明　細　書

発明の名称 ：

音声信号処理装置、音声信号処理方法、及び音声信号処理プログラム

技術分野

[0001] 本発明は、音声信号を処理する音声信号処理装置、音声信号処理方法、及び音声信号処理プログラムに関する。

背景技術

[0002] ＩＰ（Internet Protocol）電話のように、音声信号を符号化及びパケット化してインターネット網で伝送する際には、ネットワークの輻輳等が原因でパケットが失われることがある（以下、この現象を「パケットロス」と言う）。パケットロスが発生すると、必要な音声符号が失われるため音声復号ができず音切れが発生する。パケットロスにより生じる音切れを防止する技術として、音声パケットロス隠蔽技術がある。音声パケットロス隠蔽技術は、パケットロスを検出し、ロスしたパケットに対応する擬似的な音声信号（以下、「隠蔽信号」）を生成する。

[0003] 音声符号化手法として、符号化器・復号器の内部状態を更新しながら音声符号化を行う手法を用いている場合には、本来受信するはずの符号化パラメータが得られないため、音声パケットロス隠蔽技術は、擬似的に生成したパラメータにより復号器の内部状態更新も行う。

[0004] 符号化器・復号器の内部状態を更新しながら音声符号化を行う手法として、CELP（Code Excited Linear Prediction）符号化が広く用いられている。CELP符号化では、自己回帰モデルを仮定し、全極型合成フィルタa(i)によって励振信号e(n)をフィルタリングすることで音声信号を合成する。すなわち、次式に従い音声信号s(n)を合成する。a(i)は線形予測係数（LP（Linear Prediction）係数）であり、次数としてP=16などの値を用いる。

[数1]

$$s(n) = e(n) - \sum_{i=1}^{P} a(i) \cdot s(n-i)$$

[0005] CELP符号化では、線形予測係数を数学的に等価に表現したISF（Immittance Spectral Frequency）パラメータや、過去の励振信号を内部状態としてもつ。パケットロスが起こった場合、これらを擬似的に生成するため、本来であれば復号により得られていたパラメータとの間で乖離が起こる。パラメータの乖離によって起こる合成音声の不整合は、受聴者からは雑音と知覚され、主観的な品質を大きく損ねる。

[0006] 以下、音声符号化手法にCELP符号化を用いた場合を例に、音声パケットロス隠蔽を行う音声復号器の構成及び動作について説明する。

[0007] 音声復号器の構成図及び動作を図１、図２に示す。図１に示すように、音声復号器１は、パケットロス検出部１１、音声符号復号部１２、隠蔽信号生成部１３、及び内部状態バッファ１４を備える。

[0008] パケットロス検出部１１は、音声パケットを正常に受信した場合には、制御信号、及び音声パケットに含まれる音声符号を音声符号復号部１２に送る（正常受信：図２のステップＳ１００でＹＥＳの場合）。その後、音声符号復号部１２は、後述するように、音声符号の復号及び内部状態更新を行う（図２のステップＳ２００、Ｓ４００）。一方、パケットロス検出部１１は、音声パケットが正常に受信できなかった場合には、隠蔽信号生成部１３に制御信号を送る（パケットロス：図２のステップＳ１００でＮＯの場合）。その後、隠蔽信号生成部１３は、後述するように、隠蔽信号の生成及び内部状態更新を行う（図２のステップＳ３００、Ｓ４００）。図２のステップＳ１００～Ｓ４００の処理は通信終了まで（ステップＳ５００でＹＥＳと判断されるまで）繰り返される。

[0009] 音声符号は、少なくとも符号化されたISFパラメータ

[数2]

$\dot{\omega}_i$

、符号化された第一から第四サブフレームのピッチラグT^j_p、第一から第四サブフレームの符号化された適応符号帳ゲインg^j_p、第一から第四サブフレームの符号化された固定符号帳ゲインg^j_c、第一から第四サブフレームの符号化さ

れた固定符号帳ベクトル$c^j(n)$を含む。ISFパラメータの代わりに数学的に等価な表現であるLSF（line spectral frequency）パラメータを用いてもよい。以下の議論では、ISFパラメータを用いた説明を行うが、LSFパラメータを用いた場合も同じ議論が成り立つ。

[0010] 内部状態バッファには、過去のISFパラメータ

[数3]

$\dot{\omega}_i^{-1}$

及び

[数4]

$\dot{\omega}_i^{-1}$

の等価表現であるISP（Immittance Spectral Pair）パラメータ

[数5]

$\dot{\mathbf{q}}_i^{-1}$

、ISF残差パラメータ

[数6]

$\dot{r}_i^{-1}$

、過去のピッチラグ$T^j{}_p$、過去の適応符号帳ゲイン$g^j{}_p$、過去の固定符号帳ゲイン$g^j{}_c$、適応符号帳$u(n)$を含む。それぞれ過去何サブフレーム分のパラメータを含むかは設計方針による。本明細書では、１フレームが４サブフレームを含むことを仮定するが、設計方針により他の値としてもよい。

[0011] ＜正常受信の場合＞

図３には音声符号復号部１２の機能構成例を示す。この図３に示すように、音声符号復号部１２は、ISF復号部１２０、安定性処理部１２１、LP係数算出部１２２、適応符号帳算出部１２３、固定符号帳復号部１２４、ゲイン復号部１２５、励振ベクトル合成部１２６、ポストフィルタ１２７、及び合成フィルタ１２８を備える。ただし、ポストフィルタ１２７は必須の構成要素ではない。なお、図３では、説明の便宜上、音声符号復号部１２内に内部状

態バッファ１４を二点鎖線で示しているが、この内部状態バッファ１４は、音声符号復号部１２の内部に含まれるものではなく、図１に示す内部状態バッファ１４である。これ以降の音声符号復号部の構成図でも同様である。

[0012] LP係数算出部１２２の構成図を図４に、符号化されたISFパラメータからLP係数を算出する処理フローを図５に、それぞれ示す。図４に示すように、LP係数算出部１２２は、ISF-ISP変換部１２２Ａ、ISP補間部１２２Ｂ、及びISP-LPC変換部１２２Ｃを備える。

[0013] まず、符号化されたISFパラメータからLP係数を算出する処理（図５）に関連する機能構成及び動作について説明する。

[0014] ISF復号部１２０は、符号化されたISFパラメータを復号してISF残差パラメータ

[数7]

$\dot{r}_i^0$

を求め、ISFパラメータ

[数8]

$\dot{\omega}_i$

を次式に従い算出する（図５のステップＳ１）。ここで、$mean_i$は、事前に学習等で求めた平均ベクトルである。

[数9]

$$\dot{\omega}_i = mean_i + \dot{r}_i^0 + \frac{1}{3}\dot{r}_i^{-1}$$

[0015] なお、ここでは、ISFパラメータの算出にMA予測を用いる例について述べたが、以下のようにAR予測を用いてISFパラメータの算出を行うような構成としてもよい。ここで、直前フレームのISFパラメータを

[数10]

$\dot{\omega}_i^{-1}$

、AR予測の重み係数をρ_iとした。

[数11]

$$\dot{\omega}_i = mean_i + \rho_i(\dot{\omega}_i^{-1} - mean_i)$$

[0016] 安定性処理部１２１は、フィルタの安定性を確保するためにISFパラメータの各要素間に50Hz以上の間隔をあけるよう、次式に従う処理を行う（図５のステップＳ２）。ISFパラメータは、音声スペクトル包絡の形状を線スペクトルで表現したものであり、互いの距離が近づく程、スペクトルのピークが大きくなり共振が起こる。そのため、スペクトルのピークでのゲインが大きくなりすぎないよう、安定性確保の処理が必要となる。ここで、min_distは最小のISF間隔であり、isf_minは、min_distの間隔を確保するために必要なISFの最小値である。isf_minは、隣のISFの値にmin_distの間隔を加算することにより順次更新を行う。一方、isf_maxは、min_distの間隔を確保するために必要なISFの最大値である。isf_maxは、隣のISFの値からmin_distの間隔を減算することにより順次更新を行う。

[数12]

```
isf_min = min_dist = 50
for i=0 to 14
        if ω̇i<isf_min then ω̇i=isf_min
        isf_min = ω̇i + min_dist
isf_max = 6400 - min_dist
if ω̇14>isf_max
        for i=14 down to 1
                if ω̇i > isf_max then ω̇i=isf_max
                isf_max = ω̇i - min_dist
```

[0017] LP係数算出部１２２内のISF-ISP変換部１２２Ａは、次式に従い

[数13]

$\dot{\omega}_i$

をISPパラメータ

[数14]

$\dot{\mathbf{q}}_i$

に変換する（図５のステップＳ３）。ここで、Cは事前に定めた定数である。

[数15]

$\dot{\mathbf{q}}_i = \cos(C \cdot \dot{\omega}_i)$

[0018] ISP補間部１２２Ｂは、内部状態バッファ１４に含まれる過去のISPパラメータ

[数16]

$\dot{\mathbf{q}}_i^{-1}$

と、上記ISPパラメータ

[数17]

$\dot{\mathbf{q}}_i$

から、以下の式に従いサブフレーム毎のISPパラメータを算出する（図５のステップＳ４）。補間にあたっては、別の係数を用いてもよい。

[数18]

$$\mathbf{q}_i^{(1)} = 0.75 \cdot \dot{\mathbf{q}}_i^{-1} + 0.25 \cdot \dot{\mathbf{q}}_i$$
$$\mathbf{q}_i^{(2)} = 0.5 \cdot \dot{\mathbf{q}}_i^{-1} + 0.5 \cdot \dot{\mathbf{q}}_i$$
$$\mathbf{q}_i^{(3)} = 0.25 \cdot \dot{\mathbf{q}}_i^{-1} + 0.75 \cdot \dot{\mathbf{q}}_i$$
$$\mathbf{q}_i^{(4)} = \dot{\mathbf{q}}_i$$

[0019] ISP-LPC変換部１２２Ｃは、サブフレーム毎のISPパラメータをLP係数

[数19]

$\dot{a}_i^j \quad (0 < i \le P、\ 0 \le j < 4)$

に変換する（図５のステップＳ５）。具体的変換手順として、非特許文献１に記載の処理手順を用いることができる。ここで、先読み信号に含まれるサブフレームの数を４としたが、サブフレームの数は設計方針により変更してもよい。

[0020] 次に、音声符号復号部１２におけるその他の構成及び動作について説明する。

[0021] 適応符号帳算出部１２３は、符号化されたピッチラグを復号して、第一から第四サブフレームのピッチラグT^j_pを算出する。次に、適応符号帳算出部１２３は、適応符号帳u(n)を用いて次式に従い、サブフレーム毎に適応符号帳ベクトルを算出する。適応符号帳ベクトルは、適応符号帳u(n)をFIRフィルタInt(i)により補間することで算出する。ここで、適応符号帳の長さをN_{adapt}とした。補間に用いるフィルタInt(i)は、事前に定めた長さ2l+1のFIRフィルタであり、L’はサブフレームのサンプル数である。補間フィルタInt(i)を用いることにより、ピッチラグを小数点以下の精度まで利用することができる。補間フィルタの詳細については、非特許文献１記載の方法を用いることができる。

[数20]

$$v^j(n)=\sum_{i=-l}^{l} Int(i)\cdot u(n+N_{adapt}-\hat{T}_p^{(j)}+i) \qquad (0\le n<L')$$

[0022] 固定符号帳復号部１２４は、符号化された固定符号帳ベクトルを復号して、第一から第四サブフレームの固定符号帳ベクトル$c^j(n)$を取得する。

[0023] ゲイン復号部１２５は、符号化された適応符号帳ゲイン及び符号化された固定符号帳ゲインを復号して、第一から第四サブフレームの適応符号帳ゲイン及び固定符号帳ゲインを取得する。例えば、非特許文献１に記載の以下の手法により、適応符号帳ゲイン及び固定符号帳ゲインの復号を行うことができる。非特許文献１記載の以下の手法によれば、AMR-WBのゲイン符号化のようにフレーム間予測を用いないため、パケットロス耐性を高めることができる。

[0024] 例えば、ゲイン復号部１２５は、以下の処理フローに従い、固定符号帳ゲインを取得する。

[0025] 最初に、ゲイン復号部１２５は、固定符号帳ベクトルのパワーを算出する。ここで、サブフレームの長さをN_sとする。

[数21]

$$E_c=10\log\left(\frac{1}{N_s}\sum_{i=0}^{N_s-1}c^2(i)\right)$$

[0026] 次に、ゲイン復号部１２５は、ベクトル量子化されたゲインパラメータを復号し、適応符号帳ゲイン

[数22]

$\hat{g}_p$

と量子化固定符号帳ゲイン

[数23]

$\hat{E}_i$

を得る。量子化固定符号帳ゲインと上記固定符号帳ベクトルのパワーから、以下の通り予測固定符号帳ゲインを算出する。

[数24]

$$g_c' = 10^{0.05\left(\hat{E}_i - E_c\right)}$$

[0027] 最後に、ゲイン復号部１２５は、予測係数

[数25]

$\hat{\gamma}$

を復号し、予測ゲインに乗算することにより、固定符号帳ゲインを得る。

[数26]

$$\hat{g}_c = \hat{\gamma} \cdot g_c'$$

[0028] 励振ベクトル合成部１２６は、次式のように、適応符号帳ベクトルに適応符号帳ゲインを乗算するとともに、固定符号帳ベクトルに固定符号帳ゲインを乗算し、これらの和を求めることにより、励振信号を取得する。

[数27]

$$e^j(n) = g_p^j \cdot v^j(n) + g_c^j \cdot c^j(n)$$

[0029] ポストフィルタ１２７は、励振信号ベクトルに対して、例えば、ピッチ強調、ノイズ強調、低域強調といった後処理を加える。ピッチ強調、ノイズ強調、低域強調は、非特許文献１に記載された手法を用いることができる。

[0030] 合成フィルタ１２８は、線形予測逆フィルタリングにより、励振信号を駆

動音源とする復号信号を合成する。

[数28]

$$\hat{s}(n) = e^{j}(n) - \sum_{i=1}^{P} \dot{a}_i^j \cdot \hat{s}(n-i)$$

[0031] なお、符号化器において、プリエンファシスを行っている場合には、ディエンファシスを行う。

[数29]

$$\hat{s}_{de-emph}(n) = \hat{s}(n) + \beta \cdot \hat{s}(n-1)$$

[0032] 一方、符号化器においてプリエンファシスを行っていない場合には、ディエンファシスを行わない。

[0033] 以下、内部状態更新に関する動作を説明する。

[0034] LP係数算出部１２２は、パケットロス時のパラメータ補間のために、ISFパラメータの内部状態を次式で算出されるベクトルにより更新する。

[数30]

$$\vec{\omega}_i = \beta\omega_i^C + (1-\beta)\frac{\omega_i^{(-3)} + \omega_i^{(-2)} + \omega_i^{(-1)}}{3}$$

[0035] ここで、$\omega_i^{(-j)}$はバッファに格納されたｊフレーム前のISFパラメータである。ω_i^Cは事前に学習等で求めた発話区間でのISFパラメータである。βは定数であり、例えば0.75のような値とすることができるが、これには限られない。ω_i^C、βは、例えば非特許文献１記載のISFコンシールメントのように、符号化対象フレームの性質を表すインデクスにより変化させてもよい。

[0036] さらに、LP係数算出部１２２は、次式に従いISF残差パラメータの内部状態も更新する。

[数31]

$$\dot{r}_i^{-1} = \dot{r}_i^0$$

[0037] 励振ベクトル合成部１２６は、以下の式に従い、励振信号ベクトルにより内部状態を更新する。

[数32]

$$u(n)=u(n+L) \qquad (0\le n<N-L)$$
$$u(n+N-L+jL')=e^{j}(n) \qquad (0\le n<L')$$

[0038] さらに、励振ベクトル合成部１２６は、ゲインパラメータの内部状態を次式により更新する。

[数33]

$$g_c^{(-M_{la}+j)}=g_c^{j}$$

[0039] 適応符号帳算出部１２３は、ピッチラグのパラメータの内部状態を次式により更新する。

[数34]

$$T_p^{(-M_{la}+j)}=T_p^{j}$$

なお、ここでは、（－２≦ｊ＜M_{la}）とするが、ｊの範囲として設計方針により異なる値を選択してもよい。

[0040] ＜パケットロスの場合＞

図６には、隠蔽信号生成部１３の機能構成例を示す。この図６に示すように、隠蔽信号生成部１３は、LP係数補間部１３０、ピッチラグ補間部１３１、ゲイン補間部１３２、雑音信号生成部１３３、ポストフィルタ１３４、合成フィルタ１３５、適応符号帳算出部１３６、及び励振ベクトル合成部１３７を備える。ただし、ポストフィルタ１３４は必須の構成要素ではない。

[0041] LP係数補間部１３０は、

[数35]

$$\dot{\omega}_i$$

を次式により算出する。なお、$\omega_i^{(-j)}$は、バッファに格納されたｊフレーム前のISFパラメータである。

[数36]

$$\dot{\omega}_i=\alpha\omega_i^{(-1)}+(1-\alpha)\vec{\omega}_i$$

ここで、

[数37]

$\vec{\omega}_i$

はパケットを正常に受信した際に算出されたISFパラメータの内部状態である。αも定数であり、0.9のような値とすることができるが、これに限定されない。αは、例えば非特許文献１記載のISFコンシールメントのように、符号化対象フレームの性質を表すインデクスにより変化させてもよい。

[0042] ISFパラメータからLP係数を得る手順は、正常にパケットを受信した場合と同様である。

[0043] ピッチラグ補間部１３１は、ピッチラグに関する内部状態パラメータ

[数38]

$T_p^{(-M_{la}+j)}$

を用いて、ピッチラグの予測値

[数39]

$\hat{T}_p$

を算出する。具体的な処理手順として、非特許文献１の手法を用いることができる。

[0044] ゲイン補間部１３２は、固定符号帳ゲインの補間のために、非特許文献１に記載の以下の式に従う手法を用いることができる。

[数40]

$$g_s \simeq 0.4 \cdot g_c^{-1} + 0.3 \cdot g_c^{-2} + 0.2 \cdot g_c^{-3} + 0.1 \cdot g_c^{-4}$$

[0045] 雑音信号生成部１３３は、固定符号帳ベクトルと同じ長さの白色雑音を生成し、固定符号帳ベクトルとして代用する。

[0046] ポストフィルタ１３４、合成フィルタ１３５、適応符号帳算出部１３６及び励振ベクトル合成部１３７の動作は、前述した正常にパケットを受信した場合の動作と同様である。

[0047] 内部状態更新は、ISF残差パラメータを除き、正常にパケットを受信した場合と同様である。ISF残差パラメータの更新は、LP係数補間部１３０により、

次式に従い行われる。

[数41]

$$\dot{r}_i^0 = \dot{\omega}_i^0 - mean_i - \frac{1}{3}\dot{r}_i^{-1}$$

先行技術文献

特許文献

[0048] 特許文献1：国際公開2002/035520号公報

特許文献2：国際公開2008/108080号公報

非特許文献

[0049] 非特許文献1：ITU-T Recommendation G.718、2008年6月

発明の概要

発明が解決しようとする課題

[0050] 上記の通り、CELP符号化は内部状態を持つため、パケットロス時に補間により得られたパラメータと、本来復号に用いるべきパラメータとの乖離により音質の劣化が起こる。特にISFパラメータは、フレーム内・フレーム間での予測符号化が行われるため、パケットロスによる影響がパケットロス復帰後にも継続する課題がある。

[0051] より具体的には、音声の開始部分付近で起こったパケットロスから復帰した最初のフレームにおいて、急激にパワーが上昇する課題が確認されている。これは、励振信号のパワーが高くなる音声の開始部分において、パケットロス時の補間処理により得られたISF係数から算出したLP係数のインパルス応答が、本来復号器が想定しているよりも高い利得を持つことにより引き起こされ、主観品質上不快な不連続音として知覚される。

[0052] 特許文献１の方法は、ロスしたフレームに対して補間したISF係数を生成するが、ロスから復帰した最初のフレームについては、通常の復号によりISFパラメータを生成するため、上記急激なパワーの上昇を抑えることができない。

[0053] 一方、特許文献２の方法では、符号化側で求めたゲイン調整用パラメータ（正規化予測残差パワー）を伝送し、復号側でのパワー調整に用いることで、パケットロスしたフレームの励振信号のパワーを抑え、急激なパワーの上昇を防止することができる。

[0054] 図７には、特許文献２の技術に相当する音声復号器１Ｘの機能構成例を、図８には、隠蔽信号生成部１３Ｘの機能構成例をそれぞれ示す。特許文献２において、音声パケットは、従来手法において述べたパラメータに加えて、少なくとも正規化予測残差パワーの補助情報を含む。

[0055] 音声信号生成部１Ｘが備える正規化予測残差パワー復号部１５は、音声パケットから正規化予測残差パワーの補助情報を復号して、参照正規化予測残差パワーを算出し、隠蔽信号生成部１３Ｘに出力する。

[0056] 隠蔽信号生成部１３Ｘの構成要素のうち、正規化予測残差調整部１３８以外の構成要素は、前述した従来技術と同じであるので、以下では正規化予測残差調整部１３８についてのみ述べる。

[0057] 正規化予測残差調整部１３８は、LP係数補間部１３０が出力したLP係数から正規化予測残差パワーを算出する。次に、正規化予測残差調整部１３８は、正規化予測残差パワーと参照正規化予測残差パワーを用いて、合成フィルタゲイン調整係数を算出する。最後に、正規化予測残差調整部１３８は、合成フィルタゲイン調整係数を励振信号に乗算して、合成フィルタ１３５に出力する。

[0058] 上記の特許文献２の技術によれば、パケットロス時の隠蔽信号のパワーを正常受信時と同様に抑えることができるものの、低ビットレート音声符号化においては、上記ゲイン調整用パラメータの伝送に必要なビットレートを確保するのが困難である。また、隠蔽信号生成部における処理であるため、リカバリフレームにおいてISFパラメータの不一致により引き起こされるパワーの急激な変化には対応することが困難である。

[0059] そこで、本発明は、音声開始時点でのパケットロスから復帰した際に起こりうる不連続音を低減し、主観品質を改善することを目的とする。

課題を解決するための手段

[0060] 本発明の一実施形態に係る音声信号処理装置は、パケットロスが起こった後に最初に正常に受信した音声パケットに対して、音声パケット復号の結果得られる復号音声の振幅が急激に大きくなることにより生じる不連続性の推定を行う不連続推定器と、復号音声の不連続性を修正する不連続修正器と、を備える。

[0061] 上記の不連続推定器は、復号音声の不連続性を、励振信号のパワーにより推定してもよい。

[0062] また、上記の不連続推定器は、復号音声の不連続性を、励振信号算出に用いる量子化された符号帳ゲインにより推定してもよい。

[0063] 上記の音声信号処理装置は、符号化器より伝送された不連続性に関する補助情報を復号する補助情報復号器、をさらに備え、上記の不連続推定器は、補助情報復号器が補助情報符号を復号して出力する不連続性に関する補助情報を用いて、復号音声の不連続性を推定してもよい。

[0064] 上記の不連続修正器は、不連続性の推定結果に応じて、ISFパラメータ又はLSFパラメータ（以下「ISF/LSFパラメータ」と表記する）を修正してもよい。

[0065] より具体的には、上記の不連続修正器は、合成フィルタの安定性保証のために与えるISF/LSFパラメータの各要素間の間隔を、不連続性の推定結果に応じて変化させてもよい。

[0066] このとき、上記の不連続修正器は、合成フィルタの安定性保証のために与えるISF/LSFパラメータの各要素間の間隔として、安定性保証のために通常与える間隔よりも大きな間隔を与えてもよい。

[0067] また、上記の不連続修正器は、予め定めた次元までのISF/LSFパラメータを等分して得られる間隔を、合成フィルタの安定性保証のために与えるISF/LSFパラメータの各要素間の間隔として用いてもよい。

[0068] さらに、上記の不連続修正器は、予め定めたベクトルにより、ISF/LSFパラメータの一部もしくは全部を置き換えてもよい。

[0069] 本発明の一実施形態に係る音声信号処理装置は、ISF/LSFパラメータを量子化するISF/LSF量子化器と、ISF/LSFパラメータに関する隠蔽情報である隠蔽ISF/LSFパラメータを生成するISF/LSF隠蔽器と、前記ISF/LSF量子化器の量子化過程で得られる量子化ISF/LSFパラメータと前記ISF/LSF隠蔽器が生成する隠蔽ISF/LSFパラメータとの距離を用いて、パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性を推定する不連続推定器と、不連続性に関する補助情報を符号化する補助情報符号化器と、を備える。

[0070] 本発明の一実施形態に係る音声信号処理装置は、パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性を推定する不連続推定器と、不連続性に関する補助情報を符号化する補助情報符号化器と、前記不連続推定器が不連続性を推定しない場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF量子化に用い、前記不連続推定器が不連続性を推定した場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF量子化に用いることを回避するISF/LSF量子化器と、を備える。

[0071] 本発明の一実施形態に係る音声信号処理装置は、パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性に関する補助情報を復号し出力する補助情報復号器と、復号音声の不連続性を修正する不連続修正器と、前記補助情報復号器の出力により不連続性を推定しない場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF算出に用い、前記補助情報復号器の出力により不連続性を推定した場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF算出に用いることを回避するISF/LSF復号器と、を備える。

[0072] 音声信号処理装置は、過去の所定数のフレームのパケット受信状態を判定する受信状態判定部、をさらに備え、不連続修正器は、不連続性の推定結果に加え、パケット受信状態の判定結果も基礎として、不連続性の修正を行う構成を採用してもよい。

[0073] さて、本発明の一実施形態に係る音声信号処理装置は、音声信号処理方法

に係る発明、及び音声信号処理プログラムに係る発明として捉えることもでき、以下のように記述することができる。

[0074] 本発明の一実施形態に係る音声信号処理方法は、音声信号処理装置により実行される音声信号処理方法であって、パケットロスが起こった後に最初に正常に受信した音声パケットに対して、音声パケット復号の結果得られる復号音声の振幅が急激に大きくなることにより生じる不連続性の推定を行うステップと、復号音声の不連続性を修正するステップと、を備える。

[0075] 本発明の一実施形態に係る音声信号処理方法は、音声信号処理装置により実行される音声信号処理方法であって、ISF/LSFパラメータを量子化するステップと、ISF/LSFパラメータに関する隠蔽情報である隠蔽ISF/LSFパラメータを生成するステップと、前記ISF/LSFパラメータの量子化過程で得られる量子化ISF/LSFパラメータと生成された隠蔽ISF/LSFパラメータとの距離を用いて、パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性を推定するステップと、不連続性に関する補助情報を符号化するステップと、を備える。

[0076] 本発明の一実施形態に係る音声信号処理方法は、音声信号処理装置により実行される音声信号処理方法であって、パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性を推定するステップと、不連続性に関する補助情報を符号化するステップと、不連続性が推定されない場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF量子化に用い、不連続性が推定された場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF量子化に用いることを回避するステップと、を備える。

[0077] 本発明の一実施形態に係る音声信号処理方法は、音声信号処理装置により実行される音声信号処理方法であって、パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性に関する補助情報を復号し出力するステップと、復号音声の不連続性を修正するステップと、前記補助情報が不連続性の推定を示さない場合は、過去の量子化ISF/LSF残差パラメータ

を当該フレームにおけるISF/LSF算出に用い、前記補助情報が不連続性の推定を示す場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF算出に用いることを回避するステップと、を備える。

[0078] 本発明の一実施形態に係る音声信号処理プログラムは、コンピュータを、パケットロスが起こった後に最初に正常に受信した音声パケットに対して、音声パケット復号の結果得られる復号音声の振幅が急激に大きくなることにより生じる不連続性の推定を行う不連続推定器と、復号音声の不連続性を修正する不連続修正器、として動作させるための音声信号処理プログラムである。

[0079] 本発明の一実施形態に係る音声信号処理プログラムは、コンピュータを、ISF/LSFパラメータを量子化するISF/LSF量子化器と、ISF/LSFパラメータに関する隠蔽情報である隠蔽ISF/LSFパラメータを生成するISF/LSF隠蔽器と、前記ISF/LSF量子化器の量子化過程で得られる量子化ISF/LSFパラメータと前記ISF/LSF隠蔽器が生成する隠蔽ISF/LSFパラメータとの距離を用いて、パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性を推定する不連続推定器と、不連続性に関する補助情報を符号化する補助情報符号化器、として動作させるための音声信号処理プログラムである。

[0080] 本発明の一実施形態に係る音声信号処理プログラムは、コンピュータを、パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性を推定する不連続推定器と、不連続性に関する補助情報を符号化する補助情報符号化器と、前記不連続推定器が不連続性を推定しない場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF量子化に用い、前記不連続推定器が不連続性を推定した場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF量子化に用いることを回避するISF/LSF量子化器、として動作させるための音声信号処理プログラムである。

[0081] 本発明の一実施形態に係る音声信号処理プログラムは、コンピュータを、パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不

連続性に関する補助情報を復号し出力する補助情報復号器と、復号音声の不連続性を修正する不連続修正器と、前記補助情報復号器の出力により不連続性を推定しない場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF算出に用い、前記補助情報復号器の出力により不連続性を推定した場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF算出に用いることを回避するISF/LSF復号器、として動作させるための音声信号処理プログラムである。

発明の効果

[0082] 以上のような本発明によれば、音声開始時点でのパケットロスから復帰した際に起こりうる上記不連続音を低減し、主観品質を改善することができる。

図面の簡単な説明

[0083] [図1]音声復号器の構成図である。

[図2]音声復号器の処理フローである。

[図3]音声符号復号部の機能構成図である。

[図4]LP係数算出部の機能構成図である。

[図5]LP係数を算出する処理フローである。

[図6]隠蔽信号生成部の機能構成図である。

[図7]特許文献２の音声復号器の構成図である。

[図8]特許文献２の隠蔽信号生成部の機能構成図である。

[図9]第１実施形態における音声符号復号部の機能構成図である。

[図10]第１実施形態におけるLP係数算出部の処理フローである。

[図11]第１実施形態における音声符号復号部の機能構成図である。

[図12]第１実施形態変形例１における第２安定性処理部の処理フローである。

[図13]第２実施形態における音声符号復号部の機能構成図である。

[図14]第２実施形態におけるLP係数算出部の機能構成図である。

[図15]第２実施形態におけるLP係数算出の処理フローである。

[図16]第４実施形態における音声符号化器の構成図である。

[図17]第４実施形態における音声符号化器の構成図である。

[図18]第４実施形態におけるLP分析・符号化器の構成図である。

[図19]第４実施形態におけるLP分析・符号化器の処理フローである。

[図20]第４実施形態における音声符号復号部の機能構成図である。

[図21]第４実施形態におけるLP係数算出部の処理フローである。

[図22]第５実施形態におけるLP分析・符号化器の構成図である。

[図23]第５実施形態におけるLP分析・符号化器の処理フローである。

[図24]第４実施形態における音声符号復号部の機能構成図である。

[図25]第５実施形態におけるLP係数算出部の処理フローである。

[図26]第７実施形態における音声復号器の構成図である。

[図27]第７実施形態における音声復号器の処理フローである。

[図28]第７実施形態における音声符号復号部の機能構成図である。

[図29]第７実施形態におけるLP係数算出の処理フローである。

[図30]コンピュータのハードウェア構成例を示す図である。

[図31]コンピュータの外観図である。

[図32]（ａ）、（ｂ）、（ｃ）および（ｄ）は音声信号処理プログラムのさまざまな構成例を示す図である。

発明を実施するための形態

[0084] 以下、図面を用いて、本発明に係る音声信号処理装置、音声信号処理方法、及び音声信号処理プログラムの好適な実施形態について詳細に説明する。なお、図面の説明においては同一要素には同一符号を付し、重複する説明を省略する。

[0085] ［第１実施形態］

第１実施形態における音声信号処理装置は、前述した図１の音声復号器１と同様の構成を備え、音声符号復号部に特徴があるため、以下、音声符号復号部について説明する。

[0086] 図９には第１実施形態における音声符号復号部１２Ａの機能構成図を、図

１０にはLP係数算出処理のフロー図を、それぞれ示す。図９の音声符号復号部１２Ａは、前述した図３の構成に対し、不連続検出部１２９が追加されている。従来技術との差異は、LP係数算出処理のみであるため、ここではLP係数算出処理に関わる各部の動作について述べる。

[0087] 不連続検出部１２９は、復号して得られた固定符号帳ゲインg_c^0、及び内部状態に含まれる固定符号帳ゲインg_c^{-1}を参照し、次式に従いゲインの変化と閾値とを比較する（図１０のステップＳ１１）。

[数42]

$$\log(g_c^0) - \log(g_c^{-1}) > Thres$$

[0088] ゲインの変化が閾値を越える場合、不連続発生を検出し（以下、単に「不連続を検出し」ともいう）、不連続発生の検出結果に係る制御信号を安定性処理部１２１に出力する。

[0089] なお、ゲインの変化と閾値との比較にあたっては、次式を用いても良い。

[数43]

$$g_c^0 - g_c^{-1} > Thres$$

[0090] さらに、現フレームに含まれる第１から第４サブフレームの固定符号帳ゲインのうち最大のものを$g_c^{(c)}$、内部状態に含まれる固定符号帳ゲインのうち最小のものを$g_c^{(p)}$として、次式によりゲインの変化と閾値との比較を行ってもよい。

[数44]

$$\log(g_c^{(c)}) - \log(g_c^{(p)}) > Thres$$

[0091] 当然、次式を用いてもよい。

[数45]

$$g_c^{(c)} - g_c^{(p)} > Thres$$

[0092] 第１実施形態の上記の例では、直前フレーム（ロストフレーム）の第４サブフレームの固定符号帳ゲインg_c^{-1}と、現フレームの第１サブフレームの固定符号帳ゲインg_c^0を用いて不連続検出を行う例を示したが、内部状態に含まれ

る固定符号帳ゲイン及び現フレームに含まれる固定符号帳ゲインについて、それぞれ平均値を算出した上で、ゲインの変化と閾値との比較を行うようにしてもよい。

[0093] ISF復号部１２０は、従来技術と同様の動作を行う（図１０のステップＳ１２）。

[0094] 安定性処理部１２１は、不連続検出部１２９が不連続を検出した場合、以下の処理によりISFパラメータを修正する（図１０のステップＳ１３）。

[0095] 最初に、安定性処理部１２１は、内部状態バッファ１４に記憶されたISFパラメータ

[数46]

$$\dot{\omega}_i^{-1}$$

について、各要素間に通常よりM$_{-1}$倍の間隔をあける処理を行う。通常よりも非常に大きな間隔を与えることによりスペクトル包絡における過大なピークとディップを抑制する効果を与える。ここで、min_distは最小のISF間隔であり、isf_minは、min_distの間隔を確保するために必要なISFの最小値である。isf_minは、隣のISFの値にmin_distの間隔を加算することにより順次更新を行う。一方、isf_maxは、min_distの間隔を確保するために必要なISFの最大値である。isf_maxは、隣のISFの値からmin_distの間隔を減算することにより順次更新を行う。

[数47]

$isf_min = min_dist = 50M_{-1}$

for i=0 to 14

if $\dot{\omega}_i^{-1}$<isf_min then $\dot{\omega}_i^{-1}$=isf_min

$isf_min = \dot{\omega}_i^{-1} + min_dist$

$isf_max = 6400 - min_dist$

if $\dot{\omega}_{14}^{-1}$>isf_max

for i=14 down to 1

if $\dot{\omega}_i^{-1}$ > isf_max then $\dot{\omega}_i^{-1}$=isf_max

$isf_max = \dot{\omega}_i^{-1} - min_dist$

[0096] 次に、安定性処理部１２１は、現フレームのISFパラメータについて、各要素間に通常よりM_0倍の間隔をあける処理を行う。ここでは、$1 < M_0 < M_{-1}$とするが、M_{-1}あるいはM_0のいずれか一方を１とし、他方を１より大きな値に設定してもよい。

[数48]

$isf_min = min_dist = 50M_0$

for i=0 to 14

if $\dot{\omega}_i^{0}$<isf_min then $\dot{\omega}_i^{0}$=isf_min

$isf_min = \dot{\omega}_i^{0} + min_dist$

$isf_max = 6400 - min_dist$

if $\dot{\omega}_{14}^{0}$>isf_max

for i=14 down to 1

if $\dot{\omega}_i^{0}$ > isf_max then $\dot{\omega}_i^{0}$=isf_max

$isf_max = \dot{\omega}_i^{0} - min_dist$

[0097] また、安定性処理部１２１は、不連続検出器が不連続を検出しない場合、通常の復号過程で実施するのと同様に、以下の処理を行う。

[数49]

```
isf_min = min_dist = 50
for i=0 to 14
        if ω_i^0 <isf_min then ω_i^0=isf_min
        isf_min = ω_i^0 + min_dist
isf_max = 6400 - min_dist
if ω_14^0 >isf_max
        for i=14 down to 1
                if ω_i^0 > isf_max then ω_i^0=isf_max
                isf_max = ω_i^0 - min_dist
```

[0098] 不連続を検出した場合の要素間の最小の間隔はISFの周波数に応じて変えてもよい。不連続を検出した場合の要素間の最小の間隔は、通常の復号処理の最小の要素間の間隔と異なっていれば良い。

[0099] LP係数算出部１２２内のISF-ISP変換部１２２Ａは、次式に従い、ISFパラメータ

[数50]

$\dot{\omega}_i$、$\dot{\omega}_i^{-1}$

をそれぞれ、ISPパラメータ

[数51]

$\dot{\mathbf{q}}_i$、$\dot{\mathbf{q}}_i^{-1}$

に変換する（図１０のステップＳ１４）。ここで、Cは事前に定めた定数である。

[数52]

$$\dot{\mathbf{q}}_i = \cos(C \cdot \dot{\omega}_i)$$

[0100] ISP補間部１２２Ｂは、過去のISPパラメータ

[数53]

$\dot{\mathbf{q}}_i^{-1}$

と、上記ISPパラメータ

[数54]

$$\dot{\mathbf{q}}_i$$

から、以下の式に従い、サブフレーム毎のISPパラメータを算出する（図１０のステップＳ１５）。補間にあたっては、別の係数を用いてもよい。

[数55]

$$\mathbf{q}_i^{(1)} = 0.75 \cdot \dot{\mathbf{q}}_i^{-1} + 0.25 \cdot \dot{\mathbf{q}}_i$$
$$\mathbf{q}_i^{(2)} = 0.5 \cdot \dot{\mathbf{q}}_i^{-1} + 0.5 \cdot \dot{\mathbf{q}}_i$$
$$\mathbf{q}_i^{(3)} = 0.25 \cdot \dot{\mathbf{q}}_i^{-1} + 0.75 \cdot \dot{\mathbf{q}}_i$$
$$\mathbf{q}_i^{(4)} = \dot{\mathbf{q}}_i$$

[0101] ISP-LPC変換部１２２Ｃは、サブフレーム毎のISPパラメータをLP係数

[数56]

$$\dot{a}_i^j \quad (0 < i \le P、\ 0 \le j < 4)$$

に変換する（図１０のステップＳ１６）。ここで、先読み信号に含まれるサブフレームの数を４としたが、サブフレームの数は設計方針により変更してもよい。具体的変換手順として、非特許文献１に記載の処理手順を用いることができる。

[0102] さらに、ISF-ISP変換部１２２Ａは、内部状態バッファ１４に記憶されたISFパラメータ

[数57]

$$\dot{\omega}_i^{-1}$$

を次式に従い更新する。

[数58]

$$\dot{\omega}_i^{-1} = \dot{\omega}_i^0$$

このとき、不連続が検出された場合でも、ISF-ISP変換部１２２Ａは、以下の手順を実施することにより、ISFパラメータの算出結果を用いて、内部状態バッファに記憶されたISFパラメータ

[数59]

$\dot{\omega}_i^{-1}$

を更新してもよい。

[数60]

```
isf_min = min_dist = 50
for i=0 to 14
        if ω̇_i^0<isf_min then ω̇_i^0=isf_min
        isf_min = ω̇_i^0 + min_dist
isf_max = 6400 - min_dist
if ω̇_14^0>isf_max
        for i=14 down to 1
                if ω̇_i^0 > isf_max then ω̇_i^0=isf_max
                isf_max = ω̇_i^0 - min_dist
```

[0103] 以上の第１実施形態のように、復号音声の不連続性は、励振信号算出に用いる量子化された符号帳ゲインによって推定することができ、また、不連続性の推定結果に応じて、ISF/LSFパラメータ（例えば合成フィルタの安定性保証のために与えるISF/LSFパラメータの各要素間の間隔）を修正することができる。これにより、音声開始時点でのパケットロスから復帰した際に起こりうる不連続音を低減し、主観品質を改善することができる。

[0104] [第１実施形態の変形例]

図１１には、第１実施形態の変形例に係る音声符号復号部１２Ｓの機能構成図を示す。図３の従来技術の構成との差異は、不連続検出部１２９及び第２安定性処理部１２１Ｓのみであるため、これらの動作について述べる。第２安定性処理部１２１Ｓは、ゲイン調整部１２１Ｘ及びゲイン乗算部１２１Ｙを備えており、第２安定性処理部１２１Ｓの処理フローは図１２に示す。

[0105] 不連続検出部１２９は、第１実施形態の不連続検出部１２９と同様にして、復号して得られた固定符号帳ゲインg_c^0及び内部状態に含まれる固定符号帳ゲインg_c^{-1}を参照し、ゲインの変化と閾値とを比較する。そして、不連続検出

部１２９は、ゲインの変化が閾値を越えたか否かに関する情報を含む制御信号を、ゲイン調整部１２１Ｘに送る。

[0106] ゲイン調整部１２１Ｘは、ゲインの変化が閾値を越えたか否かに関する情報を制御信号から読み出し、ゲインの変化が閾値を越えた場合には、事前に定めたゲインg_{on}をゲイン乗算部１２１Ｙに出力する。一方、ゲイン調整部１２１Ｘは、ゲインの変化が閾値を越えていない場合は、事前に定めたゲインg_{off}をゲイン乗算部１２１Ｙに出力する。このようなゲイン調整部１２１Ｘの動作は図１２のステップＳ１８に対応する。

[0107] ゲイン乗算部１２１Ｙは、合成フィルタ１２８が出力する合成信号に上記ゲインg_{on}又はゲインg_{off}を乗算し（図１２のステップＳ１９）、得られた復号信号を出力する。

[0108] ここで、LP係数算出部１２２からLP係数またはISFパラメータを出力して、第２安定性処理部１２１Ｓに入力する構成（図１１にてLP係数算出部１２２からゲイン調整部１２１Ｘへの点線で示す構成）にしてもよい。この場合、乗算するべきゲインは、LP係数算出部１２２で算出したLP係数またはISFパラメータを用いて決定される。

[0109] 以上の変形例のように、音声符号復号部１２Ｓに第２安定性処理部１２１Ｓを追加し、ゲインの変化が閾値を越えたか否かに応じてゲインを調整することで、適正な復号信号を得ることができる。

[0110] なお、第２安定性処理部１２１Ｓは、上記算出したゲインを励振信号に乗算し、合成フィルタ１２８に出力してもよい。

[0111] ［第２実施形態］

第２実施形態における音声信号処理装置は、前述した図１の音声復号器１と同様の構成を備え、音声符号復号部に特徴があるため、以下、音声符号復号部について説明する。図１３には音声符号復号部１２Ｂの機能構成例を、図１４にはLP係数の算出処理に係る機能構成例を、図１５にはLP係数の算出処理のフローを、それぞれ示す。図１３の音声符号復号部１２Ｂは、前述した図３の構成に対し、不連続検出部１２９が追加されている。

[0112] ISF復号部１２０は、従来技術と同様にしてISFパラメータを算出する（図１５のステップＳ２１）。

[0113] 安定性処理部１２１は、従来技術と同様にして、フィルタの安定性を確保するためにISFパラメータ

[数61]

$\dot{\omega}_i$

の各要素間に50Hz以上の間隔をあけるよう処理を行う（図１５のステップＳ２２）。

[0114] ISF-ISP変換部１２２Ａは、第１実施形態と同様にして、安定性処理部１２１が出力したISFパラメータをISPパラメータに変換する（図１５のステップＳ２３）。

[0115] ISP補間部１２２Ｂは、第１実施形態と同様にして、過去のISPパラメータ

[数62]

$\dot{\mathbf{q}}_i^{-1}$

と、ISF-ISP変換部１２２Ａによる変換で得られたISPパラメータ

[数63]

$\dot{\mathbf{q}}_i$

から、サブフレーム毎のISPパラメータを算出する（図１５のステップＳ２４）。

[0116] ISP-LPC変換部１２２Ｃは、第１実施形態と同様にして、サブフレーム毎のISPパラメータをLP係数

[数64]

$\dot{a}_i^j \quad (0 < i \le P、\ 0 \le j < 4)$

に変換する（図１５のステップＳ２５）。ここで、先読み信号に含まれるサブフレームの数を４としたが、サブフレームの数は設計方針により変更してもよい。

[0117] 内部状態バッファ１４は、新たなISFパラメータによって、過去に記憶した

ISFパラメータを更新する。

[0118] 不連続検出部１２９は、パケットロスしたフレームにおける第４サブフレームのLP係数を内部状態バッファ１４から読み出し、パケットロスしたフレームにおける第４サブフレームのLP係数のインパルス応答のパワーを算出する。パケットロスしたフレームにおける第４サブフレームのLP係数は、パケットロス時に図６の隠蔽信号生成部１３に含まれるLP係数補間部１３０が出力して内部状態バッファ１４に蓄積した係数を用いることができる。

[数65]

$$E_{-1} = 10\log\left(\sum_{n=0}^{L'-1} h_{-1}^2(n)\right)$$

$$h_{-1}(n) = \delta(n) - \sum_{i=1}^{P} \dot{a}_i^{(-1)} \cdot h_{-1}(n-i)$$

[0119] そして、不連続検出部１２９は、例えば以下の式により不連続を検出する（図１５のステップＳ２６）。

[数66]

$$E_0 - E_{-1} > Thres$$

[0120] ゲインの変化が閾値を越えない場合（図１５のステップＳ２７でＮＯの場合）、不連続検出部１２９は不連続発生を検出せず、ISP-LPC変換部１２２CからLP係数を出力して処理を終了する。一方、ゲインの変化が閾値を越える場合（図１５のステップＳ２７でＹＥＳの場合）、不連続検出部１２９は、不連続発生を検出し、不連続発生の検出結果に係る制御信号を安定性処理部１２１に送る。制御信号を受け取った場合、安定性処理部１２１は、第１実施形態と同様にして、ISFパラメータを修正する（図１５のステップＳ２８）。以下、ISF-ISP変換部１２２Ａ、ISP補間部１２２Ｂ、及びISP-LPC変換部１２２Cの動作（図１５のステップＳ２９、Ｓ２Ａ、Ｓ２Ｂ）は、上記と同様である。

[0121] 以上の第２実施形態のように、復号音声の不連続性は、励振信号のパワーによって推定することができ、第１実施形態と同様に、不連続音を低減し主

観品質を改善することができる。

[0122] [第３実施形態]

不連続を検出した際に、別の方法によりISFパラメータを修正してもよい。第３実施形態は、安定性処理部１２１のみが第１実施形態と異なるので、安定性処理部１２１の動作のみについて述べる。

[0123] 不連続検出部１２９が不連続を検出した場合、安定性処理部１２１は、以下の処理を行いISFパラメータを修正する。

[0124] 内部状態バッファ１４に記憶されたISFパラメータ

[数67]

$$\dot{\omega}_i^{-1}$$

について、安定性処理部１２１は、低次P’次元（0＜P’≦P）までのISFパラメータを次式により置き換える。ここで、

[数68]

$$\delta^{-1} = \dot{\omega}_{P'-1}^{-1} / P'$$

とする。

[数69]

$$\dot{\omega}_i^{-1} = \dot{\omega}_{i-1}^{-1} + \delta^{-1}$$
$$\dot{\omega}_0^{-1} = \delta^{-1} \qquad (0 \le i < P')$$

[0125] また、安定性処理部１２１は、次のように事前に学習により得られたP’次元ベクトルで低次P’次元のISFパラメータを上書きしてもよい。

[数70]

$$\dot{\omega}_i^{-1} = \boldsymbol{\omega}_i^0 \qquad (0 \le i < P')$$

[0126] 次に、現フレームのISFパラメータについて、安定性処理部１２１は、第１実施形態のように各要素間に通常よりM_0倍の間隔をあける処理を行ってもよいし、次式に従い決定してもよい。ここで、

[数71]

$$\delta^0 = \dot{\omega}_{P'-1}^0 / P'$$

とする。

[数72]

$$\dot{\omega}_i^0 = \dot{\omega}_{i-1}^0 + \delta^0$$
$$\dot{\omega}_0^0 = \delta^0$$

[0127] また、安定性処理部１２１は、事前に学習したP’次元ベクトルで上書きしてもよい。

[数73]

$$\dot{\omega}_i^0 = \boldsymbol{\omega}_i^0 \quad (0 \le i < P')$$

[0128] さらには、上記P’次元ベクトルは復号過程で学習されてもよく、例えば、

[数74]

$$\boldsymbol{\omega}_i^0 = (1-\lambda)\boldsymbol{\omega}_i^{-1} + \lambda\dot{\omega}_i^{-1}$$

としてもよい。ただし、復号開始時のフレームにおいてはω_i^{-1}を予め決められたP’次元ベクトルω_i^{init}としてもよい。

[0129] 内部状態バッファ１４は、新たなISFパラメータによって、過去に記憶したISFパラメータを更新する。

[0130] 以上の第３実施形態のように、予め定めた次元までのISF/LSFパラメータを等分して得られる間隔を、合成フィルタの安定性保証のために与えるISF/LSFパラメータの各要素間の間隔として用いることができ、第１、第２実施形態と同様に、不連続音を低減し主観品質を改善することができる。

[0131] ［第４実施形態］

第４実施形態では、符号化側が、不連続の発生を検出して、不連続判定符号（検出結果を示す符号）を音声符号に含めて復号側へ伝送し、復号側が、音声符号に含まれる不連続判定符号に基づいて安定性処理の処理内容を決定する実施形態を説明する。

[0132] （符号化側について）

図１６に符号化器２の機能構成例を、図１７に符号化器２における処理のフロー図を示す。図１６に示すように、符号化器２は、LP分析・符号化部２

１、残差符号化部２２、及び符号多重化部２３を備える。

[0133] このうちLP分析・符号化部２１の機能構成例を図１８に、LP分析・符号化部２１における処理のフロー図を図１９に示す。図１８に示すように、LP分析・符号化部２１は、LP分析部２１０、LP-ISF変換部２１１、ISF符号化部２１２、不連続判定部２１３、ISF隠蔽部２１４、ISF-LP変換部２１５、及びISFバッファ２１６を備える。

[0134] LP分析・符号化部２１において、LP分析部２１０は、入力信号に対して線形予測分析を行い、線形予測係数を求める（図１７のステップＴ４１、図１８のステップＵ４１）。線形予測係数の算出に当たっては、音声信号から自己相関関数を算出した上で、レビンソン・ダービン法等を用いることができる。

[0135] LP-ISF変換部２１１は、第１実施形態と同様にして、算出した線形予測係数をISFパラメータに変換する（ステップＴ４２、Ｕ４２）。線形予測係数からISFパラメータへの変換には、非特許文献に記載の方法を用いてもよい。

[0136] ISF符号化部２１２は、ISFパラメータを事前に定めた方法により符号化してISF符号を算出し（ステップＴ４３、Ｕ４３）、符号化の過程で得られる量子化ISFパラメータを、不連続判定部２１３、ISF隠蔽部２１４及びISF-LP変換部２１５に出力する（ステップＵ４７）。ここで、量子化ISFパラメータは、ISF符号を逆量子化して得られるISFパラメータと等しい。符号化の方法としてベクトル符号化や直前フレームのISF及び事前に学習により定めた平均ベクトルからの誤差ベクトルをベクトル量子化等により符号化してもよい。

[0137] 不連続判定部２１３は、不連続判定部２１３が内蔵する内部バッファ（不図示）に格納された不連続判定フラグを符号化して、得られた不連続判定符号を出力する（ステップＵ４７）。また、不連続判定部２１３は、ISFバッファ２１６から読み出した隠蔽ISFパラメータ

[数75]

$$\widetilde{\omega}_i$$

と、量子化ISFパラメータ

[数76]

$$\dot{\omega}_i$$

を用いて、次式に従い不連続の判定を行い（ステップＴ４４、Ｕ４６）、その判定結果を不連続判定部２１３の内部バッファに格納する。ここで、$Thres_\omega$は事前に定めた閾値、P’は次式を満たす整数である（0＜P’≦P）。

[数77]

$$\sum_{i=0}^{P'-1}(\dot{\omega}_i - \tilde{\omega}_i)^2 > Thres_\omega$$

[0138] ここでは、ISFパラメータ同士のユークリッド距離を用いて不連続判定を行う例を述べたが、別の方法により不連続判定を行ってもよい。

[0139] ISF隠蔽部２１４は、デコーダ側のISF隠蔽部と同様の処理により、量子化ISFパラメータから隠蔽ISFパラメータを算出して、得られた隠蔽ISFパラメータをISFバッファ２１６に出力する（ステップＵ４４、Ｕ４５）。ISF隠蔽処理の処理手順は、デコーダ側のパケットロス隠蔽部と同じ処理であれば、どんな方法でもよい。

[0140] ISF-LP変換部２１５は、上記量子化ISFパラメータを変換して量子化線形予測係数を算出し、得られた量子化線形予測係数を残差符号化部２２へ出力する（ステップＴ４５）。ISFパラメータを量子化線形予測係数に変換する方法としては、非特許文献に記載の方法を用いてもよい。

[0141] 残差符号化部２２は、量子化線形予測係数を用いて音声信号をフィルタリングし、残差信号を算出する（ステップＴ４６）。

[0142] 次に、残差符号化部２２は、残差信号を、CELP又はTCX（Transform Coded Excitation）を用いる符号化手段、CELPとTCXとを切り替えて用いる符号化手段等により符号化し、残差符号を出力する（ステップＴ４７）。残差符号化部２２の処理は、本発明とは関連性が低いので、説明を省略する。

[0143] 符号多重化部２３は、ISF符号、不連続判定符号、及び残差符号を所定の順序でまとめて、得られた音声符号を出力する（ステップＴ４８）。

[0144] （復号側について）

第４実施形態における音声信号処理装置は、前述した図１の音声復号器１と同様の構成を備え、音声符号復号部に特徴があるため、以下、音声符号復号部について説明する。図２０には音声符号復号部１２Ｄの機能構成例を、図２１にはLP係数の算出処理のフローを、それぞれ示す。図２０の音声符号復号部１２Ｄは、前述した図３の構成に対し、不連続検出部１２９が追加されている。

[0145] ISF復号部１２０は、ISF符号を復号して安定性処理部１２１及び内部状態バッファ１４に出力する（図２１のステップＳ４１）。

[0146] 不連続検出部１２９は、不連続判定符号を復号し、得られた不連続検出結果を安定性処理部１２１に出力する（図２１のステップＳ４２）。

[0147] 安定性処理部１２１は、不連続検出結果に応じた安定性処理を行う（図２１のステップＳ４３）。安定性処理部の処理手順は、第１実施形態及び第３実施形態と同様の方法を用いることができる。

[0148] なお、安定性処理部１２１は、不連続判定符号から得られた不連続検出結果に加えて、音声符号に含まれている他のパラメータも基礎として、以下のように安定性処理を行ってもよい。例えば、安定性処理部１２１は、ISF安定度stabを次式に従い算出し、ISF安定度が閾値を越える場合には、たとえ不連続判定符号から不連続が検出された旨の不連続検出結果が得られていても、不連続が検出されていないかのように安定性処理を行う構成にしてもよい。ここで、Cは事前に定めた定数である。

[数78]

$$stab = 1.25 - \sum_{i=0}^{P'-1} (\dot{\omega}_i^0 - \dot{\omega}_i^{-1})^2 \Big/ C$$

[0149] LP係数算出部１２２内のISF-ISP変換部１２２Ａは、第１実施形態と同様の処理手順によりISFパラメータをISPパラメータに変換する（図２１のステップＳ４４）。

[0150] ISP補間部１２２Ｂは、第１実施形態と同様の処理手順によりサブフレーム毎のISPパラメータを算出する（図２１のステップＳ４５）。

[0151] ISP-LPC変換部１２２Ｃは、第１実施形態と同様の処理手順により、サブフ

レーム毎に算出したISPパラメータをLPCパラメータに変換する（図２１のステップＳ４６）。

[0152] 以上のような第４実施形態では、符号化側において、不連続判定（一例として、隠蔽ISFパラメータと量子化ISFパラメータ同士のユークリッド距離を用いた不連続判定）を行い、その判定結果に関する補助情報を符号化して復号側へ出力し、復号側において、復号して得られた補助情報を用いて不連続性の推定を行う。このように符号化側と復号側とで連携しながら、符号化側での不連続判定結果に応じた適切な処理を実行することができる。

[0153] ［第５実施形態］

（符号化側について）

符号化器の機能構成は、第４実施形態に係る図１６の機能構成と同じであり、符号化器の処理フローは、第４実施形態に係る図１７の処理フローと同じである。ここでは、第４実施形態とは異なる第５実施形態におけるLP分析・符号化部について述べる。

[0154] 図２２にLP分析・符号化部の機能構成例を、図２３にLP分析・符号化部の処理フローを示す。図２２に示すように、LP分析・符号化部２１Ｓは、LP分析部２１０、LP-ISF変換部２１１、ISF符号化部２１２、不連続判定部２１３、ISF隠蔽部２１４、ISF-LP変換部２１５、及びISFバッファ２１６を備える。

[0155] このようなLP分析・符号化部２１Ｓにおいて、LP分析部２１０は、第４実施形態と同様の処理により、入力信号に対して線形予測分析を行い、線形予測係数を求める（図２３のステップＵ５１）。

[0156] LP-ISF変換部２１１は、第４実施形態と同様の処理により、算出した線形予測係数をISFパラメータに変換する（図２３のステップＵ５２）。線形予測係数からISFパラメータへの変換には、非特許文献に記載の方法を用いてもよい。

[0157] ISF符号化部２１２は、不連続判定部２１３の内部バッファ（不図示）に格納された不連続判定フラグを読み出す（図２３のステップＵ５３）。

[0158] ＜不連続判定フラグが不連続の検出を表す場合＞

ISF符号化部２１２は、次式で算出されるISF残差パラメータr_iをベクトル量子化してISF符号を算出する（図２３のステップＵ５４）。ここで、LP-ISF変換部で算出したISFパラメータをω_i、事前に学習で求めた平均ベクトルを$mean_i$とした。

[数79]

$$r_i = \omega_i - mean_i$$

[0159] 次に、ISF符号化部２１２は、ISF残差パラメータr_iを量子化して得られた量子化ISF残差パラメータ

[数80]

$$\hat{r}_i$$

を用いてISF残差パラメータバッファを次式に従い更新する（図２３のステップＵ５５）。

[数81]

$$\dot{r}_i^{-1} = \hat{r}_i$$

[0160] ＜不連続判定フラグが不連続の検出を表さない場合＞

ISF符号化部２１２は、次式で算出されるISF残差パラメータr_iをベクトル量子化してISF符号を算出する（図２３のステップＵ５４）。ここで、直前のフレームで復号により得られたISF残差パラメータを

[数82]

$$\dot{r}_i^{-1}$$

とした。

[数83]

$$r_i = \omega_i - mean_i - \frac{1}{3}\dot{r}_i^{-1}$$

[0161] 次に、ISF符号化部２１２は、ISF残差パラメータr_iを量子化して得られた量子化ISF残差パラメータ

[数84]

$$\hat{r}_i$$

を用いて、ISF残差パラメータバッファを次式に従い更新する（図２３のステップＵ５５）。

[数85]

$$\dot{r}_i^{-1} = \hat{r}_i$$

[0162] 以上の手順により、ISF符号化部２１２は、ISF符号を算出し、符号化の過程で得られる量子化ISFパラメータを、不連続判定部２１３、ISF隠蔽部２１４及びISF-LP変換部２１５に出力する。

[0163] ISF隠蔽部２１４は、第４実施形態と同様、デコーダ側のISF隠蔽部と同様の処理により、量子化ISFパラメータから隠蔽ISFパラメータを算出して、ISFバッファ２１６に出力する（図２３のステップＵ５６、Ｕ５８）。ISF隠蔽処理の処理手順は、デコーダ側のパケットロス隠蔽部と同じ処理であれば、どんな方法でもよい。

[0164] 不連続判定部２１３は、第４実施形態と同様の処理により、不連続の判定を行い、判定結果を不連続判定部２１３の内部バッファ（不図示）に格納する（図２３のステップＵ５７）。

[0165] ISF-LP変換部２１５は、第４実施形態と同様にして、上記量子化ISFパラメータを変換して、量子化線形予測係数を算出し、残差符号化部２２（図１６）へ出力する（図２３のステップＵ５８）。

[0166] （復号側について）

第５実施形態における音声信号処理装置は、前述した図１の音声復号器１と同様の構成を備え、音声符号復号部に特徴があるため、以下、音声符号復号部について説明する。図２４には音声符号復号部１２Ｅの機能構成例を、図２５にはLP係数の算出処理のフローを、それぞれ示す。図２４の音声符号復号部１２Ｅは、前述した図３の構成に対し、不連続検出部１２９が追加されている。

[0167] 不連続検出部１２９は、不連続判定符号を復号し、得られた不連続判定フラグをISF復号部１２０へ出力する（図２５のステップＳ５１）。

[0168] ISF復号部１２０は、不連続判定フラグの値に応じて、次の通りISFパラメータを算出し、ISFパラメータを安定性処理部１２１及び内部状態バッファ１４に出力する（図２５のステップＳ５２）。

[0169] ＜不連続判定フラグが不連続の検出を表す場合＞

ISF復号部１２０は、ISF符号を復号して得られる量子化ISF残差パラメータを

[数86]

$$\dot{r}_i$$

、事前に学習で求めた平均ベクトルを$mean_i$として、次式に従い量子化ISFパラメータ

[数87]

$$\dot{\omega}_i$$

を求める。

[数88]

$$\dot{\omega}_i = mean_i + \dot{r}_i$$

[0170] 次に、ISF復号部１２０は、内部状態バッファ１４に記憶されたISF残差パラメータを次式に従い更新する。

[数89]

$$\dot{r}_i^{-1} = \dot{r}_i$$

[0171] ＜不連続判定フラグが不連続の検出を表さない場合＞

ISF復号部１２０は、直前のフレームで復号により得られたISF残差パラメータ

[数90]

$$\dot{r}_i^{-1}$$

を内部状態バッファ１４から読み出し、

得られたISF残差パラメータ

[数91]

$$\dot{r}_i^{-1}$$

、事前に学習で求めた平均ベクトル$mean_i$、及びISF符号を復号して得られる量子化ISF残差パラメータ

[数92]

$$\dot{r}_i$$

から、次式に従い量子化ISFパラメータ

[数93]

$$\dot{\omega}_i$$

を求める。

[数94]

$$\dot{\omega}_i = mean_i + \dot{r}_i + \frac{1}{3}\dot{r}_i^{-1}$$

[0172] 次に、ISF復号部１２０は、内部状態バッファ１４に記憶されたISF残差パラメータを次式に従い更新する。

[数95]

$$\dot{r}_i^{-1} = \dot{r}_i$$

[0173] 安定性処理部１２１は、第１実施形態で述べた不連続が検出されない場合と同様の処理を行う（図２５のステップＳ５３）。

[0174] LP係数算出部１２２内のISF-ISP変換部１２２Ａは、第１実施形態と同様の処理手順によりISFパラメータをISPパラメータに変換する（図２５のステップＳ５４）。

[0175] ISP補間部１２２Ｂは、第１実施形態と同様の処理手順によりサブフレーム毎のISPパラメータを算出する（図２５のステップＳ５５）。

[0176] ISP-LPC変換部１２２Ｃは、第１実施形態と同様の処理手順により、サブフレーム毎に算出したISPパラメータをLPCパラメータに変換する（図２５のス

テップＳ５６）。

[0177] 以上のような第５実施形態では、符号化側において、不連続判定フラグが不連続の検出を表さない場合に、直前のフレームで復号により得られたISF残差パラメータを使用してISF残差パラメータのベクトル量子化を行い、一方、不連続判定フラグが不連続の検出を表す場合には、直前のフレームで復号により得られたISF残差パラメータの使用を回避する。同様に、復号側において、不連続判定フラグが不連続の検出を表さない場合に、直前のフレームで復号により得られたISF残差パラメータを使用して量子化ISFパラメータを算出し、一方、不連続判定フラグが不連続の検出を表す場合には、直前のフレームで復号により得られたISF残差パラメータの使用を回避する。このように符号化側と復号側とで連携しながら、不連続判定結果に応じた適切な処理を実行することができる。

[0178] ［第６実施形態］

上記第１～第５実施形態は組み合わせてもよい。例えば、第４実施形態に記載のとおり、復号側において、符号化側からの音声符号に含まれる不連続判定符号を復号して不連続を検出し、不連続が検出された場合に、以下のとおり実施してもよい。

[0179] 内部状態バッファに記憶されたISFパラメータ

[数96]

$$\dot{\omega}_i^{-1}$$

については、第３実施形態に記載のとおり、低次P’次元（0＜P’≦P）までのISFパラメータを次式により置き換える。

[数97]

$$\dot{\omega}_i^{-1} = \boldsymbol{\omega}_i^0 \qquad (0 \le i < P')$$

[0180] 一方、現フレームのISFパラメータについては、第５実施形態に記載のとおり、次式に従って算出する。

[数98]

$$\dot{\omega}_i = mean_i + \dot{r}_i$$

[0181] 以降、上記により求められたISFパラメータを用いて、第１実施形態と同様に、ISF-ISP変換部１２２Ａ、ISP補間部１２２Ｂ、ISP-LPC変換部１２２Ｃの処理によりLP係数を求める。

[0182] 上記のように第１～第５実施形態を任意に組み合わせた態様も有効である。

[0183] ［第７実施形態］

上記第１～第６実施形態及び変形例において、復号側におけるフレームロスの状況（例えば、単一フレームロスか連続フレームロスかといった状況）を加味してもよい。なお、第７実施形態では、不連続検出については、例えば音声符号に含まれる不連続判定符号を復号した結果を用いて不連続検出すればよく、その方法は上記に限定されない。

[0184] 第７実施形態における音声信号処理装置は、前述した図１の音声復号器１と同様の構成を備え、音声符号復号部に特徴があるため、以下、音声符号復号部について説明する。

[0185] 図２６には第７実施形態に係る音声復号器１Ｓの構成例を、図２７には音声復号器における処理のフロー図を、それぞれ示す。図２６に示すように、音声復号器１Ｓは、前述した音声符号復号部１２Ｇ、隠蔽信号生成部１３、及び内部状態バッファ１４に加え、過去数フレームにおけるパケット受信状態を判定しパケットロス履歴を記憶する受信状態判定部１６を備える。

[0186] 受信状態判定部１６は、パケット受信状態を判定し、判定結果に基づきパケットロス履歴情報を更新する（図２７のステップＳ５０）。

[0187] パケットロスが検出された場合（ステップＳ１００でＮＯの場合）、受信状態判定部１６は、当該フレームのパケットロス検出結果を隠蔽信号生成部１３に出力し、隠蔽信号生成部１３は前述した隠蔽信号の生成及び内部状態の更新を行う（ステップＳ３００、Ｓ４００）。なお、隠蔽信号生成部１３も、パケットロス履歴情報をパラメータの補間等に利用してもよい。

[0188] 一方、パケットロスが検出されなかった場合（ステップＳ１００でＹＥＳの場合）、受信状態判定部１６は、当該フレームのパケットロス検出結果を含むパケットロス履歴情報と、受信パケットに含まれる音声符号とを、音声符号復号部１２に出力し、音声符号復号部１２は前述した音声符号の復号及び内部状態の更新を行う（ステップＳ２００、Ｓ４００）。

[0189] 以後、ステップＳ５０～Ｓ４００の処理は通信終了まで（ステップＳ５００でＹＥＳと判断されるまで）繰り返される。

[0190] 図２８には、音声符号復号部１２Ｇの機能構成例を、図２９には、LP係数の算出処理のフロー図を示す。ここでは、パケットロス履歴情報をLP係数算出部１２２にのみ用いる例について述べるが、他の構成要素にパケットロス履歴情報を入力して用いる構成としてもよい。

[0191] 音声符号復号部１２Ｇでは、LP係数の算出処理に係る構成以外の構成については、第１実施形態と同様であるので、以下、LP係数の算出処理に係る構成及び動作について述べる。

[0192] ISF復号部１２０は、第１実施形態と同様にしてISF符号を復号してISFパラメータを安定性処理部１２１へ出力する（図２９のステップＳ７１）。

[0193] 不連続検出部１２９は、パケットロス履歴情報を参照して受信状態を判定する（ステップＳ７２）。不連続検出部１２９は、例えば、３フレーム前はパケットロス、２フレーム前は正常受信、１フレーム前はパケットロス、といったような特定の受信パターンを予め記憶しておき、該当する受信パターンが現れた場合に、受信状態フラグをオフにし、それ以外の場合は受信状態フラグをオンにするよう設計してもよい。

[0194] また、不連続検出部１２９は、前述した第１～第６実施形態のいずれかの方法と同様に、不連続検出を行う。

[0195] そして、安定性処理部１２１は、例えば以下のように、受信状態フラグおよび不連続検出の結果に応じて安定性処理を行う（ステップＳ７３）。

[0196] 受信状態フラグがオフの場合、安定性処理部１２１は、不連続検出の結果の如何を問わず、不連続が検出されなかった場合と同様の処理を行う。

[0197] 一方、受信状態フラグがオンで且つ不連続検出の結果、不連続が検出されなかった場合、安定性処理部１２１は、不連続が検出されなかった場合と同様の処理を行う。

[0198] さらに、受信状態フラグがオンで且つ不連続検出の結果、不連続が検出された場合、安定性処理部１２１は、不連続が検出された場合と同様の処理を行う。

[0199] 以後、LP係数算出部１２２内のISF-ISP変換部１２２Ａ、ISP補間部１２２Ｂ、及びISP-LPC変換部１２２Ｃの動作（ステップＳ７４～Ｓ７６）は、第１実施形態と同様である。

[0200] 以上のような第７実施形態では、不連続検出の結果と受信状態フラグの状態とに応じた安定性処理を行うことで、フレームロスの状況（例えば、単一フレームロスか連続フレームロスかといった状況）を加味した、より精度の良い処理を実行できる。

[0201] ［音声信号処理プログラムについて］

以下、コンピュータを、本発明に係る音声信号処理装置として動作させる音声信号処理プログラムについて説明する。

[0202] 図３２は、音声信号処理プログラムのさまざまな構成例を示す図である。図３０は、コンピュータのハードウェア構成例を示す図であり、図３１は、コンピュータの外観図である。図３２（ａ）～（ｄ）にそれぞれ示す音声信号処理プログラムＰ１～Ｐ４（以下「音声信号処理プログラムＰ」と総称する）は、図３１および図３２に示すコンピュータＣ１０を音声信号処理装置として動作させることができる。なお、本明細書にて説明する音声信号処理プログラムＰは、図３１および図３２に示すようなコンピュータに限定されず、携帯電話、携帯情報端末、携帯型パーソナルコンピュータといった任意の情報処理装置を、当該音声信号処理プログラムＰに従って動作させることができる。

[0203] 音声信号処理プログラムＰは、記録媒体Ｍに格納されて提供され得る。なお、記録媒体Ｍとしては、フレキシブルディスク、ＣＤ－ＲＯＭ、ＤＶＤ、

あるいはＲＯＭ等の記録媒体、あるいは半導体メモリ等が例示される。

[0204] 図３０に示すように、コンピュータＣ１０は、フレキシブルディスクドライブ装置、ＣＤ－ＲＯＭドライブ装置、ＤＶＤドライブ装置等の読み取り装置Ｃ１２と、作業用メモリ（ＲＡＭ）Ｃ１４と、記録媒体Ｍに記憶されたプログラムを記憶するメモリＣ１６と、ディスプレイＣ１８と、入力装置であるマウスＣ２０及びキーボードＣ２２と、データ等の送受信を行うための通信装置Ｃ２４と、プログラムの実行を制御する中央演算部（ＣＰＵ）Ｃ２６とを備える。

[0205] コンピュータＣ１０は、記録媒体Ｍが読み取り装置Ｃ１２に挿入されると、記録媒体Ｍに格納された音声信号処理プログラムＰに読み取り装置Ｃ１２からアクセス可能になり、音声信号処理プログラムＰによって音声信号処理装置として動作することが可能になる。

[0206] 音声信号処理プログラムＰは、図３１に示すように、搬送波に重畳されたコンピュータデータ信号Ｗとしてネットワークを介して提供されるものであってもよい。この場合、コンピュータＣ１０は、通信装置Ｃ２４によって受信した音声信号処理プログラムＰをメモリＣ１６に格納し、音声信号処理プログラムＰを実行することができる。

[0207] 音声信号処理プログラムＰは、図３２（ａ）～（ｄ）に示すさまざまな構成を採りうる。これらは、特許請求の範囲に記載した音声信号処理プログラムに係る請求項１８～２１の構成にそれぞれ対応する。例えば、図３２（ａ）に示す音声信号処理プログラムＰ１は、不連続推定モジュールＰ１１、および不連続修正モジュールＰ１２を備える。図３２（ｂ）に示す音声信号処理プログラムＰ２は、ISF/LSF量子化モジュールＰ２１、ISF/LSF隠蔽モジュールＰ２２、不連続推定モジュールＰ２３、および補助情報符号化モジュールＰ２４を備える。図３２（ｃ）に示す音声信号処理プログラムＰ３は、不連続推定モジュールＰ３１、補助情報符号化モジュールＰ３２、およびISF/LSF量子化モジュールＰ３３を備える。図３２（ｄ）に示す音声信号処理プログラムＰ４は、補助情報復号モジュールＰ４１、不連続修正モジュールＰ４

２、およびISF/LSF復号モジュールＰ４３を備える。

[0208] 以上説明したさまざまな実施形態によって、音声開始時点でのパケットロスから復帰した際に起こりうる不連続音を低減し、主観品質を改善することができる。

[0209] 発明の１つ目の特徴である安定性処理部は、パケットロス後に最初に正常に受信したパケットにおいて不連続が検出された場合に、例えばISFパラメータの各要素間に与える間隔を通常よりも大きくとることで、LP係数の利得が大きくなりすぎることを防ぐことができる。LP係数の利得と励振信号のパワーがともに増加することを防ぐことができるので、合成信号の不連続を軽減し、主観品質劣化を低減する。また、安定性処理部は、LP係数等を利用して算出したゲインを合成信号に乗算することにより合成信号の不連続性を軽減してもよい。

[0210] また、発明の２つ目の特徴である不連続検出器は、パケットロス後に最初に正常に受信したパケットに含まれる励振信号のゲインを監視し、励振信号のゲインにおいて一定以上増加が起こるパケットについて不連続性を推定する。

符号の説明

[0211] １、１Ｓ、１Ｘ…音声復号器、１１…パケットロス検出部、１２、１２Ａ、１２Ｂ、１２Ｄ、１２Ｅ、１２Ｇ、１２Ｓ…音声符号復号部、１３、１３Ｘ…隠蔽信号生成部、１４…内部状態バッファ、１５…正規化予測残差パワー復号部、１６…受信状態判定部、２１、２１Ｓ…分析・符号化部、２２…残差符号化部、２３…符号多重化部、１２０…ISF復号部、１２１、１２１Ｓ…安定性処理部、１２１Ｘ…ゲイン調整部、１２１Ｙ…ゲイン乗算部、１２２…LP係数算出部、１２２Ａ…ISF-ISP変換部、１２２Ｂ…ISP補間部、１２２Ｃ…ISP-LPC変換部、１２３…適応符号帳算出部、１２４…固定符号帳復号部、１２５…ゲイン復号部、１２６…励振ベクトル合成部、１２７…ポストフィルタ、１２８…合成フィルタ、１２９…不連続検出部、１３０…LP係数補間部、１３１…ピッチラグ補間部、１３２…ゲイン補間部、１３３…雑音

信号生成部、１３４…ポストフィルタ、１３５…合成フィルタ、１３６…適応符号帳算出部、１３７…励振ベクトル合成部、１３８…正規化予測残差調整部、２１０…LP分析部、２１１…LP-ISF変換部、２１２…ISF符号化部、２１３…不連続判定部、２１４…ISF隠蔽部、２１５…ISF-LP変換部、２１６…ISFバッファ、Ｃ１０…コンピュータ、Ｃ１２…読み取り装置、Ｃ１４…作業用メモリ、Ｃ１６…メモリ、Ｃ１８…ディスプレイ、Ｃ２０…マウス、Ｃ２２…キーボード、Ｃ２４…通信装置、Ｃ２６…ＣＰＵ、Ｍ…記録媒体、Ｐ１～Ｐ４…音声信号処理プログラム、Ｐ１１…不連続推定モジュール、Ｐ１２…不連続修正モジュール、Ｐ２１…ISF/LSF量子化モジュール、Ｐ２２…ISF/LSF隠蔽モジュール、Ｐ２３…不連続推定モジュール、Ｐ２４…補助情報符号化モジュール、Ｐ３１…不連続推定モジュール、Ｐ３２…補助情報符号化モジュール、Ｐ３３…ISF/LSF量子化モジュール、Ｐ４１…補助情報復号モジュール、Ｐ４２…不連続修正モジュール、Ｐ４３…ISF/LSF復号モジュール、Ｗ…コンピュータデータ信号。

請求の範囲

[請求項1] パケットロスが起こった後に最初に正常に受信した音声パケットに対して、音声パケット復号の結果得られる復号音声の振幅が急激に大きくなることにより生じる不連続性の推定を行う不連続推定器と、

復号音声の不連続性を修正する不連続修正器と、

を備える音声信号処理装置。

[請求項2] 前記不連続推定器は、前記復号音声の不連続性を、励振信号のパワーにより推定する、請求項１に記載の音声信号処理装置。

[請求項3] 前記不連続推定器は、前記復号音声の不連続性を、励振信号算出に用いる量子化された符号帳ゲインにより推定する、請求項１に記載の音声信号処理装置。

[請求項4] 前記音声信号処理装置は、符号化器より伝送された不連続性に関する補助情報を復号する補助情報復号器、をさらに備え、

前記不連続推定器は、前記補助情報復号器が補助情報符号を復号して出力する不連続性に関する補助情報を用いて、前記復号音声の不連続性を推定する、

請求項１に記載の音声信号処理装置。

[請求項5] 前記不連続修正器は、不連続性の推定結果に応じてISF/LSFパラメータを修正する、請求項１～４の何れか一項に記載の音声信号処理装置。

[請求項6] 前記不連続修正器は、合成フィルタの安定性保証のために与えるISF/LSFパラメータの各要素間の間隔を、前記不連続性の推定結果に応じて変化させる、請求項５に記載の音声信号処理装置。

[請求項7] 前記不連続修正器は、前記合成フィルタの安定性保証のために与えるISF/LSFパラメータの各要素間の間隔として、安定性保証のために通常与える間隔よりも大きな間隔を与える、請求項６に記載の音声信号処理装置。

[請求項8] 前記不連続修正器は、予め定めた次元までのISF/LSFパラメータを

等分して得られる間隔を、前記合成フィルタの安定性保証のために与えるISF/LSFパラメータの各要素間の間隔として用いる、請求項６に記載の音声信号処理装置。

[請求項9] 前記不連続修正器は、予め定めたベクトルにより、ISF/LSFパラメータの一部もしくは全部を置き換える、請求項６に記載の音声信号処理装置。

[請求項10] ISF/LSFパラメータを量子化するISF/LSF量子化器と、

ISF/LSFパラメータに関する隠蔽情報である隠蔽ISF/LSFパラメータを生成するISF/LSF隠蔽器と、

前記ISF/LSF量子化器の量子化過程で得られる量子化ISF/LSFパラメータと前記ISF/LSF隠蔽器が生成する隠蔽ISF/LSFパラメータとの距離を用いて、パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性を推定する不連続推定器と、

不連続性に関する補助情報を符号化する補助情報符号化器と、

を備える音声信号処理装置。

[請求項11] パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性を推定する不連続推定器と、

不連続性に関する補助情報を符号化する補助情報符号化器と、

前記不連続推定器が不連続性を推定しない場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF量子化に用い、前記不連続推定器が不連続性を推定した場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF量子化に用いることを回避するISF/LSF量子化器と、

を備える音声信号処理装置。

[請求項12] パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性に関する補助情報を復号し出力する補助情報復号器と、

復号音声の不連続性を修正する不連続修正器と、

前記補助情報復号器の出力により不連続性を推定しない場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF算出に用い、前記補助情報復号器の出力により不連続性を推定した場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF算出に用いることを回避するISF/LSF復号器と、

を備える音声信号処理装置。

[請求項13] 前記音声信号処理装置は、過去の所定数のフレームのパケット受信状態を判定する受信状態判定部、をさらに備え、

前記不連続修正器は、不連続性の推定結果に加え、前記パケット受信状態の判定結果も基礎として、不連続性の修正を行う、

請求項１～９および１２の何れか一項に記載の音声信号処理装置。

[請求項14] 音声信号処理装置により実行される音声信号処理方法であって、

パケットロスが起こった後に最初に正常に受信した音声パケットに対して、音声パケット復号の結果得られる復号音声の振幅が急激に大きくなることにより生じる不連続性の推定を行うステップと、

復号音声の不連続性を修正するステップと、

を備える音声信号処理方法。

[請求項15] 音声信号処理装置により実行される音声信号処理方法であって、

ISF/LSFパラメータを量子化するステップと、

ISF/LSFパラメータに関する隠蔽情報である隠蔽ISF/LSFパラメータを生成するステップと、

前記ISF/LSFパラメータの量子化過程で得られる量子化ISF/LSFパラメータと生成された隠蔽ISF/LSFパラメータとの距離を用いて、パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性を推定するステップと、

不連続性に関する補助情報を符号化するステップと、

を備える音声信号処理方法。

[請求項16] 音声信号処理装置により実行される音声信号処理方法であって、

パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性を推定するステップと、

不連続性に関する補助情報を符号化するステップと、

不連続性が推定されない場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF量子化に用い、不連続性が推定された場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF量子化に用いることを回避するステップと、

を備える音声信号処理方法。

[請求項17] 音声信号処理装置により実行される音声信号処理方法であって、

パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性に関する補助情報を復号し出力するステップと、

復号音声の不連続性を修正するステップと、

前記補助情報が不連続性の推定を示さない場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF算出に用い、前記補助情報が不連続性の推定を示す場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF算出に用いることを回避するステップと、

を備える音声信号処理方法。

[請求項18] コンピュータを、

パケットロスが起こった後に最初に正常に受信した音声パケットに対して、音声パケット復号の結果得られる復号音声の振幅が急激に大きくなることにより生じる不連続性の推定を行う不連続推定器と、

復号音声の不連続性を修正する不連続修正器、

として動作させるための音声信号処理プログラム。

[請求項19] コンピュータを、

ISF/LSFパラメータを量子化するISF/LSF量子化器と、

ISF/LSFパラメータに関する隠蔽情報である隠蔽ISF/LSFパラメータを生成するISF/LSF隠蔽器と、

前記ISF/LSF量子化器の量子化過程で得られる量子化ISF/LSFパラメータと前記ISF/LSF隠蔽器が生成する隠蔽ISF/LSFパラメータとの距離を用いて、パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性を推定する不連続推定器と、

不連続性に関する補助情報を符号化する補助情報符号化器、

として動作させるための音声信号処理プログラム。

[請求項20] コンピュータを、

パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性を推定する不連続推定器と、

不連続性に関する補助情報を符号化する補助情報符号化器と、

前記不連続推定器が不連続性を推定しない場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF量子化に用い、前記不連続推定器が不連続性を推定した場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF量子化に用いることを回避するISF/LSF量子化器、

として動作させるための音声信号処理プログラム。

[請求項21] コンピュータを、

パケットロスが起こった後に最初に正常に受信した音声パケットに生じる不連続性に関する補助情報を復号し出力する補助情報復号器と、

復号音声の不連続性を修正する不連続修正器と、

前記補助情報復号器の出力により不連続性を推定しない場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF算出に用い、前記補助情報復号器の出力により不連続性を推定した場合は、過去の量子化ISF/LSF残差パラメータを当該フレームにおけるISF/LSF算出に用いることを回避するISF/LSF復号器、

として動作させるための音声信号処理プログラム。