34 を重み付けして加算し、得られた出力音声 6 を出力する。重み付け加算の制御方法の動作としては、加算制御値 35 が大きく (背景雑音らしさが高く) なるにつれて復号音声 5 に対する重みを小さく、変形復号音声 34 に対する重みを大きく制御する。逆に加算制御値 35 が小さく (背景雑音らしさが低く) なるにつれて復号音声 5 に対する重みを大きく、変形復号音声 34 に対する重みを小さく制御する。

20

【0046】

なお、フレーム間での重みの急変に伴う出力音声 6 の品質劣化を抑制するために、加算制御値 35 または重み付け係数をサンプル毎に徐々に変化するように平滑化を行うことが望ましい。

【0047】

図 2 には、この重み付け加算部 18 における、加算制御値に基づく重み付け加算の制御例を示す。

【0048】

図 2 (a) では、加算制御値 35 に対する 2 つの閾値 v_1 と v_2 を用いて線形制御している場合である。加算制御値 35 が v_1 未満の場合には、復号音声 5 に対する重み付け係数 w_S を 1、変形復号音声 34 に対する重み付け係数 w_N を 0 とする。加算制御値 35 が v_2 以上の場合には、復号音声 5 に対する重み付け係数 w_S を 0、変形復号音声 34 に対する重み付け係数 w_N を A_N とする。そして加算制御値 35 が v_1 以上で v_2 未満の場合には、復号音声 5 に対する重み付け係数 w_S を 1 ~ 0、変形復号音声 34 に対する重み付け係数 w_N を 0 ~ A_N の間で線形的に計算して与えている。

30

【0049】

この様に制御することで、確実に背景雑音区間であると判断できる場合 (v_2 以上) には変形復号信号 34 のみが出力され、確実に音声区間であると判断できる場合 (v_1 未満) には復号音声 5 そのものが出力され、音声区間か背景雑音区間か判断がつかない場合 (v_1 以上 v_2 未満) には、どちらの傾向が強いかにより依存した比率で復号音声 5 と変形復号音声 34 が混合された結果が出力される。

40

【0050】

なお、ここで確実に背景雑音区間であると判断できる場合 (v_2 以上) に変形復号信号 34 に乗じる重み付け係数値 A_N として 1 以下の値を与えれば、結果的に背景雑音区間の振幅抑圧効果が得られる。逆に 1 以上の値を与えれば、背景雑音区間の振幅強調効果が得られる。背景雑音区間は、音声符号化復号化処理によって振幅低下が起こる場合が多く、その場合には背景雑音区間の振幅強調を行うことによって、背景雑音の再現性を向上することができる。振幅抑圧と振幅強調のどちらを行うかは適用対象、使用者の要求などに依

50

存する。

【0051】

図2(b)では、新たな閾値 v_3 を追加し、 v_1 と v_3 間、 v_3 と v_2 間で重み付け係数を線形的に計算して与えた場合である。閾値 v_3 の位置における重み付け係数の値を調整することで、音声区間か背景雑音区間か判断がつかない場合(v_1 以上 v_2 未満)における混合比率を更に細かく設定することができる。一般に位相の相関が低い2つの信号を加算した場合、得られる信号のパワーは加算前の2つの信号のパワーの合計より小さくなる。 v_1 以上 v_2 未満の範囲における2つの重み付け係数の合計を1ないし w_N より大きくすることで、このパワー低下を抑制することができる。なお、図2(a)によって得られた重み付け係数の平方根をとって更に定数を乗じた値を新たに重み付け係数とするこ

10

【0052】

図2(c)では、図2(a)の v_1 未満の範囲における変形復号音声34に与える重み付け係数 w_N として0より大きい B_N という値を与え、これに応じて v_1 以上 v_2 未満の範囲における w_N も修正した場合である。背景雑音レベルが高い場合や、符号化における圧縮率が非常に高い場合など、音声区間における量子化雑音や劣化音が大きい場合には、この様に確実に音声区間と分かっている範囲においても、変形復号音声を加算することで、劣化音を聞こえにくくすることができる。

【0053】

図2(d)は、背景雑音らしさ算出部15において、推定雑音パワーを現在のパワーで除算した結果(p_N/p)を背景雑音らしさ(加算制御値35)として出力した場合に対応する制御例である。この場合、加算制御値35は復号音声5中に含まれる背景雑音の比率を示しているため、この値に比例した比率で混合されるように重み付け係数を算出している。具体的には、加算制御値35が1以上の場合には w_N が1で w_S が0、1未満の場合には、 w_N が加算制御値35そのもの、 w_S が $(1 - w_N)$ となっている。

20

【0054】

図3には、フーリエ変換部8における切り出し窓、逆フーリエ変換部11における接続のための窓の実際の形状例、復号音声5との時間関係を説明する説明図を示す。

【0055】

復号音声5は、音声復号部4から所定の時間長(1フレーム長)毎に出力されてくる。ここでこの1フレーム長を N サンプルとする。図3(a)は、この復号音声5の一例を示しており、 $x(0) \sim x(N-1)$ が入力された現在のフレームの復号音声5に当たる。フーリエ変換部8では、図3(a)に示されるこの復号音声5に対して図3(b)に示す変形台形窓を乗じることで、長さ $(N + NX)$ の信号を切り出す。 NX は変形台形窓の両端の1未満の値を持つ区間のそれぞれの長さである。この両端の区間は長さ $(2NX)$ のハニング窓を前半と後半に2分割したものに等しい。逆フーリエ変換部11では、逆フーリエ変換処理によって生成した信号に対して、図3(c)に示す変形台形窓を乗じ、(図3(c)に破線で示すように)前後のフレームで得られた同信号と時間関係を守りつつ信号の加算を行って、連続する変形復号音声34(図3(d))を生成する。

30

【0056】

次のフレームの信号との接続のための区間(長さ NX)については、現在のフレーム時点では変形復号音声34が確定していない。すなわち、新たに確定する変形復号音声34は、 $x(-NX) \sim x(N - NX - 1)$ である。このため、現在のフレームの復号音声5に対して得られる出力音声6は、次式の通りとなる。

40

【0057】

$$y(n) = x(n) + x(n) \cdots \text{式5} \\ (n = -NX, \dots, N - NX - 1)$$

ここで、 $y(n)$ が出力音声6である。この時、信号加工部2としての処理遅延は最低でも NX だけ必要となる。

【0058】

50

この処理遅延 NX が許容できない適用対象の場合、復号音声 5 と変形復号音声 3 4 の時間的ズレを許容して、次式のように出力音声 6 を生成することもできる。

【0059】

$$y(n) = x(n) + x(n - NX) \cdots \text{式6}$$

$$(n = 0, \dots, N - 1)$$

この場合、復号音声 5 と変形復号音声 3 4 の時間関係にズレがあるので、位相擾乱部 10 における擾乱が弱い（つまり復号音声の位相特性がある程度残っている）場合や、フレーム内でスペクトルやパワーが急変する場合には劣化を生じる場合がある。特に重み付け加算部 18 における重み付け係数が大きく変化するときと、2つの重み付け係数が拮抗している場合に劣化を生じ易い。しかし、それらの劣化は比較的少なく、信号加工部の導入効果の方が十分に大きい。よって処理遅延 NX が許容できない適用対象についても、この方法を用いることができる。

【0060】

なお、この図 3 の場合、フーリエ変換前と逆フーリエ変換後に変形台形窓を乗じており、接続部分の振幅低下を招く場合がある。この振幅低下も、位相擾乱部 10 における擾乱が弱い場合に起こりやすい。そのような場合には、フーリエ変換前の窓を方形窓に変更することで振幅低下の抑制が得られる。通常、位相擾乱部 10 によって位相が大きく変形された結果、逆フーリエ変換後の信号に最初の変形台形窓の形状が現れてこないため、前後のフレームの変形復号音声 3 4 とのスムーズな接続のために 2 つ目の窓がけが必要になる。

【0061】

なお、ここでは、信号変形部 7、信号評価部 12、重み付け加算部 18 の処理を全てフレーム毎に行ったが、これに限ったものではない。例えば、1 フレームを複数のサブフレームに分割し、信号評価部 12 の処理をサブフレーム毎に行ってサブフレーム毎の加算制御値 35 を算出し、重み付け加算部 18 における重み付け制御もサブフレーム毎に行っても良い。信号変形処理にフーリエ変換を使用しているので、フレーム長があまり短いとスペクトル特性の分析結果が不安定になり、変形復号音声 3 4 が安定しにくい。一方、背景雑音らしさはもっと短い区間に対しても比較的安定に算出できるので、サブフレーム毎に算出して重み付けを細かく制御することで音声の立ち上がり部分などにおける品質改善効果が得られる。

【0062】

また、信号評価部 12 の処理をサブフレーム毎に行って、フレーム内の全ての加算制御値を組み合わせて、少数の加算制御値 35 を算出することもできる。音声区間を背景雑音らしいと誤りたくない場合には、全ての加算制御値の内の最小値（背景雑音らしさの最小値）を選択してフレームを代表する加算制御値 35 として出力すれば良い。

【0063】

更に、復号音声 5 のフレーム長と信号変形部 7 の処理フレーム長は同一である必要はない。例えば、復号音声 5 のフレーム長が短くて、信号変形部 7 内のスペクトル分析にとって短すぎる場合には、複数フレームの復号音声 5 を蓄積して、一括して信号変形処理を行うようにすれば良い。但し、この場合には、複数フレームの復号音声 5 を蓄積するために処理遅延が発生してしまう。この他、復号音声 5 のフレーム長と全く独立に信号変形部 7 や信号加工部 2 全体の処理フレーム長を設定しても構わない。この場合、信号のバッファリングが複雑になるが、様々な復号音声 5 のフレーム長に依存することなく、信号加工処理にとって最適の処理フレーム長を選択でき、信号加工部 2 の品質が最も良くなる効果がある。

【0064】

また、ここでは、背景雑音らしさの算出に、逆フィルタ部 13、パワー算出部 14、背景雑音らしさ算出部 15、推定背景雑音レベル更新部 16、推定雑音スペクトル更新部 17 を使用したが、背景雑音らしさを評価するものであれば、この構成に限ったものではない。

【 0 0 6 5 】

この実施の形態 1 によれば、入力信号（復号音声）に対して所定の信号加工処理を行うことで、入力信号に含まれる劣化成分を主観的に気にならないようにした加工信号（変形復号音声）を生成し、所定の評価値（背景雑音らしさ）によって入力信号と加工信号の加算重みを制御するようにしたので、劣化成分が多く含まれる区間を中心に加工信号の比率を増やして、主観品質を改善できる効果がある。

【 0 0 6 6 】

また、スペクトル領域で信号加工処理を行うようにしたことで、スペクトル領域での細かい劣化成分の抑圧処理を行うことができ、更に主観品質を改善できる効果がある。

【 0 0 6 7 】

また、加工処理として振幅スペクトル成分の平滑化処理と位相スペクトル成分の擾乱付与処理を行うようにしたので、量子化雑音などによって生じる振幅スペクトル成分の不安定な変動を良好に抑圧することができ、更に、位相成分間に独特な相互関係を持ってしまい特徴的な劣化と感じられることが多い量子化雑音に対して、位相成分間の関係に擾乱を与えることができ、主観品質を改善できる効果がある。

【 0 0 6 8 】

また、従来の音声区間または背景雑音区間のどちらか、という 2 値区間判定を廃し、背景雑音らしさという連続量を算出して、これに基づいて連続的に復号音声と変形復号音声の重み付け加算係数を制御するようにしたので、区間判定誤りによる品質劣化を回避できる効果がある。

【 0 0 6 9 】

また、音声区間における量子化雑音や劣化音が大きい場合には、確実に音声区間と分かっている区間においても、変形復号音声を加算することで、劣化音を聞こえにくくすることができる効果がある。

【 0 0 7 0 】

また、背景雑音の情報が多く含まれている復号音声の加工処理によって出力音声を生成しているので、実際の背景雑音の特性を残しつつ、雑音種やスペクトル形状にあまり依存しない安定な品質改善効果が得られるし、音源符号化などによる劣化成分に対しても改善効果が得られる効果がある。

【 0 0 7 1 】

また、現在までの復号音声を用いて処理を行うので特に大きな遅延時間は不要で、復号音声と変形復号音声の加算方法によっては処理時間以外の遅延を排除することもできる効果がある。変形復号音声のレベルを上げる際には復号音声のレベルを下げていくので、従来のように量子化雑音を聞こえなくするために大きな疑似雑音を重畳することも不要で、逆に適用対象に応じて、背景雑音レベルを小さ目にしたり、大き目にしたりすることすら可能である。また、当然のことであるが、音声復号装置または信号加工部内に閉じた処理であるので従来のような新たな伝送情報の追加は不要である。

【 0 0 7 2 】

更に、この実施の形態 1 では、音声復号部と信号加工部が明確に分離されており、両者の間の情報のやりとりも少ないので、既存のものも含めて様々な音声復号装置内に導入することが容易である。

【 0 0 7 3 】

実施の形態 2 .

図 4 は、本実施の形態による音信号加工方法を雑音抑圧方法と組み合わせて適用した音信号加工装置の構成の一部を示す。図中 3 6 は入力信号、8 はフーリエ変換部、1 9 は雑音抑圧部、3 9 はスペクトル変形部、1 2 は信号評価部、1 8 は重み付き加算部、1 1 は逆フーリエ変換部、4 0 は出力信号である。スペクトル変形部 3 9 は、振幅平滑化部 9、位相擾乱部 1 0 より構成されている。

以下、図に基づいて動作を説明する。

【 0 0 7 4 】

まず、入力信号 3 6 が、フーリエ変換部 8 と信号評価部 1 2 に入力される。

【 0 0 7 5 】

フーリエ変換部 8 は、入力された現フレームの入力信号 3 6 と必要に応じ前フレームの入力信号 3 6 の最新部分を合わせた信号に対して、窓がけを行い、窓がけ後の信号に対してフーリエ変換処理を行うことで周波数毎のスペクトル成分を算出し、これを雑音抑圧部 1 9 に出力する。なお、フーリエ変換処理および窓がけ処理については実施の形態 1 と同様である。

【 0 0 7 6 】

雑音抑圧部 1 9 は、フーリエ変換部 8 より入力された周波数毎のスペクトル成分から、雑音抑圧部 1 9 内部に格納してある推定雑音スペクトルを減算し、得られた結果を雑音抑圧スペクトル 3 7 として重み付け加算部 1 8 とスペクトル変形部 3 9 内の振幅平滑化部 9 10 に出力する。これは、いわゆるスペクトルサブトラクション処理の主部に相当する処理である。そして、雑音抑圧部 1 9 は、背景雑音区間であるか否かの判定を行い、背景雑音区間であればフーリエ変換部 8 より入力された周波数毎のスペクトル成分を用いて、内部の推定雑音スペクトルを更新する。なお、背景雑音区間であるか否かの判定は、後述する信号評価部 1 2 の出力結果を流用して行うことで処理を簡易化することも可能である。

【 0 0 7 7 】

スペクトル変形部 3 9 内の振幅平滑化部 9 は、雑音抑圧部 1 9 より入力された雑音抑圧スペクトル 3 7 の振幅成分に対して平滑化処理を行い、平滑化後の雑音抑圧スペクトルを位相擾乱部 1 0 に出力する。ここで用いる平滑化処理としては、周波数軸方向、時間軸方向の何れを用いても、雑音抑圧部が発生させた劣化音の抑制効果が得られる。具体的な平滑化方法については実施の形態 1 と同様のものを用いることができる。 20

【 0 0 7 8 】

スペクトル変形部 3 9 内の位相擾乱部 1 0 は、振幅平滑化部 9 から入力された平滑化後の雑音抑圧スペクトルの位相成分に擾乱を与え、擾乱後のスペクトルを変形雑音抑圧スペクトル 3 8 として重み付き加算部 1 8 に出力する。各位相成分に擾乱を与える方法については実施の形態 1 と同様のものを用いることができる。

【 0 0 7 9 】

信号評価部 1 2 は、入力信号 3 6 を分析して背景雑音らしさを算出し、これを加算制御値 3 5 として重み付け加算部 1 8 に出力する。なお、この信号評価部 1 2 内の構成と各処理については、実施の形態 1 と同様のものを用いることができる。 30

【 0 0 8 0 】

重み付き加算部 1 8 は、信号評価部 1 2 から入力された加算制御値 3 5 に基づいて、雑音抑圧部 1 9 から入力された雑音抑圧スペクトル 3 7 とスペクトル変形部 3 9 から入力された変形雑音抑圧スペクトル 3 8 を重み付けして加算し、得られたスペクトルを逆フーリエ変換部 1 1 に出力する。重み付け加算の制御方法の動作としては、実施の形態 1 と同様に、加算制御値 3 5 が大きく（背景雑音らしさが高く）なるにつれて雑音抑圧スペクトル 3 7 に対する重みを小さく、変形雑音抑圧スペクトル 3 8 に対する重みを大きく制御する。逆に加算制御値 3 5 が小さく（背景雑音らしさが低く）なるにつれて雑音抑圧スペクトル 3 7 に対する重みを大きく、変形雑音抑圧スペクトル 3 8 に対する重みを小さく制御する。 40

【 0 0 8 1 】

そして、最後の処理として、逆フーリエ変換部 1 1 は、重み付き加算部 1 8 から入力されたスペクトルに対して逆フーリエ変換処理を行うことで、信号領域に戻し、前後のフレームとの滑らかな接続のための窓がけを行いつつ接続していき、得られた信号を出力信号 4 0 として出力する。接続のための窓がけと接続処理については、実施の形態 1 と同様である。

【 0 0 8 2 】

この実施の形態 2 によれば、雑音抑圧処理等によって劣化したスペクトルに対して所定の加工処理を行うことで、劣化成分を主観的に気にならないようにした加工スペクトル (50

変形雑音抑圧スペクトル)を生成し、所定の評価値(背景雑音らしさ)によって加工前のスペクトルと加工スペクトルの加算重みを制御するようにしたので、劣化成分が多く含まれて主観品質の低下につながっている区間(背景雑音区間)を中心に加工スペクトルの比率を増やして、主観品質を改善できる効果がある。

【0083】

また、スペクトル領域での重み付け加算を行うようにしたので、実施の形態1に比べると加工処理のためのフーリエ変換と逆フーリエ変換が不要となり、処理が簡易になる効果がある。なお、この実施の形態2におけるフーリエ変換部8と逆フーリエ変換11は、雑音抑圧部19のために元々必要な構成である。

【0084】

また、加工処理として振幅スペクトル成分の平滑化処理と位相スペクトル成分の擾乱付与処理を行うようにしたので、量子化雑音などによって生じる振幅スペクトル成分の不安定な変動を良好に抑圧することができ、更に、位相成分間に独特な相互関係を持ってしまい特徴的な劣化と感じられることが多い量子化雑音や劣化成分に対して、位相成分間の関係に擾乱を与えることができ、主観品質を改善できる効果がある。

【0085】

また、背景雑音区間であるか否かという2値区間判定ではなく、背景雑音らしさという連続量を算出して、これに基づいて連続的に重み付け加算係数を制御するようにしたので、区間判定誤りによる品質劣化を回避できる効果がある。

【0086】

また、背景雑音区間以外における劣化音が大きい場合には、図2(c)のような重み付け加算を行うことで、確実に背景雑音区間以外と分かっている区間においても変形雑音抑圧スペクトルを加算し、劣化音を聞こえにくくすることができる効果がある。

【0087】

また、雑音抑圧スペクトルに対して、単純な処理を直接施して変形雑音抑圧スペクトルを生成しているので、雑音種やスペクトル形状にあまり依存しない安定な品質改善効果が得られる効果がある。

【0088】

また、現在までの雑音抑圧スペクトルを用いて処理を行うので、雑音抑圧部19の遅延時間に追加して、大きな遅延時間がいらぬ特長を持つ。変形雑音抑圧スペクトルの加算レベルをあげる際には元々の雑音抑圧スペクトルの加算レベルを下げていくので、量子化雑音を聞こえなくするために比較的大きな雑音を重畳することも不要で、背景雑音レベルを小さくすることができる効果がある。また、当然のことであるが、この処理を音声符号化処理の前処理などとして用いる場合にも、符号化部内に閉じた処理となるので従来のような新たな伝送情報の追加は不要である。

【0089】

実施の形態3.

図1との対応部分に同一符号を付けた図5は、本実施の形態による音信号加工方法を適用した音声復号装置の全体構成を示し、図中20は信号変形部7の変形強度を制御する情報を出力する変形強度制御部である。変形強度制御部20は、聴覚重み付け部21、フーリエ変換部22、レベル判定部23、連続性判定部24、変形強度算出部25より構成されている。

【0090】

以下、図に基づいて動作を説明する。

【0091】

音声復号部4から出力された復号音声5が、信号加工部2内の信号変形部7、変形強度制御部20、信号評価部12、重み付き加算部18に入力される。

【0092】

変形強度制御部20内の聴覚重み付け部21は、音声復号部4より入力された復号音声5に対して、聴覚重み付け処理を行い、得られた聴覚重み付け音声をフーリエ変換部2

10

20

30

40

50

2 に出力する。ここで、聴覚重み付け処理としては、音声符号化処理（音声復号部 4 で行った音声復号処理と対を成すもの）で使用されているものと同様な処理を行う。

【0093】

C E L P などの符号化処理で良く用いられる聴覚重み付け処理は、符号化対象の音声を分析して線形予測係数（L P C）を算出し、これに定数乗算を行って 2 つの変形 L P C を求め、この 2 つの変形 L P C をフィルタ係数とする A R M A フィルタを構成し、このフィルタを用いたフィルタリング処理によって聴覚重み付けを行う、というものである。復号音声 5 に対して符号化処理と同様の聴覚重み付けを行うためには、受信した音声符号 3 を復号して得られた L P C、もしくは復号音声 5 を再分析して算出した L P C を出発点として、2 つの変形 L P C を求め、これを用いて聴覚重み付けフィルタを構成すれば良い。

10

【0094】

C E L P などの符号化処理では、聴覚重み付け後の音声上での歪みを最小化するように符号化を行うので、聴覚重み付け後の音声において、振幅が大きいスペクトル成分は、量子化雑音の重畳が少ない、ということになる。従って、符号化時の聴覚重み付け音声に近い音声を復号化部 1 内で生成できれば、信号変形部 7 における変形強度の制御情報として有用である。

【0095】

なお、音声復号部 4 における音声復号処理にスペクトルポストフィルタなどの加工処理が含まれている場合（C E L P の場合にはほとんどに含まれている）には、本来であればまず復号音声 5 からスペクトルポストフィルタなどの加工処理の影響を除去した音声を生成するか、音声復号部 4 内からこの加工処理直前の音声を抽出するかして、該音声に対して聴覚重み付けを行うことによって、符号化時の聴覚重み付け音声に近い音声を得られる。しかし、背景雑音区間の品質改善を主な目的とする場合には、この区間におけるスペクトルポストフィルタなどの加工処理の影響は少なく、その影響を除去しなくても効果に大差は出ない。この実施の形態 3 は、スペクトルポストフィルタなどの加工処理の影響除去を行わない構成としている。

20

【0096】

なお、当然のことであるが、符号化処理において聴覚重み付けを行っていない場合や、その効果が小さくて無視しても良い場合には、この聴覚重み付け部 2 1 は不要となる。その場合、信号変形部 7 内のフーリエ変換部 8 の出力を、後述するレベル判定部 2 3 と連続性判定部 2 4 に与えればよいので、フーリエ変換部 2 2 も不要とできる。

30

【0097】

更に、スペクトル領域でも非線型振幅変換処理など聴覚重み付けに近い効果をもたらす方法があるので、符号化処理内で使用している聴覚重み付け方法との誤差を無視して構わない場合には、信号変形部 7 内のフーリエ変換部 8 の出力をこの聴覚重み付け部 2 1 への入力とし、聴覚重み付け部 2 1 がこの入力に対してスペクトル領域での聴覚重み付けを行い、フーリエ変換部 2 2 を省略して、後述するレベル判定部 2 3 と連続性判定部 2 4 に聴覚重み付けされたスペクトルを出力するように構成することも可能である。

【0098】

変形強度制御部 2 0 内のフーリエ変換部 2 2 は、聴覚重み付け部 2 1 より入力された聴覚重み付け音声と必要に応じ前フレームの聴覚重み付け音声の最新部分を合わせた信号に対して、窓がけを行い、窓がけ後の信号に対してフーリエ変換処理を行うことで周波数毎のスペクトル成分を算出し、これを聴覚重み付けスペクトルとしてレベル判定部 2 3 と連続性判定部 2 4 に出力する。なお、フーリエ変換処理および窓がけ処理については実施の形態 1 のフーリエ変換部 8 と同様である。

40

【0099】

レベル判定部 2 3 は、フーリエ変換部 2 2 から入力された聴覚重み付けスペクトルの各振幅成分の値の大きさに基づいて、各周波数毎の第一の変形強度を算出し、これを変形強度算出部 2 5 に出力する。聴覚重み付けスペクトルの各振幅成分の値が小さい程量子化雑音の比率が大きいため、第一の変形強度を強くすればよい。最も単純には、全振幅成分の

50

平均値を求めて、この平均値に所定の閾値 T_h を加算して、これを上回る成分に対しては第一の変形強度を 0、これを下回る成分に対しては第一の変形強度を 1 とすればよい。図 6 には、この閾値 T_h を用いた場合の聴覚重み付けスペクトルと第一の変形強度の関係を示す。なお、第一の変形強度の算出方法はこれに限定されるものではない。

【 0 1 0 0 】

連続性判定部 2 4 は、フーリエ変換部 2 2 から入力された聴覚重み付けスペクトルの各振幅成分または各位相成分の時間方向の連続性を評価し、この評価結果に基づいて、各周波数毎の第二の変形強度を算出し、これを変形強度算出部 2 5 に出力する。聴覚重み付けスペクトルの振幅成分の時間方向の連続性、位相成分の（フレーム間の時間推移による位相の回転を補償した後の）連続性が低い周波数成分については、良好な符号化が行われていたとは考えにくいので、第二の変形強度を強くする。この第二の変形強度の算出についても、最も単純には所定の閾値を用いた判定によって 0 または 1 を与える方法を用いることができる。

10

【 0 1 0 1 】

変形強度算出部 2 5 は、レベル判定部 2 3 より入力された第一の変形強度と、連続性判定部 2 4 より入力された第二の変形強度に基づいて、各周波数毎の最終的な変形強度を算出し、これを信号変形部 7 内の振幅平滑化部 9 と位相擾乱部 1 0 に出力する。この最終的な変形強度については、第一の変形強度と第二の変形強度の最小値、重み付き平均値、最大値などを用いることができる。以上でこの実施の形態 3 にて新たに加わった変形強度制御部 2 0 の動作の説明を終了する。

20

【 0 1 0 2 】

次に、この変形強度制御部 2 0 の追加に伴って、動作に変更がある構成要素について説明する。

【 0 1 0 3 】

振幅平滑化部 9 は、変形強度制御部 2 0 より入力された変形強度に従い、フーリエ変換部 8 から入力された周波数毎のスペクトルの振幅成分に対して平滑化処理を行い、平滑化後のスペクトルを位相擾乱部 1 0 に出力する。なお、変形強度が強い周波数成分程、平滑化を強めるように制御する。平滑化強度の強さを制御する最も単純な方法は、入力された変形強度が大きいときにのみ平滑化を行うようにすればよい。この他にも平滑化を強める方法としては、実施の形態 1 で説明した平滑化の数式における平滑化係数を小さくしたり、固定的な平滑化を行った後のスペクトルと平滑化前のスペクトルを重み付き加算して最終的なスペクトルを生成するように構成しておき、平滑化前のスペクトルに対する重みを小さくするなど様々な方法を用いることができる。

30

【 0 1 0 4 】

位相擾乱部 1 0 は、変形強度制御部 2 0 より入力された変形強度に従い、振幅平滑化部 9 から入力された平滑化後のスペクトルの位相成分に擾乱を与え、擾乱後のスペクトルを逆フーリエ変換部 1 1 に出力する。なお、変形強度が強い周波数成分程、位相の擾乱を大きく与えるように制御する。擾乱の大きさを制御する最も単純な方法は、入力された変形強度が大きいときにのみ擾乱を与えるようにすればよい。この他にも擾乱を制御する方法としては、乱数で生成する位相角の範囲を大小させるなど様々な方法を用いることができる。

40

【 0 1 0 5 】

その他の構成要素については、実施の形態 1 と同様であるため説明を省略する。

【 0 1 0 6 】

なお、ここでは、レベル判定部 2 3 と連続性判定部 2 4 の両方の出力結果を使用したが一方向だけを使用するようにして、残るもう一方は省略する構成も可能である。また、変形強度によって制御する対象を、振幅平滑化部 9 と位相擾乱部 1 0 の一方のみとする構成でも構わない。

【 0 1 0 7 】

この実施の形態 3 によれば、入力信号（復号音声）または聴覚重み付けされた入力信号

50

(復号音声)の各周波数成分毎の振幅の大きさ、各周波数毎の振幅や位相の連続性の大きさに基づいて、加工信号(変形復号音声)を生成する際の変形強度を周波数毎に制御するようにしたので、実施の形態1が持つ効果に加えて、前記振幅スペクトル成分が小さいために量子化雑音や劣化成分が支配的になっている成分、スペクトル成分の連続性が低いために量子化雑音や劣化成分が多くなりがちな成分に対して重点的に加工が加えられ、量子化雑音や劣化成分が少ない良好な成分まで加工してしまうことがなくなり、入力信号や実際の背景雑音の特性を比較的良好に残しつつ量子化雑音や劣化成分を主観的に抑圧でき、主観品質を改善できる効果がある。

【0108】

実施の形態4 .

図5との対応部分に同一符号を付けた図7は、本実施の形態による音信号加工方法を適用した音声復号装置の全体構成を示し、図中41は加算制御値分割部であり、図5における信号変形部7の部分は、フーリエ変換部8、スペクトル変形部39、逆フーリエ変換部11の構成に変更している。

【0109】

以下、図に基づいて動作を説明する。

【0110】

音声復号部4から出力された復号音声5は、信号加工部2内のフーリエ変換部8、変形強度制御部20、信号評価部12に入力される。

【0111】

フーリエ変換部8は、実施の形態2と同様にして、入力された現フレームの復号音声5と必要に応じ前フレームの復号音声5の最新部分を合わせた信号に対して、窓がけを行い、窓がけ後の信号に対してフーリエ変換処理を行うことで周波数毎のスペクトル成分を算出し、これを復号音声スペクトル43として重み付き加算部18とスペクトル変形部39内の振幅平滑化部9に出力する。

【0112】

スペクトル変形部39は、実施の形態2と同様にして、入力された復号音声スペクトル43に対して、振幅平滑化部9、位相擾乱部10の処理を順に行い、得られたスペクトルを変形復号音声スペクトル44として、重み付き加算部18に出力する。

【0113】

変形強度制御部20内では、実施の形態3と同様に、入力された復号音声5に対して、聴覚重み付け部21、フーリエ変換部22、レベル判定部23、連続性判定部24、変形強度算出部25の処理を順次行い、得られた周波数毎の変形強度を加算制御値分割部41に出力する。

【0114】

なお、実施の形態3と同様に、符号化処理において聴覚重み付けを行っていない場合や、その効果が小さい場合には、聴覚重み付け部21とフーリエ変換部22は不要となる。その場合、フーリエ変換部8の出力を、レベル判定部23と連続性判定部24に与えればよい。

【0115】

また、フーリエ変換部8の出力をこの聴覚重み付け部21への入力とし、聴覚重み付け部21がこの入力に対してスペクトル領域での聴覚重み付けを行い、フーリエ変換部22を省略して、後述するレベル判定部23と連続性判定部24に聴覚重み付けされたスペクトルを出力するように構成することも可能である。この様に構成することで、処理の簡易化効果が得られる。

【0116】

信号評価部12は、実施の形態1と同様に、入力された復号音声5に対して、背景雑音らしさを求めて、これを加算制御値35として加算制御値分割部41に出力する。

【0117】

新たに加えられた加算制御値分割部41は、変形強度制御部20から入力された周波数

10

20

30

40

50

毎の変形強度と、信号評価部 1 2 から入力された加算制御値 3 5 を用いて、周波数毎の加算制御値 4 2 を生成し、これを重み付き加算部 1 8 に出力する。変形強度が強い周波数については、その周波数の加算制御値 4 2 の値を制御して、重み付き加算部 1 8 における復号音声スペクトル 4 3 の重みを弱く、変形復号音声スペクトル 4 4 の重みを強くする。逆に変形強度が弱い周波数については、その周波数の加算制御値 4 2 の値を制御して、重み付き加算部 1 8 における復号音声スペクトル 4 3 の重みを強く、変形復号音声スペクトル 4 4 の重みを弱くする。つまり、変形強度が強い周波数については、背景雑音らしさが高いわけであるので、その周波数の加算制御値 4 2 を大きくし、逆の場合には、小さくするわけである。

【 0 1 1 8 】

重み付き加算部 1 8 は、加算制御値分割部 4 1 から入力された周波数毎の加算制御値 4 2 に基づいて、フーリエ変換部 8 から入力された復号音声スペクトル 4 3 とスペクトル変形部 3 9 から入力された変形復号音声スペクトル 4 4 を重み付けして加算し、得られたスペクトルを逆フーリエ変換部 1 1 に出力する。重み付け加算の制御方法の動作としては、図 2 にて説明したのと同様に、周波数毎の加算制御値 4 2 が大きい（背景雑音らしさが高い）周波数成分に対しては復号音声スペクトル 4 3 に対する重みを小さく、変形復号音声スペクトル 4 4 に対する重みを大きく制御する。逆に周波数毎の加算制御値 4 2 が小さい（背景雑音らしさが低い）周波数成分に対しては復号音声スペクトル 4 3 に対する重みを大きく、変形復号音声スペクトル 4 4 に対する重みを小さく制御する。

【 0 1 1 9 】

そして、最後の処理として、逆フーリエ変換部 1 1 は、実施の形態 2 と同様にして、重み付き加算部 1 8 から入力されたスペクトルに対して逆フーリエ変換処理を行うことで、信号領域に戻し、前後のフレームとの滑らかな接続のための窓がけを行いつつ接続していき、得られた信号を出力音声 6 として出力する。

【 0 1 2 0 】

なお、加算制御値分割部 4 1 を廃して、信号評価部 1 2 の出力を重み付き加算部 1 8 に与え、変形強度制御部 2 0 の出力である変形強度を振幅平滑化部 9 と位相擾乱部 1 0 に与える構成も可能である。この様にしたものは、実施の形態 3 の構成における重み付き加算処理をスペクトル領域で行うようにしたものに相当する。

【 0 1 2 1 】

更に、実施の形態 3 の場合と同様に、レベル判定部 2 3 と連続性判定部 2 4 の一方だけを使用するようにして、残るもう一方は省略する構成も可能である。

この実施の形態 4 によれば、入力信号（復号音声）または聴覚重み付けされた入力信号（復号音声）の各周波数成分毎の振幅の大きさ、各周波数毎の振幅や位相の連続性の大きさに基づいて、入力信号のスペクトル（復号音声スペクトル）と加工スペクトル（変形復号音声スペクトル）の重み付け加算を周波数成分毎に独立に制御するようにしたので、実施の形態 1 が持つ効果に加えて、前記振幅スペクトル成分が小さいために量子化雑音や劣化成分が支配的になっている成分、スペクトル成分の連続性が低いために量子化雑音や劣化成分が多くなりがちな成分に対して重点的に加工スペクトルの重みを強め、量子化雑音や劣化成分が少ない良好な成分まで加工スペクトルの重みを強めてしまうことがなくなり、入力信号や実際の背景雑音の特性を比較的良好に残しつつ量子化雑音や劣化成分を主観的に抑圧でき、主観品質を改善できる効果がある。

【 0 1 2 2 】

実施の形態 3 と比較すると、平滑化と擾乱という 2 つの周波数毎の変形処理から、1 つの周波数毎の変形処理に変わっており、処理が簡易化される効果がある。

【 0 1 2 3 】

実施の形態 5 .

図 5 との対応部分に同一符号を付けた図 8 は、本実施の形態による音信号加工方法を適用した音声復号装置の全体構成を示し、図中 2 6 は背景雑音らしさ（加算制御値 3 5 ）の時間方向の変動性を判定する変動性判定部である。

10

20

30

40

50

【 0 1 2 4 】

以下、図に基づいて動作を説明する。

【 0 1 2 5 】

音声復号部 4 から出力された復号音声 5 が、信号加工部 2 内の信号変形部 7、変形強度制御部 20、信号評価部 12、重み付き加算部 18 に入力される。信号評価部 12 は、入力された復号音声 5 に対して、背景雑音らしさを評価し、評価結果を加算制御値 35 として、変動性判定部 26 と重み付き加算部 18 に出力する。

【 0 1 2 6 】

変動性判定部 26 は、信号評価部 12 より入力された加算制御値 35 を、その内部に格納している過去の加算制御値 35 と比較し、値の時間方向の変動性が高いか否かを判定し、この判定結果に基づいて第三の変形強度を算出し、これを変形強度制御部 20 内の変形強度算出部 25 に出力する。そして、入力された加算制御値 35 を用いて内部に格納している過去の加算制御値 35 を更新する。

加算制御値 35 などのフレーム（またはサブフレーム）の特性を表すパラメータの時間方向の変動性が高い場合には、復号音声 5 のスペクトルが時間方向に大きく変化している場合が多く、必要以上に強い振幅平滑化や位相擾乱付与を行うと不自然な反響感が発生してしまう。そこで、この第三の変形強度は、加算制御値 35 の時間方向の変動性が高い場合には、振幅平滑化部 9 における平滑化と位相擾乱部 10 における擾乱付与が弱くなるように設定する。なお、フレーム（またはサブフレーム）の特性を表すパラメータであれば、復号音声のパワー、スペクトル包絡パラメータなど、加算制御値 35 以外のパラメータを用いても同様の効果を得ることができる。

【 0 1 2 7 】

変動性の判定方法としては、最も単純には、前フレームの加算制御値 35 との差分の絶対値を所定の閾値と比較して、閾値を上回っていれば変動性が高い、とすれば良い。この他、前フレームおよび前々フレームの加算制御値 35 との差分の絶対値を各々算出して、その一方が所定の閾値を上回っているか否かで判定してもよい。また、信号評価部 12 がサブフレーム毎に加算制御値 35 を算出する場合には、現在のフレーム内または必要に応じて前フレーム内の全サブフレーム間の加算制御値 35 の差分の絶対値を求めて、何れかが所定の閾値を上回っているか否かで判定することもできる。そして、具体的な処理例としては、閾値を上回っていれば第三の変形強度を 0、閾値を下回っていれば第三の変形強度を 1 とする。

【 0 1 2 8 】

変形強度制御部 20 内では、入力された復号音声 5 に対して、聴覚重み付け部 21、フーリエ変換部 22、レベル判定部 23、連続性判定部 24 までは、実施の形態 3 と同様な処理を行う。

【 0 1 2 9 】

そして、変形強度算出部 25 では、レベル判定部 23 より入力された第一の変形強度、連続性判定部 24 より入力された第二の変形強度、変動性判定部 26 より入力された第三の変形強度に基づいて、各周波数毎の最終的な変形強度を算出し、これを信号変形部 7 内の振幅平滑化部 9 と位相擾乱部 10 に出力する。この最終的な変形強度の算出方法としては、第三の変形強度を全周波数に対して一定値として与え、周波数毎にこの全周波数に拡張した第三の変形強度、第一の変形強度、第二の変形強度の最小値、重み付き平均値、最大値などを求めて最終的な変形強度とする、という方法を用いることができる。

【 0 1 3 0 】

以降の信号変形部 7、重み付き加算部 18 の動作は、実施の形態 3 と同様であり、説明を省略する。

【 0 1 3 1 】

なお、ここでは、レベル判定部 23 と連続性判定部 24 の両方の出力結果を使用したが、一方だけを使用するようになり、両方とも使用しない構成も可能である。また、変形強度によって制御する対象を、振幅平滑化部 9 と位相擾乱部 10 の一方のみとしたり、第

10

20

30

40

50

三の変形強度については一方のみを制御対象とする構成でも構わない。

【0132】

この実施の形態5によれば、実施の形態3の構成に加えて、平滑化強度または擾乱付与強度を、所定の評価値（背景雑音らしさ）の時間変動性（フレームまたはサブフレーム間の変動性）の大きさによって制御するようにしたので、実施の形態3が持つ効果に加えて、入力信号（復号音声）の特性が変動している区間において必要以上に強い加工処理を抑制でき、なまけ、エコー（反響感）の発生を防止できる効果がある。

【0133】

実施の形態6 .

図5との対応部分に同一符号を付けた図9は、本実施の形態による音信号加工方法を適用した音声復号装置の全体構成を示す。図中27は摩擦音らしさ評価部、31は背景雑音らしさ評価部、45は加算制御値算出部である。摩擦音らしさ評価部27は、低域カットフィルタ28、零交差数カウント部29、摩擦音らしさ算出部30より構成される。背景雑音らしさ評価部31は、図5における信号評価部12と同じ構成であり、逆フィルタ部13、パワー算出部14、背景雑音らしさ算出部15、推定雑音パワー更新部16、推定雑音スペクトル更新部17より構成される。信号評価部12は、図5の場合と異なり、摩擦音らしさ評価部27、背景雑音らしさ評価部31、加算制御値算出部45より構成される。

10

【0134】

以下、図に基づいて動作を説明する。

20

【0135】

音声復号部4から出力された復号音声5が、信号加工部2内の信号変形部7、変形強度制御部20、信号評価部12内の摩擦音らしさ評価部27と背景雑音らしさ評価部31、そして重み付き加算部18に入力される。

【0136】

信号評価部12内の背景雑音らしさ評価部31は、実施の形態3における信号評価部12と同様に、入力された復号音声5に対して、逆フィルタ部13、パワー算出部14、背景雑音らしさ算出部15の処理を行って、得られた背景雑音らしさ46を加算制御値算出部45に出力する。また、推定雑音パワー更新部16、推定雑音スペクトル更新部17の処理を行って、各々に格納してある推定雑音パワーと推定雑音スペクトルの更新を行う。

30

【0137】

摩擦音らしさ評価部27内の低域カットフィルタ28は、入力された復号音声5に対して低周波数成分を抑圧する低域カットフィルタリング処理を行い、フィルタリング後の復号音声を零交差数カウント部29に出力する。この低域カットフィルタリング処理の目的は、復号音声に含まれる直流成分や低周波数の成分がオフセットとなって、後述する零交差数カウント部29のカウント結果が少なくなることを防止することである。従って、単純には、フレーム内の復号音声5の平均値を算出し、これを復号音声5の各サンプルから減算することでもよい。

【0138】

零交差数カウント部29は、低域カットフィルタ28より入力された音声を分析して、含まれる零交差数を数え上げ、得られた零交差数を摩擦音らしさ算出部30に出力する。零交差数を数え上げる方法としては、隣接サンプルの正負を比較し、同一でなければ零を交差している、としてカウントする方法、隣接サンプルの値の積をとって、その結果が負または零であれば零を交差している、としてカウントする方法などがある。

40

【0139】

摩擦音らしさ算出部30は、零交差数カウント部29より入力された零交差数を、所定の閾値と比較し、この比較結果に基づいて摩擦音らしさ47を求めて、これを加算制御値算出部45に出力する。例えば、零交差数が閾値より大きい場合には、摩擦音らしいと判定して摩擦音らしさを1に設定する。逆に零交差数が閾値より小さい場合には、摩擦音らしくないと判定して摩擦音らしさを0に設定する。この他、閾値を2つ以上設けて、摩擦

50

音らしさを段階的に設定したり、所定の関数を用意しておいて、零交差数から連続的な値の摩擦音らしさを算出するようにしても良い。

【 0 1 4 0 】

なお、この摩擦音らしさ評価部 2 7 内の構成は、あくまでも一例にすぎず、スペクトル傾斜の分析結果に基づいて評価するようにしたり、パワーやスペクトルの定常性に基づいて評価するようにしたり、零交差数も含めて複数のパラメータを組み合わせて評価するようにしたりしても構わない。

【 0 1 4 1 】

加算制御値算出部 4 5 は、背景雑音らしさ評価部 3 1 より入力された背景雑音らしさ 4 6 と、摩擦音らしさ評価部 2 7 より入力された摩擦音らしさ 4 7 に基づいて、加算制御値 3 5 を算出し、これを重み付き加算部 1 8 に出力する。背景雑音らしい場合と摩擦音らしい場合のどちらにおいても、量子化雑音が聞き苦しくなってしまうことが多いので、背景雑音らしさ 4 6 と摩擦音らしさ 4 7 を適切に重み付き加算することで加算制御値 3 5 を算出すればよい。

10

【 0 1 4 2 】

以降の信号変形部 7、変形強度制御部 2 0、重み付き加算部 1 8 の動作は、実施の形態 3 と同様であり、説明を省略する。

【 0 1 4 3 】

この実施の形態 6 によれば、入力信号（復号音声）の背景雑音らしさと摩擦音らしさが高い場合に、入力信号（復号音声）の代わりに加工信号（変形復号音声）をより大きく出力するようにしたので、実施の形態 3 が持つ効果に加えて、量子化雑音や劣化成分が多く発生しがちな摩擦音区間に対して重点的な加工が加えられ、摩擦音以外の区間についてもその区間に適切な加工（加工しない、低レベルの加工を行うなど）が選択されるので、主観品質を改善できる効果がある。なお、摩擦音らしさ以外にも、量子化雑音や劣化成分が多く発生しがちな部分がある程度特定できる場合には、その部分らしさを評価して、加算制御値に反映させることが可能である。その様に構成すれば、大きい量子化雑音や劣化成分を 1 つずつ抑圧していくことができるので、主観品質が一層改善できる効果がある。

20

【 0 1 4 4 】

また、当然のことであるが、背景雑音らしさ評価部を削除した構成も可能である。

【 0 1 4 5 】

実施の形態 7 .

図 1 との対応部分に同一符号を付けた図 1 0 は、本実施の形態による信号加工方法を適用した音声復号装置の全体構成を示し、図中 3 2 はポストフィルタ部である。

30

【 0 1 4 6 】

以下、図に基づいて動作を説明する。

【 0 1 4 7 】

まず音声符号 3 が音声復号装置 1 内の音声復号部 4 に入力される。

【 0 1 4 8 】

音声復号部 4 は、入力された音声符号 3 に対して復号処理を行い、得られた復号音声 5 をポストフィルタ部 3 2、信号変形部 7、信号評価部 1 2 に出力する。

40

【 0 1 4 9 】

ポストフィルタ部 3 2 は、入力された復号音声 5 に対して、スペクトル強調処理、ピッチ周期性強調処理などを行い、得られた結果をポストフィルタ復号音声 4 8 として重み付き加算部 1 8 に出力する。このポストフィルタ処理は、C E L P 復号処理の後処理として一般的に使用されているもので、符号化復号化によって発生した量子化雑音を抑圧することを目的として導入されている。スペクトル強度の弱い部分には量子化雑音が多く含まれているので、この成分の振幅を抑圧してしまうものである。なお、ピッチ周期性強調処理が行われず、スペクトル強調処理だけが行われている場合もある。

【 0 1 5 0 】

なお、実施の形態 1、実施の形態 3 ないし 6 は、このポストフィルタ処理を音声復号部

50

4 内に含まれるもの、もしくは存在しないものの何れにも適用可能なものについて説明したが、この実施の形態 7 では、音声復号部 4 内にポストフィルタ処理が含まれるものからポストフィルタ処理の全部もしくは一部をポストフィルタ部 3 2 として独立させている。

【 0 1 5 1 】

信号変形部 7 は、実施の形態 1 と同様に、入力された復号音声 5 に対して、フーリエ変換部 8、振幅平滑化部 9、位相擾乱部 10、逆フーリエ変換部 11 の処理を行い、得られた変形復号音声 3 4 を重み付き加算部 1 8 に出力する。

【 0 1 5 2 】

信号評価部 1 2 は、実施の形態 1 と同様に、入力された復号音声 5 に対して、背景雑音らしさを評価し、評価結果を加算制御値 3 5 として重み付き加算部 1 8 に出力する。

10

【 0 1 5 3 】

そして、最後の処理として、重み付き加算部 1 8 は、実施の形態 1 と同様に、信号評価部 1 2 から入力された加算制御値 3 5 に基づいて、ポストフィルタ部 3 2 から入力されたポストフィルタ復号音声 4 8 と信号変形部 7 から入力された変形復号音声 3 4 を重み付け加算し、得られた出力音声 6 を出力する。

【 0 1 5 4 】

この実施の形態 7 によれば、ポストフィルタによる加工前の復号音声に基づいて変形復号音声を生成し、更にポストフィルタによる加工前の復号音声进行分析して背景雑音らしさを求め、これに基づいてポストフィルタ復号音声と変形復号音声の加算時の重みを制御するようにしたので、実施の形態 1 が持つ効果に加えて、ポストフィルタによる復号音声の変形を含まない変形復号音声の生成ができ、ポストフィルタによる復号音声の変形に影響されずに算出した精度の高い背景雑音らしさに基づいて精度の高い加算重み制御ができるようになるので、更に主観品質が改善する効果がある。

20

【 0 1 5 5 】

背景雑音区間においては、ポストフィルタによって劣化音までも強調されて聞き苦しくなってしまうことが多く、ポストフィルタによる加工前の復号音声を出発点として変形復号音声を生じた方が、歪み音は小さくなる。また、ポストフィルタの処理が複数のモードを持っており、しばしば処理を切り替える場合には、その切り替えが背景雑音らしさの評価に影響する危険性が高く、ポストフィルタによる加工前の復号音声に対して背景雑音らしさを評価した方が安定な評価結果が得られる。

30

【 0 1 5 6 】

なお、実施の形態 3 の構成において、この実施の形態 7 と同様にポストフィルタ部の分離を行った場合には、図 5 の聴覚重み付け部 2 1 の出力結果が、より符号化処理内の聴覚重み付け音声に近づき、量子化雑音の多い成分の特定精度が上がり、より良い変形強度制御が得られ、主観品質が更に改善する効果が得られる。

【 0 1 5 7 】

また、実施の形態 6 の構成において、この実施の形態 7 と同様にポストフィルタ部の分離を行った場合には、図 9 の摩擦音らしさ評価部 2 7 における評価精度が上がり、主観品質が更に改善する効果が得られる。

【 0 1 5 8 】

40

なお、ポストフィルタ部の分離を行わない構成は、分離したこの実施の形態 7 の構成に比べると、音声復号部（ポストフィルタを含む）との接続が復号音声の 1 点だけと少なく、独立の装置、プログラムにて実現が容易である長所がある。この実施の形態 7 では、ポストフィルタを有する音声復号部に対して独立の装置、プログラムにて実現することが容易でない短所もあるが、上記の様々な効果を持つものである。

【 0 1 5 9 】

実施の形態 8 .

図 1 0 との対応部分に同一符号を付けた図 1 1 は、本実施の形態による音信号加工方法を適用した音声復号装置の全体構成を示し、図中 3 3 は音声復号部 4 内で生成されたスペクトルパラメータである。図 1 0 との相違点としては、実施の形態 3 と同様の変形強度制

50

御部 20 が追加され、スペクトルパラメータ 33 が音声復号部 4 から信号評価部 12 と変形強度制御部 20 に入力されている点である。

【0160】

以下、図に基づいて動作を説明する。

【0161】

まず音声符号 3 が音声復号装置 1 内の音声復号部 4 に入力される。

【0162】

音声復号部 4 は、入力された音声符号 3 に対して復号処理を行い、得られた復号音声 5 をポストフィルタ部 32、信号変形部 7、変形強度制御部 20、信号評価部 12 に出力する。また、復号処理の過程で生成したスペクトルパラメータ 33 を、信号評価部 12 内の推定雑音スペクトル更新部 17 と変形強度制御部 20 内の聴覚重み付け部 21 に出力する。なお、スペクトルパラメータ 33 としては、線形予測係数 (LPC)、線スペクトル対 (LSP) などが一般的に用いられていることが多い。

10

【0163】

変形強度制御部 20 内の聴覚重み付け部 21 は、音声復号部 4 より入力された復号音声 5 に対して、やはり音声復号部 4 から入力されたスペクトルパラメータ 33 を用いて聴覚重み付け処理を行い、得られた聴覚重み付け音声をフーリエ変換部 22 に出力する。具体的な処理としては、スペクトルパラメータ 33 が線形予測係数 (LPC) である場合にはこれをそのまま用い、スペクトルパラメータ 33 が LPC 以外のパラメータである場合には、このスペクトルパラメータ 33 を LPC に変換して、この LPC に定数乗算を行って 2 つの変形 LPC を求め、この 2 つの変形 LPC をフィルタ係数とする ARMA フィルタを構成し、このフィルタを用いたフィルタリング処理によって聴覚重み付けを行う。なお、この聴覚重み付け処理は、音声符号化処理 (音声復号部 4 で行った音声復号処理と対を成すもの) で使用されているものと同様な処理を行うことが望ましい。

20

【0164】

変形強度制御部 20 内では、上記聴覚重み付け部 21 の処理に続いて、実施の形態 3 と同様に、フーリエ変換部 22、レベル判定部 23、連続性判定部 24、変形強度算出部 25 の処理を行い、得られた変形強度を信号変形部 7 に対して出力する。

【0165】

信号変形部 7 は、実施の形態 3 と同様に、入力された復号音声 5 と変形強度に対して、フーリエ変換部 8、振幅平滑化部 9、位相擾乱部 10、逆フーリエ変換部 11 の処理を行い、得られた変形復号音声 34 を重み付き加算部 18 に出力する。

30

【0166】

信号評価部 12 内では、実施の形態 1 と同様に、入力された復号音声 5 に対して、まず逆フィルタ部 13、パワー算出部 14、背景雑音らしさ算出部 15 の処理を行って背景雑音らしさを評価し、評価結果を加算制御値 35 として重み付き加算部 18 に出力する。また、推定雑音パワー更新部 16 の処理を行って、内部の推定雑音パワーを更新する。

【0167】

そして、推定雑音スペクトル更新部 17 は、音声復号部 4 から入力されたスペクトルパラメータ 33 と背景雑音らしさ算出部 15 から入力され背景雑音を用いて、その内部に格納してある推定雑音スペクトルを更新する。例えば、入力された背景雑音らしさが高い時に、実施の形態 1 に示した式に従い、スペクトルパラメータ 33 を推定雑音スペクトルに反映させることで更新を行う。

40

以降のポストフィルタ部 32、重み付き加算部 18 の動作については、実施の形態 7 と同様であるため、説明を省略する。

【0168】

この実施の形態 8 によれば、音声復号処理の過程で生成されたスペクトルパラメータを流用して、聴覚重み付け処理、推定雑音スペクトルの更新を行うようにしたので、実施の形態 3 及び実施の形態 7 が持つ効果に加えて、処理が簡易化される効果がある。

【0169】

50

更に、符号化処理とまったく同じ聴覚重み付け処理が実現され、量子化雑音の多い成分の特定精度が上がり、より良い変形強度制御が得られ、主観品質が改善する効果が得られる。

【 0 1 7 0 】

また、背景雑音らしさの算出に用いる推定雑音スペクトルの（音声符号化処理に入力された音声のスペクトルに近いという意味での）推定精度が上がり、結果として得られる安定した高精度の背景雑音らしさに基づいて精度の高い加算重み制御ができるようになり、主観品質が改善する効果がある。

【 0 1 7 1 】

なお、この実施の形態 8 では、ポストフィルタ部 3 2 を音声復号部 4 から分離した構成であったが、分離していない構成においても、実施の形態 8 のように音声復号部 4 が出力したスペクトルパラメータ 3 3 を流用して信号加工部 2 の処理を行うことができる。この場合でも、上記実施の形態 8 と同様の効果が得られる。

10

【 0 1 7 2 】

実施の形態 9 .

上記図 7 に示す実施の形態 4 の構成において、加算制御値分割部 4 1 が、重み付け加算部 1 8 にて加算される変形復号音声スペクトル 4 4 の周波数毎の重みを乗じた後のスペクトルの概形が、量子化雑音の推定スペクトル形状に一致するように、出力する変形強度を制御することも可能である。

【 0 1 7 3 】

20

図 1 2 は、この場合の復号音声スペクトル 4 3 と、変形復号音声スペクトル 4 4 に周波数毎の重みを乗じた後のスペクトルの一例を示す模式図である。

【 0 1 7 4 】

復号音声スペクトル 4 3 には、符号化方式に依存したスペクトル形状を持つ量子化雑音が重畳している。C E L P 系の音声符号化方式においては、聴覚重み付け処理後の音声における歪みを最小化するように符号の探索を行う。このため、量子化雑音は、聴覚重み付け処理後の音声においては、平坦なスペクトル形状を持つことになり、最終的な量子化雑音のスペクトル形状は、聴覚重み付け処理の逆特性のスペクトル形状を持つことになる。よって、聴覚重み付け処理のスペクトル特性を求め、この逆特性のスペクトル形状を求めて、変形復号音声スペクトルのスペクトル形状がこれに合うように、加算制御値分割部 4 1 の出力を制御することは可能である。

30

【 0 1 7 5 】

この実施の形態 9 によれば、最終的な出力音声 6 に含まれる変形復号音声成分のスペクトル形状を量子化雑音の推定スペクトルの概形に一致するようにしたので、実施の形態 4 が持つ効果に加えて、必要最低限のパワーの変形復号音声の加算によって音声区間における聞き苦しい量子化雑音を聞こえにくくすることができる効果がある。

【 0 1 7 6 】

実施の形態 1 0 .

上記実施の形態 1、実施の形態 3 ないし 8 の構成において、振幅平滑化部 9 の処理内で、平滑化後の振幅スペクトルが推定量子化雑音の振幅スペクトル形状に一致するように加工することも可能である。なお、推定量子化雑音の振幅スペクトル形状の算出は、実施の形態 9 と同様にして行えばよい。

40

【 0 1 7 7 】

この実施の形態 1 0 によれば、変形復号音声のスペクトル形状を量子化雑音の推定スペクトル形状に一致するようにしたので、実施の形態 1、実施の形態 3 ないし 8 が持つ効果に加えて、必要最低限のパワーの変形復号音声の加算によって音声区間における聞き苦しい量子化雑音を聞こえにくくすることができる効果がある。

【 0 1 7 8 】

実施の形態 1 1 .

上記実施の形態 1、実施の形態 3 ないし 1 0 では、信号加工部 2 を復号音声 5 の加工に

50

使用しているが、この信号加工部 2 のみを取り出して、音響信号復号部（音響信号符号化に対する復号部）、雑音抑圧処理の後段に接続するなど、他の信号加工処理に使用することもできる。但し、解消したい劣化成分の特性に応じて、信号変形部における変形処理、信号評価部における評価方法を変更、調整することが必要になる。

【0179】

この実施の形態 1 1 によれば、復号音声以外の劣化成分を含む信号に対して、主観的に好ましくない成分を感じにくく加工することが可能である。

【0180】

実施の形態 1 2 .

上記実施の形態 1 ないし 1 1 では、現在のフレームまでの信号を用いて該信号の加工を行っているが、処理遅延の発生を許して次フレーム以降の信号も使用する構成も可能である。

10

【0181】

この実施の形態 1 2 によれば、次のフレーム以降の信号を参照できるので、振幅スペクトルの平滑化特性の改善、連続性判定の精度向上、雑音らしさなどの評価精度の向上効果が得られる。

【0182】

実施の形態 1 3 .

上記実施の形態 1、実施の形態 3、実施の形態 5 ないし 1 2 では、フーリエ変換によってスペクトル成分を算出し、変形処理を行って、逆フーリエ変換によって信号領域に戻しているが、フーリエ変換の代わりにバンドパスフィルタ群の各出力に対して、変形処理を行い、帯域別信号の加算によって信号を再構築する構成も可能である。

20

【0183】

この実施の形態 1 3 によれば、フーリエ変換を使用しない構成でも同様の効果が得られる。

【0184】

実施の形態 1 4 .

上記実施の形態 1 ないし 1 3 では、振幅平滑化部 9 と位相擾乱部 1 0 の両方を備えた構成であったが、振幅平滑化部 9 と位相擾乱部 1 0 の一方を省略した構成も可能であるし、更に別の変形部を導入した構成も可能である。

30

【0185】

この実施の形態 1 4 によれば、解消したい量子化雑音や劣化音の特性によっては、導入効果がない変形部を省略することで処理が簡易化できる効果がある。また、適切な変形部を導入することで、振幅平滑化部 9 と位相擾乱部 1 0 では解消できない量子化雑音や劣化音を解消できる効果が期待できる。

【図面の簡単な説明】

【0186】

【図 1】この発明の実施の形態 1 による音声復号方法を適用した音声復号装置の全体構成を示す図である。

【図 2】この発明の実施の形態 1 の重み付け加算部 1 8 における加算制御値に基づく重み付け加算の制御例を示す図である。

40

【図 3】この発明の実施の形態 1 のフーリエ変換部 8 における切り出し窓、逆フーリエ変換部 1 1 における接続のための窓の実際の形状例、復号音声 5 との時間関係を説明する説明図である。

【図 4】この発明の実施の形態 2 の音信号加工方法を雑音抑圧方法と組み合わせて適用した音声復号装置の構成の一部を示す図である。

【図 5】この発明の実施の形態 3 による音声復号方法を適用した音声復号装置の全体構成を示す図である。

【図 6】この発明の実施の形態 3 の聴覚重み付けスペクトルと第一の変形強度の関係を示す図である。

50

【図7】この発明の実施の形態4による音声復号方法を適用した音声復号装置の全体構成を示す図である。

【図8】この発明の実施の形態5による音声復号方法を適用した音声復号装置の全体構成を示す図である。

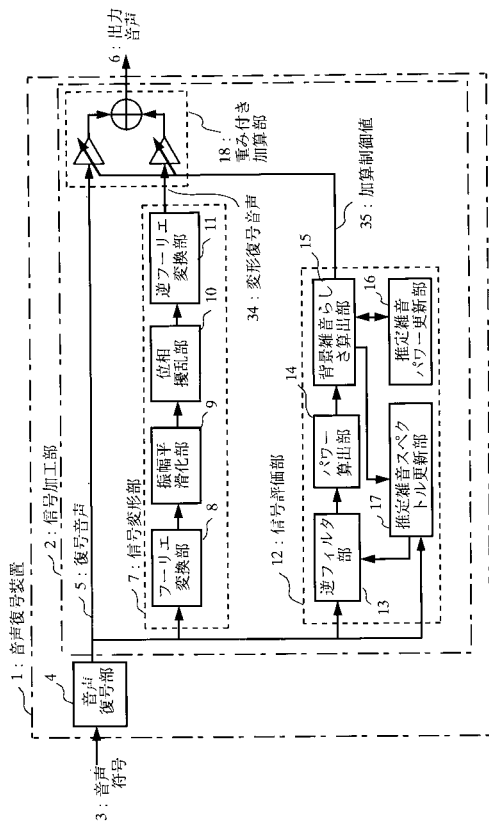
【図9】この発明の実施の形態6による音声復号方法を適用した音声復号装置の全体構成を示す図である。

【図10】この発明の実施の形態7による音声復号方法を適用した音声復号装置の全体構成を示す図である。

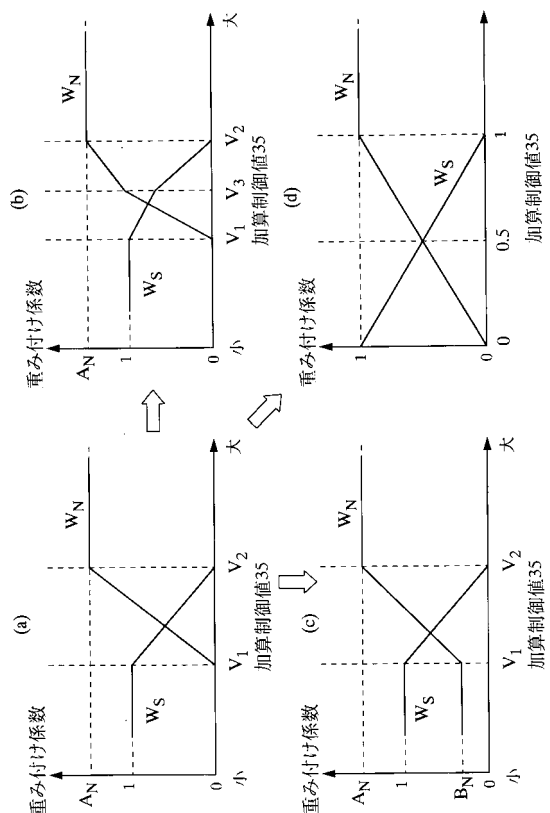
【図11】この発明の実施の形態8による音声復号方法を適用した音声復号装置の全体構成を示す図である。

【図12】この発明の実施の形態9を適用した復号音声スペクトル43と、変形復号音声スペクトル44に周波数毎の重みを乗じた後のスペクトルの一例を示す模式図である。

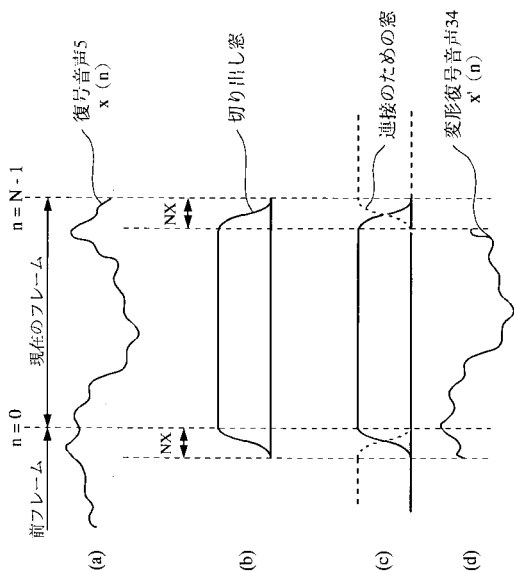
【図1】



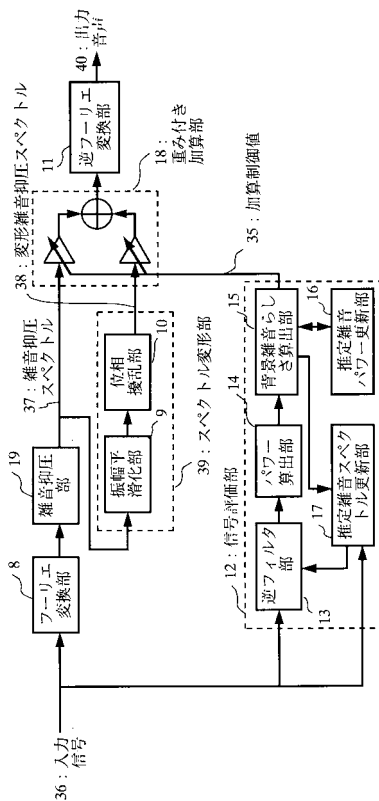
【図2】



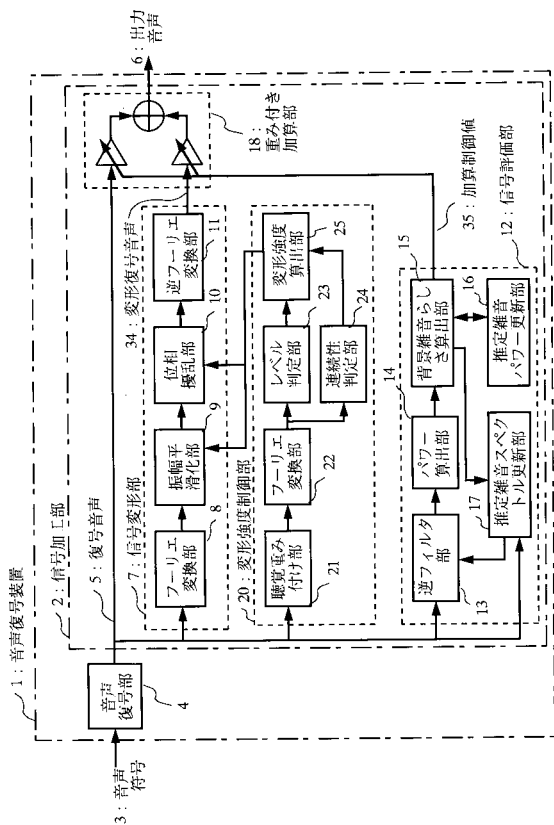
【図3】



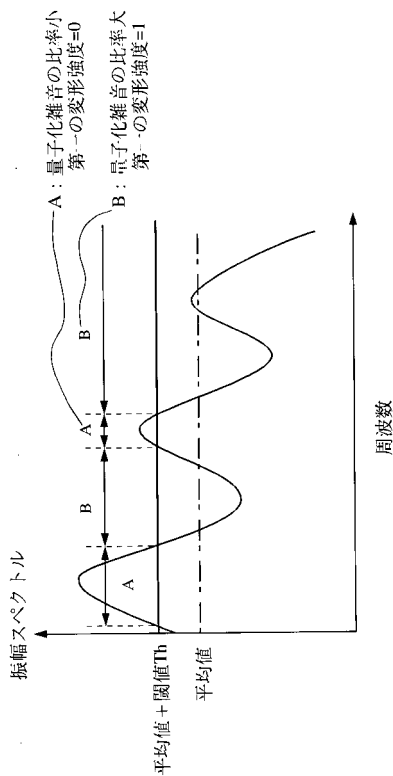
【図4】



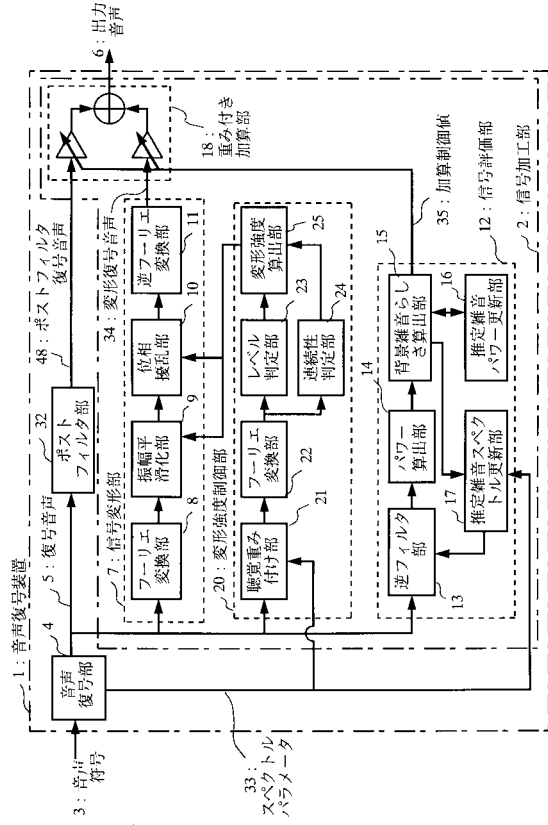
【図5】



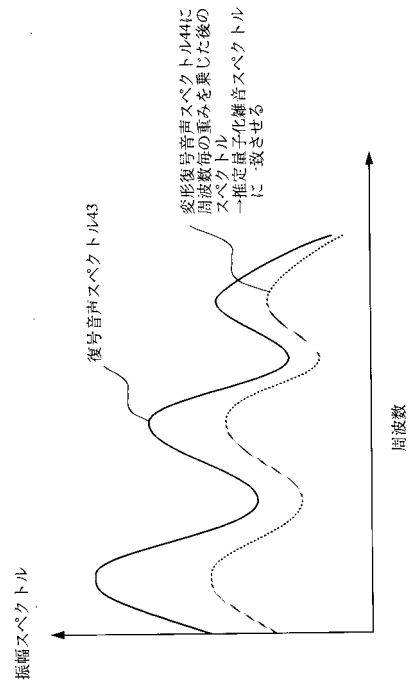
【図6】



【図11】



【図12】



フロントページの続き

(56)参考文献 特開平05 - 224698 (JP, A)
特開平09 - 054600 (JP, A)
特開平08 - 248997 (JP, A)
特開平06 - 051791 (JP, A)

(58)調査した分野(Int.Cl., DB名)
G10L 11/00 - 21/06
JSTPlus (JDreamII)