

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6259378号
(P6259378)

(45) 発行日 平成30年1月10日(2018.1.10)

(24) 登録日 平成29年12月15日(2017.12.15)

(51) Int.Cl.			F I		
G 1 O L	19/02	(2013.01)	G 1 O L	19/02	1 6 O A
G 1 O L	25/18	(2013.01)	G 1 O L	25/18	
G 1 O L	25/27	(2013.01)	G 1 O L	25/27	1 0 0
G 1 O L	19/06	(2013.01)	G 1 O L	19/06	Z

請求項の数 24 (全 53 頁)

(21) 出願番号	特願2014-171776 (P2014-171776)	(73) 特許権者	000004226 日本電信電話株式会社 東京都千代田区大手町一丁目5番1号
(22) 出願日	平成26年8月26日(2014.8.26)	(73) 特許権者	504137912 国立大学法人 東京大学 東京都文京区本郷七丁目3番1号
(65) 公開番号	特開2016-45462 (P2016-45462A)	(74) 代理人	100121706 弁理士 中尾 直樹
(43) 公開日	平成28年4月4日(2016.4.4)	(74) 代理人	100128705 弁理士 中村 幸雄
審査請求日	平成28年9月7日(2016.9.7)	(74) 代理人	100147773 弁理士 義村 宗洋
		(72) 発明者	守谷 健弘 東京都千代田区大手町一丁目5番1号 日 本電信電話株式会社内 最終頁に続く

(54) 【発明の名称】 周波数領域パラメータ列生成方法、周波数領域パラメータ列生成装置及びプログラム

(57) 【特許請求の範囲】

【請求項1】

pを1以上の整数とし、[1], [2], ..., [p]を所定の時間区間の時間領域の音信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

上記周波数領域パラメータ列 [1], [2], ..., [p]を入力として、変換後周波数領域パラメータ列~ [1], ~ [2], ..., ~ [p]を求めるパラメータ列変換ステップを含み、

上記パラメータ列変換ステップは、

上記変換後周波数領域パラメータ列~ [1], ~ [2], ..., ~ [p]における各変換後周波数領域パラメータ~ [i] (i=1, 2, ..., p) の値を、

[i]と [i]に近接する1つまたは複数の周波数領域パラメータとの値の関係に基づく線形変換により求める

周波数領域パラメータ列生成方法。

【請求項2】

請求項1に記載の周波数領域パラメータ列生成方法であって、

上記線形変換は、

上記周波数領域パラメータ列 [1], [2], ..., [p]よりも、上記変換後周波数領域パラメータ列のほうが、パラメータ値の間隔が均等間隔に近くなるか遠くなり、かつ、低い周波

数に対応するパラメータの方が高い周波数に対応するパラメータよりも上記周波数領域パラメータと上記変換後周波数領域パラメータとの差が小さくなるような線形変換である
周波数領域パラメータ列生成方法。

【請求項3】

請求項1に記載の周波数領域パラメータ列生成方法であって、

上記線形変換は、

上記周波数領域パラメータ列 [1], [2], ..., [p]よりも、上記変換後周波数領域パラメータ列のほうが、パラメータ値の間隔が均等間隔に近くなるか遠くなり、かつ、高い周波数に対応するパラメータの方が低い周波数に対応するパラメータよりも上記周波数領域パラメータと上記変換後周波数領域パラメータとの差が小さくなるような線形変換である

10

周波数領域パラメータ列生成方法。

【請求項4】

pを1以上の整数とし、[1], [2], ..., [p]を所定の時間区間の時間領域の音信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

上記周波数領域パラメータ列 [1], [2], ..., [p]を入力として、変換後周波数領域パラメータ列 \sim [1], \sim [2],..., \sim [p]を求めるパラメータ列変換ステップを含み、

上記パラメータ列変換ステップは、

20

上記変換後周波数領域パラメータ列 \sim [1], \sim [2],..., \sim [p]における各 \sim [i] ($i=1, 2, \dots, p$)を、

[i]が [i+1]と [i-1]との中点よりも [i+1]に近い場合には、 \sim [i]が \sim [i+1]と \sim [i-1]との中点よりも \sim [i+1]に近く、かつ、[i+1]-[i]よりも \sim [i+1]- \sim [i]の方が値が小さく、かつ、iの値が小さいほど \sim [i]-[i]の絶対値が小さくなるように求め、

[i]が [i+1]と [i-1]との中点よりも [i-1]に近い場合には、 \sim [i]が \sim [i+1]と \sim [i-1]との中点よりも \sim [i-1]に近く、かつ、[i]-[i-1]よりも \sim [i]- \sim [i-1]の方が値が小さく、かつ、iの値が小さいほど [i]- \sim [i]の絶対値が小さくなるように求める

30

周波数領域パラメータ列生成方法。

【請求項5】

pを1以上の整数とし、[1], [2], ..., [p]を所定の時間区間の時間領域の音信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

上記周波数領域パラメータ列 [1], [2], ..., [p]を入力として、変換後周波数領域パラメータ列 \sim [1], \sim [2],..., \sim [p]を求めるパラメータ列変換ステップを含み、

上記パラメータ列変換ステップは、

40

上記変換後周波数領域パラメータ列 \sim [1], \sim [2],..., \sim [p]における各 \sim [i] ($i=1, 2, \dots, p$)を、

[i]が [i+1]と [i-1]との中点よりも [i+1]に近い場合には、 \sim [i]が \sim [i+1]と \sim [i-1]との中点よりも \sim [i+1]に近く、かつ、[i+1]-[i]よりも \sim [i+1]- \sim [i]の方が値が小さく、かつ、iの値が大きいほど \sim [i]-[i]の絶対値が小さくなるように求め、

[i]が [i+1]と [i-1]との中点よりも [i-1]に近い場合には、 \sim [i]が \sim [i+1]と \sim [i-1]との中点よりも \sim [i-1]に近く、かつ、[i]-[i-1]よりも \sim [i]- \sim [i-1]の方が値が小さく、かつ、iの値が大きいほど [i]- \sim [i]の絶対値が小さくなるように求める

50

周波数領域パラメータ列生成方法。

【請求項 6】

p を 1 以上の整数とし、 $[1], [2], \dots, [p]$ を所定の時間区間の時間領域の音信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

上記周波数領域パラメータ列 $[1], [2], \dots, [p]$ を入力として、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を求めるパラメータ列変換ステップを含み、

上記パラメータ列変換ステップは、

上記変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ における各 $\sim [i]$ ($i=1, 2, \dots, p$) を、

$[i]$ が $[i+1]$ と $[i-1]$ との midpoint よりも $[i+1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との midpoint よりも $\sim [i+1]$ に近く、かつ、 $[i+1] - [i]$ よりも $\sim [i+1] - \sim [i]$ の方が値が大きく、かつ、 i の値が小さいほど $\sim [i] - [i]$ の絶対値が小さくなるように求め、

$[i]$ が $[i+1]$ と $[i-1]$ との midpoint よりも $[i-1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との midpoint よりも $\sim [i-1]$ に近く、かつ、 $[i] - [i-1]$ よりも $\sim [i] - \sim [i-1]$ の方が値が大きく、かつ、 i の値が小さいほど $[i] - \sim [i]$ の絶対値が小さくなるように求める

周波数領域パラメータ列生成方法。

【請求項 7】

p を 1 以上の整数とし、 $[1], [2], \dots, [p]$ を所定の時間区間の時間領域の音信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

上記周波数領域パラメータ列 $[1], [2], \dots, [p]$ を入力として、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を求めるパラメータ列変換ステップを含み、

上記パラメータ列変換ステップは、

上記変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ における各 $\sim [i]$ ($i=1, 2, \dots, p$) を、

$[i]$ が $[i+1]$ と $[i-1]$ との midpoint よりも $[i+1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との midpoint よりも $\sim [i+1]$ に近く、かつ、 $[i+1] - [i]$ よりも $\sim [i+1] - \sim [i]$ の方が値が大きく、かつ、 i の値が大きいほど $\sim [i] - [i]$ の絶対値が小さくなるように求め、

$[i]$ が $[i+1]$ と $[i-1]$ との midpoint よりも $[i-1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との midpoint よりも $\sim [i-1]$ に近く、かつ、 $[i] - [i-1]$ よりも $\sim [i] - \sim [i-1]$ の方が値が大きく、かつ、 i の値が大きいほど $[i] - \sim [i]$ の絶対値が小さくなるように求める

周波数領域パラメータ列生成方法。

【請求項 8】

p を 1 以上の整数とし、 $a[1], a[2], \dots, a[p]$ を所定の時間区間の時間領域の音信号に対応する線形予測係数列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

$[1], [2], \dots, [p]$ を、

上記線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来する L S P パラメータ列、

上記線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来する L S F パラメータ列、

10

20

30

40

50

上記線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来し、かつ、 $[1], [2], \dots, [p]$ の全てが0から までの間に存在し、かつ、線形予測係数列に含まれる全ての線形予測係数が0である場合には $[1], [2], \dots, [p]$ が0から までの間に均等間隔に存在する周波数領域パラメータ列、のいずれかとし、

1及び 2をそれぞれ1以下の正の定数である補正係数とし、 K を予め定められた $p \times p$ の帯行列とし、

次式により定義される変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を生成するパラメータ列変換ステップを含む

【数41】

$$\begin{pmatrix} \tilde{\omega}[1] \\ \tilde{\omega}[2] \\ \vdots \\ \tilde{\omega}[p] \end{pmatrix} = K \begin{pmatrix} \omega[1] - \frac{\pi}{p+1} \\ \omega[2] - \frac{2\pi}{p+1} \\ \vdots \\ \omega[p] - \frac{p\pi}{p+1} \end{pmatrix} (\gamma_2 - \gamma_1) + \begin{pmatrix} \omega[1] \\ \omega[2] \\ \vdots \\ \omega[p] \end{pmatrix}$$

10

周波数領域パラメータ列生成方法。

20

【請求項9】

p を1以上の整数とし、 $a[1], a[2], \dots, a[p+1]$ を所定の時間区間の時間領域の音信号に対応する $(p+1)$ 次の線形予測係数列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

$[1], [2], \dots, [p+1]$ を、

上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来するISPパラメータ列、

上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来するISFパラメータ列、

上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来し、かつ、 $[1], [2], \dots, [p]$ の全てが0から までの間に存在し、かつ、線形予測係数列に含まれる全ての線形予測係数が0である場合には $[1], [2], \dots, [p]$ が0から までの間に均等間隔に存在する周波数領域パラメータ列、のいずれかとし、

30

1及び 2をそれぞれ1以下の正の定数である補正係数とし、 K を予め定められた $p \times p$ の帯行列とし、

次式により定義される変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を生成するパラメータ列変換ステップを含む

【数42】

$$\begin{pmatrix} \tilde{\omega}[1] \\ \tilde{\omega}[2] \\ \vdots \\ \tilde{\omega}[p] \end{pmatrix} = K \begin{pmatrix} \omega[1] - \frac{\pi}{p+1} \\ \omega[2] - \frac{2\pi}{p+1} \\ \vdots \\ \omega[p] - \frac{p\pi}{p+1} \end{pmatrix} (\gamma_2 - \gamma_1) + \begin{pmatrix} \omega[1] \\ \omega[2] \\ \vdots \\ \omega[p] \end{pmatrix}$$

40

周波数領域パラメータ列生成方法。

【請求項10】

50

請求項 1 から 8 のいずれかに記載の周波数領域パラメータ列生成方法であって、

1 を 1 以下の正の定数とし、

上記線形予測係数列を $a[1], a[2], \dots, a[p]$ とし、

上記周波数領域パラメータ列 $[1], [2], \dots, [p]$ における各 $[i]$ ($i=1, 2, \dots, p$) は

$a_{-1}[i]=a[i] \times (1)^i$ として、

$a_{-1}[1], a_{-1}[2], \dots, a_{-1}[p]$ と等価な周波数領域のパラメータまたはその量子化値、

または、

上記音信号に対応する隣接するサンプル点間の周波数の間隔が均等な周波数領域の係数列を線形離散化サンプル点列とし、

上記音信号に対応する隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列を非線形離散化サンプル点列とし、

上記線形離散化サンプル点列と上記非線形離散化サンプル点列のサンプル点数を N とし、

上記非線形離散化サンプル点列で k 番目のインデックスに対応するサンプル点の周波数の

、上記線形離散化サンプル点列での相対的な周波数位置を $f(k)$ とし、

【数 4 3】

$$g(k) = \begin{cases} f(k)/k & \text{for } 0 < k < N-1 \\ 1 & \text{for } k = 0 \end{cases}$$

とし、

$a_{-1}[i]=a[i] \times (1)^{ig(k)}$ として、

$a_{-1}[1], a_{-1}[2], \dots, a_{-1}[p]$ と等価な周波数領域のパラメータまたはその量子化値、

である

周波数領域パラメータ列生成方法。

【請求項 1 1】

請求項 8 または 9 に記載の周波数領域パラメータ列生成方法であって、

上記帯行列 K は、対角要素が 0 以上の値であり、行方向において対角要素に隣接する要素が 0 以下の値であり、対角要素の左上の成分の絶対値は右下の成分の絶対値よりも大きい

周波数領域パラメータ列生成方法。

【請求項 1 2】

請求項 8 または 9 に記載の周波数領域パラメータ列生成方法であって、

上記帯行列 K は、対角要素が 0 以上の値であり、行方向において対角要素に隣接する要素が 0 以下の値であり、対角要素の左上の成分の絶対値は右下の成分の絶対値よりも小さい

周波数領域パラメータ列生成方法。

【請求項 1 3】

請求項 1 から 8、10 のいずれかに記載の周波数領域パラメータ列生成方法であって、

上記所定の時間区間の音信号に対応するパワースペクトル系列であって、隣接するサンプル点間の周波数の間隔が不均等な周波数領域のパワースペクトル系列を得るパワースペクトル系列生成ステップと、

上記パワースペクトル系列生成ステップが得たパワースペクトル系列を用いて、上記パワースペクトル系列がサンプル点の周波数の間隔が均等であると見做した線形予測分析を行い、上記線形予測係数列 $a[1], a[2], \dots, a[p]$ を得る線形予測係数算出ステップと、

上記線形予測係数算出ステップが得た上記線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来する L S P パラメータ列、上記線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来する L S F パラメータ列、上記線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来し、かつ、 $[1], [2], \dots, [p]$ の全てが 0 から \quad までの間に存在し、かつ、線形予測係数列に含まれる全ての線形予測係数が 0 である場合には $[1], [2], \dots, [p]$ が 0 から \quad までの間に均等間隔に存在する周波数領域パラメータ列、のいずれかを上記 $[1], [2], \dots, [p]$ として求める周波数領域パラメータ列計算ステップと、

10

20

30

40

50

を更に含む周波数領域パラメータ列生成方法。

【請求項 14】

請求項 9 に記載の周波数領域パラメータ列生成方法であって、

上記所定の時間区間の音信号に対応するパワースペクトル系列であって、隣接するサンプル点間の周波数の間隔が不均等な周波数領域のパワースペクトル系列を得るパワースペクトル系列生成ステップと、

上記パワースペクトル系列生成ステップが得たパワースペクトル系列を用いて、上記パワースペクトル系列がサンプル点の周波数の間隔が均等であると見做した線形予測分析を行い、上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ を得る線形予測係数算出ステップと、

上記線形予測係数算出ステップが得た上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来する I S P パラメータ列、上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来する I S F パラメータ列、上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来し、かつ、 $[1], [2], \dots, [p]$ の全てが 0 から までの間に存在し、かつ、線形予測係数列に含まれる全ての線形予測係数が 0 である場合には $[1], [2], \dots, [p]$ が 0 から までの間に均等間隔に存在する周波数領域パラメータ列、のいずれかを上記 $[1], [2], \dots, [p+1]$ として求める周波数領域パラメータ列計算ステップと、

を更に含む周波数領域パラメータ列生成方法。

【請求項 15】

p を 1 以上の整数とし、 $[1], [2], \dots, [p]$ を所定の時間区間の時間領域の音信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

上記周波数領域パラメータ列 $[1], [2], \dots, [p]$ を入力として、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を求めるパラメータ列変換部を含み、

上記パラメータ列変換部は、

上記変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ における各変換後周波数領域パラメータ $\sim [i]$ ($i=1, 2, \dots, p$) の値を、

$[i]$ と $[i]$ に近接する 1 つまたは複数の周波数領域パラメータとの値の関係に基づく線形変換により求める

周波数領域パラメータ列生成装置。

【請求項 16】

p を 1 以上の整数とし、 $[1], [2], \dots, [p]$ を所定の時間区間の時間領域の音信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

上記周波数領域パラメータ列 $[1], [2], \dots, [p]$ を入力として、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を求めるパラメータ列変換部を含み、

上記パラメータ列変換部は、

上記変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ における各 $\sim [i]$ ($i=1, 2, \dots, p$) を、

$[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i+1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i+1]$ に近く、かつ、 $[i+1] - [i]$ よりも $\sim [i+1] - \sim [i]$ の方が値が小さく、かつ、 i の値が小さいほど $\sim [i] - [i]$ の絶対値が小さくなるように求め、

$[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i-1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i-1]$ に近く、かつ、 $[i] - [i-1]$ よりも $\sim [i] - \sim [i-1]$ の方が値が小さく、かつ、 i の値が小さいほど $[i] - \sim [i]$ の絶対値が小さくなるように求める

10

20

30

40

50

周波数領域パラメータ列生成装置。

【請求項 17】

pを1以上の整数とし、[1], [2], ..., [p]を所定の時間区間の時間領域の音信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

上記周波数領域パラメータ列 [1], [2], ..., [p]を入力として、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を求めるパラメータ列変換部を含み、

上記パラメータ列変換部は、

上記変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ における各 $\sim [i]$ ($i=1, 2, \dots, p$)を、

[i]が [i+1]と [i-1]との中点よりも [i+1]に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i+1]$ に近く、かつ、[i+1]- [i]よりも $\sim [i+1]-\sim [i]$ の方が値が小さく、かつ、iの値が大きいくほど $\sim [i]- [i]$ の絶対値が小さくなるように求め、

[i]が [i+1]と [i-1]との中点よりも [i-1]に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i-1]$ に近く、かつ、[i]- [i-1]よりも $\sim [i]-\sim [i-1]$ の方が値が小さく、かつ、iの値が大きいくほど [i]- $\sim [i]$ の絶対値が小さくなるように求める

周波数領域パラメータ列生成装置。

【請求項 18】

pを1以上の整数とし、[1], [2], ..., [p]を所定の時間区間の時間領域の音信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

上記周波数領域パラメータ列 [1], [2], ..., [p]を入力として、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を求めるパラメータ列変換部を含み、

上記パラメータ列変換部は、

上記変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ における各 $\sim [i]$ ($i=1, 2, \dots, p$)を、

[i]が [i+1]と [i-1]との中点よりも [i+1]に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i+1]$ に近く、かつ、[i+1]- [i]よりも $\sim [i+1]-\sim [i]$ の方が値が大きく、かつ、iの値が小さいほど $\sim [i]- [i]$ の絶対値が小さくなるように求め、

[i]が [i+1]と [i-1]との中点よりも [i-1]に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i-1]$ に近く、かつ、[i]- [i-1]よりも $\sim [i]-\sim [i-1]$ の方が値が大きく、かつ、iの値が小さいほど [i]- $\sim [i]$ の絶対値が小さくなるように求める

周波数領域パラメータ列生成装置。

【請求項 19】

pを1以上の整数とし、[1], [2], ..., [p]を所定の時間区間の時間領域の音信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

上記周波数領域パラメータ列 [1], [2], ..., [p]を入力として、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を求めるパラメータ列変換部を含み、

上記パラメータ列変換部は、

10

20

30

40

50

上記変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ における各 $\sim [i]$ ($i=1, 2, \dots, p$)を、

$[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i+1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i+1]$ に近く、かつ、 $[i+1]-[i]$ よりも $\sim [i+1]-\sim [i]$ の方が値が大きく、かつ、 i の値が大きいくほど $\sim [i]-[i]$ の絶対値が小さくなるように求め、

$[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i-1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i-1]$ に近く、かつ、 $[i]-[i-1]$ よりも $\sim [i]-\sim [i-1]$ の方が値が大きく、かつ、 i の値が大きいくほど $[i]-\sim [i]$ の絶対値が小さくなるように求める

10

周波数領域パラメータ列生成装置。

【請求項 20】

p を1以上の整数とし、 $a[1], a[2], \dots, a[p]$ を所定の時間区間の時間領域の音信号に対応する線形予測係数列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

$[1], [2], \dots, [p]$ を、

上記線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来するL S Pパラメータ列、

上記線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来するL S Fパラメータ列、

20

上記線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来し、かつ、 $[1], [2], \dots, [p]$ の全てが0から までの間に存在し、かつ、線形予測係数列に含まれる全ての線形予測係数が0である場合には $[1], [2], \dots, [p]$ が0から までの間に均等間隔に存在する周波数領域パラメータ列、のいずれかとし、

1及び 2をそれぞれ1以下の正の定数である補正係数とし、 K を予め定められた $p \times p$ の帯行列とし、

次式により定義される変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を生成するパラメータ列変換部を含む

【数 4 4】

$$\begin{pmatrix} \tilde{\omega}[1] \\ \tilde{\omega}[2] \\ \vdots \\ \tilde{\omega}[p] \end{pmatrix} = K \begin{pmatrix} \omega[1] - \frac{\pi}{p+1} \\ \omega[2] - \frac{2\pi}{p+1} \\ \vdots \\ \omega[p] - \frac{p\pi}{p+1} \end{pmatrix} (\gamma_2 - \gamma_1) + \begin{pmatrix} \omega[1] \\ \omega[2] \\ \vdots \\ \omega[p] \end{pmatrix}$$

30

周波数領域パラメータ列生成装置。

40

【請求項 21】

p を1以上の整数とし、 $a[1], a[2], \dots, a[p+1]$ を所定の時間区間の時間領域の音信号に対応する $(p+1)$ 次の線形予測係数列とし、

上記線形予測係数列は、隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做した線形予測分析により得られるものであり、

$[1], [2], \dots, [p+1]$ を、

上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来するI S Pパラメータ列、

上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来するI S Fパラメータ列、

上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来し、かつ、 $[1], [2], \dots, [p]$ の全て

50

が 0 から までの間に存在し、かつ、線形予測係数列に含まれる全ての線形予測係数が 0 である場合には [1], [2]..., [p]が 0 から までの間に均等間隔に存在する周波数領域パラメータ列、のいずれかとし、

1及び 2をそれぞれ 1 以下の正の定数である補正係数とし、Kを予め定められた $p \times p$ の帯行列とし、

次式により定義される変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を生成するパラメータ列変換部を含む

【数 4 5】

$$\begin{pmatrix} \tilde{\omega}[1] \\ \tilde{\omega}[2] \\ \vdots \\ \tilde{\omega}[p] \end{pmatrix} = K \begin{pmatrix} \omega[1] - \frac{\pi}{p+1} \\ \omega[2] - \frac{2\pi}{p+1} \\ \vdots \\ \omega[p] - \frac{p\pi}{p+1} \end{pmatrix} (\gamma_2 - \gamma_1) + \begin{pmatrix} \omega[1] \\ \omega[2] \\ \vdots \\ \omega[p] \end{pmatrix}$$

10

周波数領域パラメータ列生成装置。

【請求項 2 2】

20

請求項 1 5 から 2 0 のいずれかに記載の周波数領域パラメータ列生成装置であって、
上記所定の時間区間の音信号に対応するパワースペクトル系列であって、隣接するサンプル点間の周波数の間隔が不均等な周波数領域のパワースペクトル系列を得るパワースペクトル系列生成部と、

上記パワースペクトル系列生成部が得たパワースペクトル系列を用いて、上記パワースペクトル系列がサンプル点の周波数の間隔が均等であると見做した線形予測分析を行い、上記線形予測係数列 $a[1], a[2], \dots, a[p]$ を得る線形予測係数算出部と、

上記線形予測係数算出部が得た上記線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来する L S P パラメータ列、上記線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来する L S F パラメータ列、上記線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来し、かつ、 [1], [2]..., [p]の全てが 0 から までの間に存在し、かつ、線形予測係数列に含まれる全ての線形予測係数が 0 である場合には [1], [2]..., [p]が 0 から までの間に均等間隔に存在する周波数領域パラメータ列、のいずれかを上記 [1], [2]..., [p]として求める周波数領域パラメータ列計算部と、

30

を更に含む周波数領域パラメータ列生成装置。

【請求項 2 3】

請求項 2 1 に記載の周波数領域パラメータ列生成装置であって、
上記所定の時間区間の音信号に対応するパワースペクトル系列であって、隣接するサンプル点間の周波数の間隔が不均等な周波数領域のパワースペクトル系列を得るパワースペクトル系列生成部と、

40

上記パワースペクトル系列生成部が得たパワースペクトル系列を用いて、上記パワースペクトル系列がサンプル点の周波数の間隔が均等であると見做した線形予測分析を行い、上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ を得る線形予測係数算出部と、

上記線形予測係数算出部が得た上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来する I S P パラメータ列、上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来する I S F パラメータ列、上記線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来し、かつ、 [1], [2]..., [p]の全てが 0 から までの間に存在し、かつ、線形予測係数列に含まれる全ての線形予測係数が 0 である場合には [1], [2]..., [p]が 0 から までの間に均等間隔に存在する周波数領域パラメータ列、のいずれかを上記 [1], [2]..., [p+1]として求める周波数領域パラメータ列計算部と、

50

を更に含む周波数領域パラメータ列生成装置。

【請求項 2 4】

請求項 1 から 1 4 のいずれかに記載の周波数領域パラメータ列生成方法の各ステップをコンピュータに実行させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、音信号の符号化技術などの信号処理技術において、音信号に由来する周波数領域のサンプル列を、当該周波数領域のサンプル列におけるサンプル点の周波数領域での間隔を伸縮した系列を生成する技術に関する。

10

【背景技術】

【0002】

低ビット（例えば10kbit/s～20kbit/s程度）の音信号の符号化方法として、DFT（離散フーリエ変換）やMDCT（変形離散コサイン変換）などの周波数領域での直交変換係数に対する適応符号化が知られている。例えば標準規格技術であるMPEG USAC(Unified Speech and Audio Coding)は、TCX(transform coded excitation:変換符号化励振)符号化モードを持ち、この中ではMDCT係数をフレームごとに正規化して量子化後に可変長符号化している（例えば、非特許文献1参照）。

【0003】

従来のTCXに基づく符号化装置の構成例を図1に示す。以下、図1の各部について説明する。

20

【0004】

<周波数領域変換部11>

周波数領域変換部11には、時間領域の音信号が入力される。音信号は、例えば音声信号又は音響信号である。

【0005】

周波数領域変換部11は、所定の時間長のフレーム単位で、入力された時間領域の音信号を周波数領域のN点のMDCT係数列 $X(0), X(1), \dots, X(N-1)$ に変換する。Nは正の整数である。

【0006】

変換されたMDCT係数列 $X(0), X(1), \dots, X(N-1)$ は、包絡正規化部14に出力される。

30

【0007】

<線形予測分析部12>

線形予測分析部12には、時間領域の音信号が入力される。

【0008】

線形予測分析部12は、フレーム単位で入力された音信号に対する線形予測分析を行うことにより、線形予測係数 $\hat{c}_1, \hat{c}_2, \dots, \hat{c}_p$ を生成する。また、線形予測分析部12は、生成された線形予測係数 $\hat{c}_1, \hat{c}_2, \dots, \hat{c}_p$ を符号化して線形予測係数符号を生成する。線形予測係数符号の例は、線形予測係数 $\hat{c}_1, \hat{c}_2, \dots, \hat{c}_p$ に対応するLSP(Line Spectrum Pairs)パラメータ列の量子化値の列に対応する符号であるLSP符号である。pは2以上の整数である。

40

【0009】

また、線形予測分析部12は、生成された線形予測係数符号に対応する線形予測係数である量子化線形予測係数 $\hat{c}_1, \hat{c}_2, \dots, \hat{c}_p$ を生成する。

【0010】

生成された量子化線形予測係数 $\hat{c}_1, \hat{c}_2, \dots, \hat{c}_p$ は、パワースペクトル包絡系列生成部13に出力される。また、生成された線形予測係数符号は、復号装置に出力される。

【0011】

<パワースペクトル包絡系列生成部13>

パワースペクトル包絡系列生成部13には、線形予測分析部12が生成した量子化線形

50

予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ が入力される。

【0012】

パワースペクトル包絡系列生成部13は、量子化線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ を用いて、以下の式(P1)により定義される平滑化パワースペクトル包絡系列 $\hat{W}(0), \hat{W}(1), \dots, \hat{W}(N-1)$ を生成する。 \cdot を実数として $\exp(\cdot)$ はネイピア数を底とする指数関数、 j は虚数単位、 σ^2 は予測残差エネルギーである。 γ は、1以下の正の定数であり、以下の式(P1')により定義されるパワースペクトル包絡系列 $\hat{W}(0), \hat{W}(1), \dots, \hat{W}(N-1)$ の振幅の凹凸を鈍らせる係数、言い換えればパワースペクトル包絡系列を平滑化する係数である。

【数1】

$$\hat{W}_\gamma(k) = \frac{\sigma^2}{2\pi} \frac{1}{\left| 1 + \sum_{n=1}^p \hat{\alpha}_n \gamma^n \exp(-j2\pi kn/N) \right|^2} \quad \dots(P1)$$

$$\hat{W}(k) = \frac{\sigma^2}{2\pi} \frac{1}{\left| 1 + \sum_{n=1}^p \hat{\alpha}_n \exp(-j2\pi kn/N) \right|^2} \quad \dots(P1')$$

【0013】

生成された平滑化パワースペクトル包絡系列 $\hat{W}(0), \hat{W}(1), \dots, \hat{W}(N-1)$ は、包絡正規化部14に出力される。

【0014】

<包絡正規化部14>

包絡正規化部14には、周波数領域変換部11が生成したMDCT係数列 $X(0), X(1), \dots, X(N-1)$ 及びパワースペクトル包絡系列生成部13が出力した平滑化パワースペクトル包絡系列 $\hat{W}(0), \hat{W}(1), \dots, \hat{W}(N-1)$ が入力される。

【0015】

包絡正規化部14は、MDCT係数列の各係数 $X(i)$ を平滑化パワースペクトル包絡系列の各値 $\hat{W}(i)$ の平方根で正規化することにより、正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ を生成する。つまり、 $X_N(i) = X(i) / \sqrt{\hat{W}(i)}$ [$i=0, 1, \dots, N-1$]である。ここで、 $\sqrt{\cdot}$ は \cdot の平方根を表す。

【0016】

生成された正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ は、符号化部15に出力される。

【0017】

ここでは、聴覚的に歪が小さくなるような量子化の実現のために、包絡正規化部14は、パワースペクトル包絡を鈍らせたパワースペクトル包絡の系列である平滑化パワースペクトル包絡系列 $\hat{W}(0), \hat{W}(1), \dots, \hat{W}(N-1)$ を用いて、フレーム単位でMDCT係数列 $X(0), X(1), \dots, X(N-1)$ を正規化している。

【0018】

この結果、生成される正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ は、入力されたMDCT係数列 $X(0), X(1), \dots, X(N-1)$ ほどの大きな振幅の傾きや振幅の凹凸を持たないが、入力された音信号のパワースペクトル包絡系列と類似の大小関係を有するもの、すなわち低い周波数に対応する係数側の領域にやや大きな振幅を持ちピッチ周期に起因する微細構造を持つものとなる。

【0019】

<符号化部15>

符号化部15には、包絡正規化部14が生成した正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ が入力される。

【0020】

10

20

30

40

50

符号化部 15 は、正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ に対応する符号を生成する。

【0021】

生成された正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ に対応する符号は、復号装置に出力される。

【0022】

正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ の各係数を利得（グローバルゲイン） g で割り算し、その結果を量子化した整数値による系列である量子化正規化済係数系列 $X_Q(0), X_Q(1), \dots, X_Q(N-1)$ を符号化して得られる符号を整数信号符号とする。非特許文献1の技術では、符号化部15は、この整数信号符号のビット数が、予め配分されたビット数である配分ビット数 B 以下、かつ、なるべく大きな値となるような利得 g を決定する。そして、符号化部15は、この決定された利得 g に対応する利得符号と、この利得符号に対応する利得 g で正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ の各係数を割り算した結果を量子化した整数値による系列である量子化正規化済係数系列 $X_Q(0), X_Q(1), \dots, X_Q(N-1)$ を符号化して得られる整数信号符号とを生成する。

10

【0023】

この生成された利得符号及び整数信号符号が、正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ に対応する符号として復号装置に出力される。

【0024】

以上のように、従来のTCXに基づく符号化では、パワースペクトル包絡を鈍らせた平滑化パワースペクトル包絡系列を用いてMDCT係数列を正規化した後、正規化MDCT係数列を符号化している。この符号化方法は、上記のMPEG-4 USACなどで採用されている。

20

【先行技術文献】

【非特許文献】

【0025】

【非特許文献1】M. Neuendorf, et al., "MPEG Unified Speech and Audio Coding - The ISO/MPEG Standard for High-Efficiency Audio Coding of all Content Types," AES 132nd Convention, Budapest, Hungary, 2012.

【発明の概要】

【発明が解決しようとする課題】

30

【0026】

線形予測係数を用いて得られるパワースペクトル包絡は、おおよそ（信号のサンプル数）/（線形予測次数）の解像度で元のスペクトルを表現するものである。そして、この解像度は周波数領域において均一である。すなわち、MPEG USACなどの従来のTCXに基づく符号化におけるMDCT係数列の正規化に用いるパワースペクトル包絡系列は、周波数領域の周波数軸で均等な間隔で、言い換えれば周波数方向の均一な解像度で離散化（以下、「線形離散化」ともいう）した包絡の値であった。

【0027】

通常の音声や音楽の信号は特定の周波数領域（例えば低周波数領域）にエネルギーが集中する場合が多く、エネルギーが集中している周波数領域ではパワースペクトル包絡の変化が大きい傾向がある。全周波数領域において均一な解像度で離散化した包絡の値を用いると、エネルギーが集中している周波数領域では周波数方向の解像度が不足して、得られるパワースペクトル包絡系列は元のパワースペクトル包絡の振幅の凹凸の変化を十分な精度で表現できないことがある。このようなパワースペクトル包絡系列を用いてMDCT係数列を正規化すると、解像度が不足した部分でのMDCT係数列とパワースペクトル包絡系列との差が大きくなり、正規化MDCT係数列の値のばらつきが大きくなってしまいうため、符号化効率が低下する可能性があった。

40

【0028】

ここで、予測次数を増やせば、線形離散化でも解像度は高くすることはできるが、パラメータの情報量が増加して、符号化効率が低下する可能性がある。また、特定のフレーム

50

だけに次数を増やすと、フレーム間の処理の連続性のため処理が煩雑となる可能性がある。

【0029】

符号化処理に限らず、音信号の信号処理においては、音信号に由来する周波数領域のサンプル列として、周波数方向の不均一な解像度で音信号を離散化したサンプル列を用いることにより信号処理の精度が向上する場合がある。

【0030】

この発明は、このような技術的背景に鑑みて、少ない演算量の増加で符号化効率を改善する技術を提供することを目的とする。また、符号化以外の信号処理において、少ない演算量の増加で信号処理の精度を改善する技術を提供することを目的とする。

10

【課題を解決するための手段】

【0031】

上記の課題を解決するために、この発明の第一の態様の周波数領域パラメータ列生成方法は、 p を1以上の整数とし、 $[1], [2], \dots, [p]$ を所定の時間区間の音信号に対応する隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做したときの時間領域の信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、周波数領域パラメータ列 $[1], [2], \dots, [p]$ を入力として、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を求めるパラメータ列変換ステップを含む。パラメータ列変換ステップは、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ における各 $\sim [i]$ ($i=1, 2, \dots, p$)を、 $[i]$ と $[i]$ に近接する1つまたは複数の周波数領域パラメータとの値の関係に基づく線形変換により変換後周波数領域パラメータ $\sim [i]$ の値を求める。

20

【0032】

この発明の第二の態様の周波数領域パラメータ列生成方法は、 p を1以上の整数とし、 $[1], [2], \dots, [p]$ を所定の時間区間の音信号に対応する隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做したときの時間領域の信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、周波数領域パラメータ列 $[1], [2], \dots, [p]$ を入力として、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を求めるパラメータ列変換ステップを含む。パラメータ列変換ステップは、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ における各 $\sim [i]$ ($i=1, 2, \dots, p$)を、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i+1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i+1]$ に近く、かつ、 $[i+1]-[i]$ よりも $\sim [i+1]-\sim [i]$ の方が値が小さく、かつ、 i の値が小さいほど $[i]-[i]$ の絶対値が小さくなるように求め、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i-1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i-1]$ に近く、かつ、 $[i]-[i-1]$ よりも $\sim [i]-\sim [i-1]$ の方が値が小さく、かつ、 i の値が小さいほど $[i]-\sim [i]$ の絶対値が小さくなるように求める。

30

【0033】

この発明の第三の態様の周波数領域パラメータ列生成方法は、 p を1以上の整数とし、 $[1], [2], \dots, [p]$ を所定の時間区間の音信号に対応する隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做したときの時間領域の信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、周波数領域パラメータ列 $[1], [2], \dots, [p]$ を入力として、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を求めるパラメータ列変換ステップを含む。パラメータ列変換ステップは、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ における各 $\sim [i]$ ($i=1, 2, \dots, p$)を、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i+1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i+1]$ に近く、かつ、 $[i+1]-[i]$ よりも $\sim [i+1]-\sim [i]$ の方が値が小さく、かつ、 i の値が大きいほど $[i]-[i]$ の絶対値が小さくなるように求め、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i-1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i-1]$ に近く、かつ、 $[i]-$

40

50

$[i-1]$ よりも $\sim [i]$ 、 $\sim [i-1]$ の方が値が小さく、かつ、 i の値が大きいほど $[i]$ 、 $\sim [i]$ の絶対値が小さくなるように求める。

【0034】

この発明の第四の態様の周波数領域パラメータ列生成方法は、 p を1以上の整数とし、 $[1], [2], \dots, [p]$ を所定の時間区間の音信号に対応する隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であるときに見做したときの時間領域の信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、周波数領域パラメータ列 $[1], [2], \dots, [p]$ を入力として、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を求めるパラメータ列変換ステップを含む。パラメータ列変換ステップは、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ における各 $\sim [i]$ ($i=1, 2, \dots, p$)を、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i+1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i+1]$ に近く、かつ、 $[i+1]$ 、 $[i]$ よりも $\sim [i+1]$ 、 $\sim [i]$ の方が値が大きく、かつ、 i の値が小さいほど $\sim [i]$ 、 $\sim [i]$ の絶対値が小さくなるように求め、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i-1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i-1]$ に近く、かつ、 $[i]$ 、 $[i-1]$ よりも $\sim [i]$ 、 $\sim [i-1]$ の方が値が大きく、かつ、 i の値が小さいほど $[i]$ 、 $\sim [i]$ の絶対値が小さくなるように求める。

10

【0035】

この発明の第五の態様の周波数領域パラメータ列生成方法は、 p を1以上の整数とし、 $[1], [2], \dots, [p]$ を所定の時間区間の音信号に対応する隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であるときに見做したときの時間領域の信号に対応する線形予測係数列に由来する周波数領域パラメータ列とし、周波数領域パラメータ列 $[1], [2], \dots, [p]$ を入力として、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を求めるパラメータ列変換ステップを含む。パラメータ列変換ステップは、変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ における各 $\sim [i]$ ($i=1, 2, \dots, p$)を、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i+1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i+1]$ に近く、かつ、 $[i+1]$ 、 $[i]$ よりも $\sim [i+1]$ 、 $\sim [i]$ の方が値が大きく、かつ、 i の値が大きいほど $\sim [i]$ 、 $\sim [i]$ の絶対値が小さくなるように求め、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i-1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i-1]$ に近く、かつ、 $[i]$ 、 $[i-1]$ よりも $\sim [i]$ 、 $\sim [i-1]$ の方が値が大きく、かつ、 i の値が大きいほど $[i]$ 、 $\sim [i]$ の絶対値が小さくなるように求める。

20

30

【0036】

この発明の第六の態様の周波数領域パラメータ列生成方法は、 p を1以上の整数とし、 $a[1], a[2], \dots, a[p]$ を所定の時間区間の音信号に対応する隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であるときに見做したときの時間領域の信号に対応する線形予測係数列とし、 $[1], [2], \dots, [p]$ を、線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来するLSPパラメータ列、線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来するLSFパラメータ列、線形予測係数列 $a[1], a[2], \dots, a[p]$ に由来し、かつ、 $[1], [2], \dots, [p]$ の全てが0から までの間に存在し、かつ、線形予測係数列に含まれる全ての線形予測係数が0である場合には $[1], [2], \dots, [p]$ が0から までの間に均等間隔に存在する周波数領域パラメータ列、のいずれかとし、 1 及び 2 をそれぞれ1以下の正の定数である補正係数とし、 K を予め定められた $p \times p$ の帯行列とし、次式により定義される変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を生成するパラメータ列変換ステップを含む。

40

【数 2】

$$\begin{pmatrix} \tilde{\omega}[1] \\ \tilde{\omega}[2] \\ \vdots \\ \tilde{\omega}[p] \end{pmatrix} = K \begin{pmatrix} \omega[1] - \frac{\pi}{p+1} \\ \omega[2] - \frac{2\pi}{p+1} \\ \vdots \\ \omega[p] - \frac{p\pi}{p+1} \end{pmatrix} (\gamma_2 - \gamma_1) + \begin{pmatrix} \omega[1] \\ \omega[2] \\ \vdots \\ \omega[p] \end{pmatrix}$$

10

【0037】

この発明の第七の態様の周波数領域パラメータ列生成方法は、 p を1以上の整数とし、 $a[1], a[2], \dots, a[p+1]$ を所定の時間区間の音信号に対応する隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であるときに見做したときの時間領域の信号に対応する $(p+1)$ 次の線形予測係数列とし、 $[1], [2], \dots, [p+1]$ を、線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来するISPパラメータ列、線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来するISFパラメータ列、線形予測係数列 $a[1], a[2], \dots, a[p+1]$ に由来し、かつ、 $[1], [2], \dots, [p]$ の全てが0から までの間に存在し、かつ、線形予測係数列に含まれる全ての線形予測係数が0である場合には $[1], [2], \dots, [p]$ が0から までの間に均等間隔に存在する周波数領域パラメータ列、のいずれかとし、
1及び 2をそれぞれ1以下の正の定数である補正係数とし、 K を予め定められた $p \times p$ の帯行列とし、次式により定義される変換後周波数領域パラメータ列 $\sim [1], \sim [2], \dots, \sim [p]$ を生成するパラメータ列変換ステップを含む。

20

【数 3】

$$\begin{pmatrix} \tilde{\omega}[1] \\ \tilde{\omega}[2] \\ \vdots \\ \tilde{\omega}[p] \end{pmatrix} = K \begin{pmatrix} \omega[1] - \frac{\pi}{p+1} \\ \omega[2] - \frac{2\pi}{p+1} \\ \vdots \\ \omega[p] - \frac{p\pi}{p+1} \end{pmatrix} (\gamma_2 - \gamma_1) + \begin{pmatrix} \omega[1] \\ \omega[2] \\ \vdots \\ \omega[p] \end{pmatrix}$$

30

【発明の効果】

【0038】

符号化効率を改善することができる。または、信号処理の精度を改善することができる。

【図面の簡単な説明】

【0039】

【図 1】従来の符号化装置の例を説明するためのブロック図。

40

【図 2】第一実施形態及び第二実施形態の符号化装置の例を説明するためのブロック図。

【図 3】第一実施形態及び第二実施形態の符号化方法の例を説明するためのフローチャート。

【図 4】変換行列 U の例を説明するための図。【図 5】変換行列 U の例を説明するための図。【図 6】変換行列 U の例を説明するための図。【図 7】図 6 の変換行列 U における重心位置の性質を説明するための図。【図 8】変換行列 V の例を説明するための図。

【図 9】第一実施形態及び第二実施形態の復号装置の例を説明するためのブロック図。

【図 10】第一実施形態及び第二実施形態の復号方法の例を説明するためのフローチャー

50

ト。

【図11】第三実施形態の符号化装置の例を説明するためのブロック図。

【図12】第三実施形態の符号化方法の例を説明するためのフローチャート。

【図13】LSPパラメータの性質を説明するための図。

【図14】LSPパラメータの性質を説明するための図。

【図15】LSPパラメータの性質を説明するための図。

【図16】第三実施形態の復号装置の例を説明するためのブロック図。

【図17】第三実施形態の復号方法の例を説明するためのフローチャート。

【図18】第四実施形態の周波数領域パラメータ列生成装置の例を説明するためのブロック図。

10

【発明を実施するための形態】

【0040】

以下、この発明の実施の形態について詳細に説明する。なお、図面中において同じ機能を有する構成部には同じ番号を付し、重複説明を省略する。

【0041】

[技術的背景]

まず、従来技術で説明した符号化処理を例に技術的背景について説明する。

【0042】

この発明の一例では、パワースペクトル包絡系列を利用する際に、周波数方向の非線形な解像度による離散化によって、離散係数列であるパワースペクトル包絡系列を生成する。このとき、パワースペクトル包絡の振幅のばらつきが大きい周波数領域では細かい解像度で離散化し、パワースペクトル包絡の振幅のばらつきが小さい周波数領域では粗い解像度で離散化する。これにより、正規化MDCT係数列の値のばらつきを小さくし、符号化効率を高めることができる。

20

【0043】

例えば、エネルギーの集中している周波数領域の離散化間隔を他の周波数領域の離散化間隔よりも小さくするようにする。言い換えれば、エネルギーの集中している周波数領域の解像度を他の周波数領域の解像度よりも高くするようにする。

【0044】

従来は、周波数方向の線形な解像度による離散化によりパワースペクトル包絡を表現していた。すなわち、F=0HzからF=100Hzを全周波数領域とし、N=10として、11個のサンプル点で離散化する場合、以下の表の11個の周波数のそれぞれをサンプル点として、これらの11個のサンプル点にそれぞれ対応するパワースペクトル包絡値の系列によりパワースペクトル包絡を表現していた。

30

【表1】

インデックス	0	1	2	3	4	5	6	7	8	9	10
(i)の値											
周波数(Hz)	0	10	20	30	40	50	60	70	80	90	100

40

【0045】

つまり、隣接するサンプル点間の周波数領域での間隔が均等(上述の例では10Hz間隔)であるようなサンプル点列を用いてパワースペクトル包絡を表現していた。

【0046】

このように、パワースペクトル包絡系列を表現するためのサンプル点列を、隣接するサンプル点間の周波数領域での間隔が均等(上述の例では10Hz間隔)であるように離散化することを、「線形離散化」と呼ぶ。

【0047】

50

これに対して、本発明の一例では、周波数方向の非線形な解像度による離散化によって、パワースペクトル包絡を表現する。例えば、以下の表の11個の周波数のそれぞれをサンプル点として、これらの11個のサンプル点に対応するパワースペクトル包絡値の系列によりパワースペクトル包絡を表現する。

【表2】

インデックス (i)の値	0	1	2	3	4	5	6	7	8	9	10
周波数(Hz)	0	1	3	6	10	15	23	35	50	70	100

10

【0048】

この例では、低周波数領域の方が隣接するサンプル点間の周波数領域での間隔が狭く、高周波数領域ほど隣接する離散化サンプル点間の周波数領域での間隔が広がっている。例えば、最も低周波数領域のサンプル点の間隔、言い換えればインデックス0に対応するサンプル点とインデックス1に対応するサンプル点との周波数領域での間隔は1Hzであるが、最も高周波数領域のサンプル点の間隔、言い換えればインデックス9に対応するサンプル点とインデックス10に対応するサンプル点との周波数領域での間隔は30Hzとなっている。

20

【0049】

このように、パワースペクトル包絡系列を表現するためのサンプル点列を、隣接するサンプル点間の周波数方向での間隔が均等でないように離散化することを、「非線形離散化」と呼ぶ。

【0050】

以下では、このような周波数方向の解像度の違いを区別するため、周波数方向に等間隔なサンプル点の系列を「線形離散化サンプル点列」とも呼び、周波数方向に不均等な間隔のサンプル点の系列を「非線形離散化サンプル点列」とも呼ぶこととする。線形離散化サンプル点列の隣接するサンプル点の周波数の間隔は均等であるが、非線形離散化サンプル点列の隣接するサンプル点の周波数の間隔は不均等である。また、線形離散化サンプル点列に含まれる各サンプル点を「線形離散化サンプル点」とも呼び、非線形離散化サンプル点列に含まれる各サンプル点を「非線形離散化サンプル点」とも呼ぶ。

30

【0051】

また、線形離散化サンプル点列の各サンプル点に対応する入力された音信号のパワーの系列を「パワースペクトル系列」とも呼び、非線形離散化サンプル点列の各サンプル点に対応する入力された音信号のパワーの系列を「伸縮疑似パワースペクトル系列」とも呼ぶこととする。

【0052】

なお、上述の例では、非線形離散化サンプル点列は低周波数領域ほどサンプル点間の周波数の間隔が狭くなっているが、必ずしもこの性質である必要はなく、例えば中周波数領域の方が低周波数領域よりもサンプル点間の周波数の間隔が狭くてもよい。要するに、非線形離散化サンプル点列では、隣接するサンプル点間の周波数の間隔が不均等であればよい。

40

【0053】

[第一実施形態]

(第一実施形態の符号化)

第一実施形態の符号化装置の構成例を図2に示す。第一実施形態の符号化装置は、図2に示すように、周波数領域変換部21と、伸縮疑似パワースペクトル系列生成部22と、線形予測分析部23と、伸縮パワースペクトル包絡系列生成部24と、逆伸縮変換部25と、包絡正規化部26と、符号化部27とを例えば備えている。この符号化装置により実

50

現される第一実施形態の符号化方法の各処理の例を図3に示す。

【0054】

以下、図2の各部について説明する。

【0055】

<周波数領域変換部21>

周波数領域変換部21には、時間領域の音信号が入力される。音信号の例は、音声デジタル信号又は音響デジタル信号である。

【0056】

周波数領域変換部21は、所定の時間長のフレーム単位で、入力された時間領域の音信号を周波数領域のN点のMDCT係数列 $X(0), X(1), \dots, X(N-1)$ に変換する(ステップE1)。Nは正の整数である。

10

【0057】

変換されたMDCT係数列 $X(0), X(1), \dots, X(N-1)$ は、包絡正規化部26に出力される。

【0058】

特に断りがない限り、以降の処理はフレーム単位で行われるものとする。

【0059】

ここでのMDCT係数列 $X(0), X(1), \dots, X(N-1)$ に対応する各サンプル点は、線形離散化サンプル点である。すなわち、MDCT係数列 $X(0), X(1), \dots, X(N-1)$ に対応するサンプル点列の隣接するサンプル点の周波数の間隔は等間隔である。言い換えれば、 $i=0, 1, \dots, N-2$ として、MDCT係数列におけるインデックス*i*に対応する周波数とMDCT係数列におけるインデックス*i*

20

【0060】

<伸縮疑似パワースペクトル系列生成部22>

伸縮疑似パワースペクトル系列生成部22には、周波数領域変換部21が変換したMDCT係数列 $X(0), X(1), \dots, X(N-1)$ が入力される。

【0061】

伸縮疑似パワースペクトル系列生成部22は、まず、MDCT係数列 $X(0), X(1), \dots, X(N-1)$ の各係数の二乗値(パワー)からなる系列であるパワースペクトル系列 $Y(0), Y(1), \dots, Y(N-1)$ を生成する。すなわち、 $Y(i)=X(i)^2$ ($i=0, 1, \dots, N-1$)である。

【0062】

そして、伸縮疑似パワースペクトル系列生成部22は、パワースペクトル系列 $Y(0), Y(1), \dots, Y(N-1)$ を補間や線形変換することにより、伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を生成する(ステップE2)。

30

【0063】

生成された伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ は、線形予測分析部23に出力される。

【0064】

ここで、パワースペクトル系列 $Y(0), Y(1), \dots, Y(N-1)$ に対応するサンプル点列は線形離散化サンプル点列であり、伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ に対応するサンプル点列は非線形離散化サンプル点列である。

40

【0065】

言い換えれば、パワースペクトル系列 $Y(0), Y(1), \dots, Y(N-1)$ における各インデックス0, 1, ..., N-1に対応する周波数の間隔は等間隔である。また、伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ における各インデックス0, 1, ..., N-1に対応する周波数の間隔は不均等な間隔である。

【0066】

補間により伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を生成する場合、伸縮疑似パワースペクトル系列生成部22は例えば以下の処理を行う。パワースペクトル系列 $Y(0), Y(1), \dots, Y(N-1)$ に対応するサンプル点列の隣接するサンプル点の間の周波数を*f*とする。そして、*f*におけるパワースペクトル値がsinc関数($\text{sinc}(f)=\sin(f)/f$)に従うと

50

仮定して補間した曲線を求める。そして、その曲線における非線形離散化サンプル点列の各サンプル点（周波数）に対応する値を伸縮疑似パワースペクトル値とすることで、伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を生成する。なお、この場合、非線形離散化サンプル点列は予め与えられているものとする。

【 0 0 6 7 】

線形変換により伸縮疑似パワースペクトル系列 $\sim Y(0), \dots, \sim Y(N-1)$ を得る場合、伸縮疑似パワースペクトル系列生成部 22 は、パワースペクトル系列 $Y(0), Y(1), \dots, Y(N-1)$ からなるベクトルに予め定められた変換行列 U を左から乗じることで伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を生成する。

【 0 0 6 8 】

言い換えれば、伸縮疑似パワースペクトル系列生成部 22 は、以下の式により定義される伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を生成する。

【数 4】

$$\begin{pmatrix} \tilde{Y}(0) \\ \tilde{Y}(1) \\ \vdots \\ \tilde{Y}(N-1) \end{pmatrix} = U \begin{pmatrix} Y(0) \\ Y(1) \\ \vdots \\ Y(N-1) \end{pmatrix}$$

【 0 0 6 9 】

ここで、変換行列 U は、線形離散化サンプル点列から非線形離散化サンプル点列へのマッピングを近似する行列である。この変換は、周波数方向で等間隔のサンプル点列を周波数方向で不均等な間隔のサンプル点列に変換するものであり、いわば隣接するサンプル点間の周波数の間隔を伸縮させるものであることから、「伸縮変換」と呼ぶこととする。

【 0 0 7 0 】

伸縮変換後の非線形離散化サンプル点列のある非線形離散化サンプル点の周波数（以下、伸縮後周波数と呼ぶ。）は、伸縮変換前の線形離散化サンプル点列の線形離散化サンプル点のうち伸縮後周波数と近い周波数を持つ 1 以上の線形離散化サンプル点の周波数の重み付き和で近似できる。言い換えれば、伸縮後周波数は、当該伸縮後周波数と最も近い周波数の線形離散化サンプル点の近傍の 1 以上の線形離散化サンプル点の周波数の重み付き和で近似できる。

【 0 0 7 1 】

変換行列 U の各行は伸縮変換後の非線形離散化サンプル点列の各非線形離散化サンプル点に対応し、変換行列 U の各列は線形離散化サンプル点列の各線形離散化サンプル点に対応する。すなわち、変換行列 U の各行は、当該各行に対応する非線形離散化サンプル点の周波数を表現するための各線形離散化サンプル点の周波数についての重みの系列になっている。

【 0 0 7 2 】

本来は、伸縮後周波数を当該伸縮後周波数と最も近い周波数の線形離散化サンプル点の近傍の 1 以上の線形離散化サンプル点の周波数の重み付き和で近似するならば、その重みを負の値としてもよい。しかし、負の値を含むように変換行列 U を構成すると、精度よく伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を生成するためには後処理が必要となってしまう。そこで、この発明では、例えば変換行列 U の全ての要素を非負の値とする（つまり、 U を非負値行列とする）。これにより、後処理をすることなく精度の高い伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を生成することができる。

【 0 0 7 3 】

さらに、伸縮後周波数と離れた周波数の離散化サンプル点の周波数に乗じる重みは小さい値となることが想定されるため、小さい値の要素を 0 と見做しても、伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ の精度への影響が少ない。そこで、変換行列 U の各

10

20

30

40

50

行は、当該各行に対応する非線形離散化サンプル点の周波数と最も近い周波数の線形離散化サンプル点に対応する列の近傍要素のみ0でない値とし、残りの要素は0と例えば設定してもよい。ここで、変換行列Uにおいて0以外の値をもつ要素を「伸縮で対応する要素」と呼ぶ。変換行列Uは、例えば、伸縮で対応する要素の近傍のみ0でない値を持ち、それ以外の成分は0であるような帯状の行列（疎行列）であるといえる。

【0074】

行列のすべての値を使って変換することは演算量が多くなる可能性があるが、このように変換行列Uを疎行列とすることで少ない演算量で伸縮疑似パワースペクトルを得ることができる。行列中の0でない要素の開始サンプル点を別途記憶して、そのサンプル点からのみの少数の演算とすればよい。

10

【0075】

このように、変換行列Uを非負値行列、もしくは、疎行列とすることで、少ない演算量で精度よく伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を求めることができる。なお、変換行列Uを疎行列とする場合には負の値の要素を含んでもよい。

【0076】

変換行列Uは、非線形離散化サンプル点列と線形離散化サンプル点列との相関関係から予め学習などにより求めておくことができる。変換行列Uの求め方については後述する。

【0077】

非線形離散化サンプル点列の隣接するサンプル点の間隔の伸縮度合いは任意に設定できる。低周波数領域のサンプル点間隔を縮小し、高周波数領域のサンプル点間隔を拡大する例として下記的一般化対数を使うことができる。なお、「隣接するサンプル点の間隔」のこと「離散化幅」とも呼ぶ

20

【0078】

線形離散化サンプル点列のインデックスiのサンプル点に対応する周波数 ω_i との関係をあらず関数の例として $S(\omega_i)$ を使うことができる。非線形離散化サンプル点列の同じインデックスiのサンプル点に対応する周波数 $S(\omega_i)$ と周波数 ω_i の間には、例えば以下の関係が成り立つ。

【数5】

$$S_\lambda(\omega_i) = \begin{cases} (\omega_i^\lambda - 1) / \lambda & 0 < |\lambda| \leq 1 \\ \log \omega_i & \lambda = 0 \end{cases}$$

30

【0079】

$S(\omega_i)$ は、一般化対数関数と呼ばれる関数である。 λ は、非線形な伸縮度合い、すなわち非線形離散化サンプル点列のサンプル点の間隔の伸縮度合いを決める定数である。伸縮変換の度合いである λ は、入力信号の性質に合わせて設計することができる。 $\lambda=0$ のときは、 $S(\omega_i)$ は対数関数となる。

【0080】

変換行列Uと $S(\omega_i)$ との関係の例を図4に示す。図4(a)、図4(b)及び図4(c)は、縦軸を変換行列Uの行に対応するインデックス、横軸を変換行列Uの列に対応するインデックスとして、変換行列Uの各要素の値を色で示したものである。白色は値が0の要素を表し、黒色は0より大きな値を持つ要素を表す。 $\lambda=1$ の場合の変換行列Uは、線形離散化サンプル点列のまま何も変換しない行列であり、対角成分のみ値が1の対角行列である。これらの図から、 λ が小さくなるほど、変換行列Uにおける非零成分がなす曲線の形状、言い換えれば0より大きな値を持つ要素を近似する曲線の形状が対角線よりも下方向に非線形に曲がっていることが分かる。

40

【0081】

ここで、縦軸を線形離散化サンプル点列のサンプル点の周波数とし、縦軸の上から下に向かうほど周波数が増大するものとし、横軸を非線形離散化サンプル点列の周波数とし、横軸の左から右に向かうほど周波数が増大するものとして定義される2次元平面を考える

50

【0082】

変換行列Uにおける非零成分がなす曲線の形状は、この2次元平面上に、変換行列Uの各要素のインデックスに対応する線形離散化サンプル点の周波数及び非線形離散化サンプル点の周波数をマッピングしたときの点列を補間して得られる曲線に相当するとも言える。この曲線のことを以下「伸縮曲線」とも呼ぶ。伸縮曲線は、言い換えれば、「伸縮に対応する要素」をマッピングした点列を補間して得られる曲線である。ここで、要素のインデックスに対応する線形離散化サンプル点とは、その要素の列に対応する線形離散化サンプル点のことである。また、要素のインデックスに対応する非線形離散化サンプル点とは、その要素の行に対応する非線形離散化サンプル点のことである。

10

【0083】

なお、上記の例では低周波数領域の解像度が高周波数領域の解像度よりも高くなるような変換の例を示した。言い換えれば、低周波数領域の非線形離散化サンプル点間の間隔が高周波数領域の非線形離散化サンプル点の間隔よりも狭くなるような変換の例を示した。しかし、これはあくまで一例に過ぎない。

【0084】

解像度又は伸縮の度合いは変換行列の非零要素が存在するサンプル点の傾きに対応するので、例えば、非線形離散化サンプル点列として、低周波数領域と高周波数領域との中間の周波数領域の解像度が他の周波数領域の解像度よりも高くなるような領域と低くなる領域が中心近くにあるような非線形離散化サンプル点列を用いても良い。この場合、変換行列Uは、例えば図5のような伸縮曲線上の近傍成分のみ非零となるような変換行列となる。

20

【0085】

いずれにしても、変換行列UをM×N行列として、変換行列Uにおける伸縮曲線は、行を横軸とし、列を縦軸と見做した二次元平面において、左上端の成分に対応する点[1,1]から右下端の成分に対応する点[M,N]に向けて単調減少する曲線となる。言い換えると、変換行列Uは以下のような性質を持つ。

【0086】

変換行列Uの第i行の重心 g_i を、

【数6】

$$g_i = \left(\sum_{k=1}^N k \times U[i, k] \right) / \left(\sum_{k=1}^N U[i, k] \right)$$

と定義する。U[i, k]は変換行列Uの(i, k)要素を表す。行と列のインデックスは1から開始されるものとする。すると、変換行列Uの各行の重心の系列 g_1, g_2, \dots, g_M は $g_1 < g_2 < \dots < g_M$ という関係を満たす。なお、MとNはそれぞれ3以上の整数である。また、本実施形態においてはM=Nである。

30

【0087】

また、音声信号及び音響信号では低周波数領域又は中周波数領域にエネルギーが集中することが多いので、低周波数領域又は中周波数領域の非線形離散化サンプル点間の周波数の間隔を、高周波数領域の非線形離散化サンプル点間の周波数の間隔よりも狭くした非線形離散化サンプル点列に対応する伸縮疑似パワースペクトルを用いてもよい。この場合、変換行列Uの各行において、当該各行の非零要素のうち対角要素から最も距離の離れている要素と対角要素との距離（この距離は、列のインデックスがどれだけ離れているかを表す。）を各行のインデックスの伸縮距離としたとき、全インデックスのうちの前半分のインデックスに対する伸縮距離の平均は、後半分のインデックスに対する伸縮距離の平均よりも大きい傾向にある。また、非零成分は正の値をとる。あるいは、変換行列Uは、その上三角行列に含まれる非負成分の数が、その下三角行列に含まれる非負成分の数より少ないような疎行列となる。図6に変換行列Uの一例を示す。この例ではN=M=16である。図7には、図6の変換行列Uの重心 g_i を、行番号を横軸とし、 g_i の値を縦軸とする二次元平面にプロットしたものを示す。図7に示されているように、行番号が増加するにつれて

40

50

g_i の値は単調増加する。

【 0 0 8 8 】

伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ は、入力された音信号を周波数領域に変換した信号のパワースペクトルを非線形離散化したものに相当する。

【 0 0 8 9 】

このように、入力された音信号を周波数領域に変換した信号のエネルギーが集中している周波数領域での離散化間隔がそれ以外の周波数領域での離散化間隔よりも狭くなるような非線形な離散化間隔のサンプル点に基づいてパワースペクトルを表現する。

【 0 0 9 0 】

< 線形予測分析部 2 3 >

線形予測分析部 2 3 には、伸縮疑似パワースペクトル系列生成部 2 2 が生成した伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ が入力される。

【 0 0 9 1 】

線形予測分析部 2 3 は、伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を用いて、以下の式により定義される $\sim X(0), \sim X(1), \dots, \sim X(N-1)$ を線形予測分析して伸縮線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ を生成し、生成された伸縮線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ を符号化して伸縮線形予測係数符号と伸縮線形予測係数符号に対応する量子化された伸縮線形予測係数である量子化伸縮線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ とを生成する(ステップ E 3)。

【 数 7 】

$$\tilde{X}(k) = \sum_{n=0}^{N-1} \tilde{Y}(n) \exp\left(-j \frac{2\pi kn}{N}\right), \quad k = 0, 1, \dots, N-1$$

【 0 0 9 2 】

生成された量子化伸縮線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ は、伸縮パワースペクトル包絡系列生成部 2 4 に出力される。

【 0 0 9 3 】

また、生成された伸縮線形予測係数符号は、復号装置に送信される。

【 0 0 9 4 】

伸縮線形予測係数符号を生成するために、線形予測分析部 2 3 は、まず伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ に対応するサンプル点の周波数の間隔が均等であると見做して逆FFTに相当する演算を行うことにより、 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ に対応する時間領域の信号列である伸縮相関関数信号列 $\sim X(0), \sim X(1), \dots, \sim X(N-1)$ を求める。そして、線形予測分析部 2 3 は、求めた伸縮相関関数信号列 $\sim X(0), \sim X(1), \dots, \sim X(N-1)$ に対して線形予測分析を行って、伸縮線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ を生成する。そして、線形予測分析部 2 3 は、生成された伸縮線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ を符号化することにより、伸縮線形予測係数符号を生成する。この結果、伸縮線形予測係数符号に対応する量子化伸縮線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ も得られる。

【 0 0 9 5 】

伸縮線形予測係数は、伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ のサンプル点の周波数方向での間隔が均等な間隔であると見做したときの時間領域の信号に対応する線形予測係数である。

【 0 0 9 6 】

線形予測分析部 2 3 による伸縮線形予測係数符号の生成は、例えば従来の符号化技術によって行われる。従来の符号化技術とは、例えば、線形予測係数そのものに対応する符号を予測係数符号とする符号化技術、線形予測係数を L S P パラメータに変換して L S P パラメータに対応する符号を予測係数符号とする符号化技術、線形予測係数を P A R C O R 係数に変換して P A R C O R 係数に対応する符号を予測係数符号とする符号化技術などである。これらの従来の符号化技術において、線形予測係数を伸縮線形予測係数に置き換えて符号化することで、伸縮線形予測係数に対応する符号である伸縮線形予測係数符号が得られる。

10

20

30

40

50

【 0 0 9 7 】

< 伸縮パワースペクトル包絡系列生成部 2 4 >

伸縮パワースペクトル包絡系列生成部 2 4 には、線形予測分析部 2 3 が生成した量子化伸縮線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ が入力される。

【 0 0 9 8 】

伸縮パワースペクトル包絡系列生成部 2 4 は、量子化伸縮線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ を周波数領域に変換することにより、量子化伸縮線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ に対応する周波数領域のパワースペクトル包絡系列である伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ を生成する（ステップ E 4）。

【 0 0 9 9 】

生成された伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ は、逆伸縮変換部 2 5 に出力される。

【 0 1 0 0 】

伸縮パワースペクトル包絡系列生成部 2 4 は、量子化伸縮線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ を用いて、伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ として、例えば以下の式（2）により定義される伸縮非平滑化パワースペクトル包絡系列 $\sim W_o(0), \sim W_o(1), \dots, \sim W_o(N-1)$ 又は以下の式（3）により定義される伸縮平滑化パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ を生成する。

【 0 1 0 1 】

言い換えれば、伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ の一例が、式（2）により定義される伸縮非平滑化パワースペクトル包絡系列 $\sim W_o(0), \sim W_o(1), \dots, \sim W_o(N-1)$ 又は式（3）により定義される伸縮平滑化パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ である。

【 数 8 】

$$\tilde{W}_o(k) = \frac{\sigma^2}{2\pi} \frac{1}{\left| 1 + \sum_{n=1}^p \hat{\beta}_n \exp(-j2\pi kn/N) \right|^2} \quad \dots(2)$$

$$\tilde{W}_\gamma(k) = \frac{\sigma^2}{2\pi} \frac{1}{\left| 1 + \sum_{n=1}^p \hat{\beta}_n \gamma^n \exp(-j2\pi kn/N) \right|^2}, \quad k = 0, 1, \dots, N-1 \quad \dots(3)$$

【 0 1 0 2 】

ここで、補正係数 γ は予め定められた 1 以下の定数であり、伸縮非平滑化パワースペクトル包絡系列 $\sim W_o(0), \sim W_o(1), \dots, \sim W_o(N-1)$ の振幅の凹凸を鈍らせる係数、言い換えれば伸縮非平滑化パワースペクトル包絡系列 $\sim W_o(0), \sim W_o(1), \dots, \sim W_o(N-1)$ を平滑化する係数である。伸縮変換を行わない従来の符号化処理におけるパワースペクトル包絡系列の平滑化に用いる γ 、すなわち上述の式（P 1）における γ と同じと考えればよい。

【 0 1 0 3 】

なお、伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ に対応するサンプル点列、言い換えれば伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ における各インデックス $0, 1, \dots, N-1$ に対応する周波数の系列は、非線形離散化サンプル点列である。

【 0 1 0 4 】

このようにして、伸縮パワースペクトル包絡系列生成部 2 4 は、所定の時間区間ごとの音信号に由来する周波数領域のサンプル列のパワースペクトル包絡を平滑化した包絡を周波数方向に不均等間隔で離散化した系列である伸縮パワースペクトル包絡系列を生成する。

【 0 1 0 5 】

10

20

30

40

50

所定の時間区間ごとの音信号に由来する周波数領域のサンプル列とは、この例ではMDCT係数列 $X(0), X(1), \dots, X(N-1)$ のことである。所定の時間区間ごとの音信号に由来する周波数領域のサンプル列として、MDCT係数列 $X(0), X(1), \dots, X(N-1)$ 以外の周波数領域のサンプル列を用いてもよい。

【0106】

伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ は、サンプル点の周波数の間隔が狭いところでは細かい解像度で表現されるため、パワースペクトル包絡の振幅の凹凸の細かい変化も表現することができる。逆に、サンプル点の周波数の間隔が広いところでは粗い解像度で表現されるため、パワースペクトル包絡の大まかな変化のみしか表現されない。一般に、音声音響信号はエネルギーの集中している部分でのパワースペクトル包絡の変化が大きく、それ以外の部分でのパワースペクトル包絡の変化は小さい。よって、エネルギーの集中している周波数領域での隣接するサンプル点間の周波数の間隔がそれ以外の周波数領域での隣接するサンプル点間の周波数の間隔よりも狭いような非線形離散化サンプル点列に対応する周波数領域の入力された音信号のパワーの系列を伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ とすることで、限られたサンプル点列でより正確にパワースペクトルの振幅の凹凸の変化を表現する離散系列を得ることができる。言い換えれば、エネルギーの集中している周波数領域を他の周波数領域よりも細かい解像度で表現するように非線形離散化することで、より正確にパワースペクトルを表現する離散系列を得ることができる。

10

【0107】

こうして得られた伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ に基づいて算出したスペクトル包絡値を用いてMDCT係数列を正規化すると、正規化MDCT係数列の大きさの変化が小さくなるため、効率的に符号化できるようになる。

20

【0108】

< 逆伸縮変換部 25 >

逆伸縮変換部 25 には、伸縮パワースペクトル包絡系列生成部 24 が生成した伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ が入力される。

【0109】

逆伸縮変換部 25 は、伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ を、補間又は線形変換により線形離散化サンプル点列に対応するパワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ に変換する(ステップ E5)。

30

【0110】

変換されたパワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ は、包絡正規化部 26 に出力される。

【0111】

補間によりパワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ を得る場合、逆伸縮変換部 25 は例えば以下の処理を行う。逆伸縮変換部 25 は、伸縮疑似パワースペクトル系列生成部 22 と同様に、伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ を sinc 関数により補間した曲線(伸縮パワースペクトル包絡をなめらかにつないだ包絡)を求める。そして、その曲線上で線形離散化サンプル点列の各離散化サンプル点に対応する周波数のパワースペクトル包絡値の系列をパワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ として得る。

40

【0112】

線形変換によりパワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ を得る場合、逆伸縮変換部 25 は、伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ からなるベクトルに予め定められた変換行列 V を左から乗じることでパワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ を生成する。

【0113】

言い換えれば、逆伸縮変換部 25 は、以下の式により定義されるパワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ を生成する。

50

$$\begin{bmatrix} W(0) \\ W(1) \\ \vdots \\ W(N-1) \end{bmatrix} = V \begin{bmatrix} \tilde{W}(0) \\ \tilde{W}(1) \\ \vdots \\ \tilde{W}(N-1) \end{bmatrix}$$

【 0 1 1 4 】

ここで、変換行列Vは、変換行列Uの逆変換を近似する行列であり、非線形離散化サンプル点列から線形離散化サンプル点列へのマッピングを近似する行列である。この変換は、不均等な間隔のサンプル点列を等間隔のサンプル点列に変換するものであり、上述の「伸縮変換」とは逆の関係となるようにサンプル点の間隔を伸縮させるものであることから、「逆伸縮変換」と呼ぶこととする。ただし、変換行列Vが変換行列Uの逆行列を意味するものではない。

10

【 0 1 1 5 】

変換行列Vは、例えば、逆伸縮で対応する要素の近傍のみ0でない値を持ち、それ以外の成分は0であるような帯状の行列（疎行列）とする。

【 0 1 1 6 】

逆伸縮変換後の線形離散化サンプル点列のある線形離散化サンプル点（以下、「逆伸縮後周波数」と呼ぶ。）は、逆伸縮変換前の非線形離散化サンプル点列の非線形離散化サンプル点のうち逆伸縮後周波数と近い周波数を持つ1以上の非線形離散化サンプル点の周波数の重み付き和で近似できる。言い換えれば、逆伸縮後周波数は、当該逆伸縮後周波数と最も近い周波数の非線形離散化サンプル点の近傍の1以上の非線形離散化サンプル点の周波数の重み付き和で近似できる。

20

【 0 1 1 7 】

変換行列Vの各行は逆伸縮変換後の線形離散化サンプル点列の各線形離散化サンプル点に対応し、変換行列Vの各列は非線形離散化サンプル点列の非線形離散化サンプル点に対応する。すなわち、変換行列Vの各行は、当該行に対応する線形離散化サンプル点の周波数を表現するための、各非線形離散化サンプル点の周波数についての重みの系列になっている。

30

【 0 1 1 8 】

上記の性質から、変換行列Vの各行は、当該各行に対応する線形離散化サンプル点の周波数と最も近い周波数の非線形離散化サンプル点に対応する列の近傍要素のみ0でない値とし、残りの要素は0と例えば設定される。これは、線形離散化サンプル点の周波数とは離れた周波数の非線形離散化サンプル点の影響は極めて小さいため0とみなすことができるからである。

【 0 1 1 9 】

上述の「逆伸縮で対応する要素」とは、逆伸縮後周波数に対応する変換行列Vの行の要素の中の、当該逆伸縮後周波数を近似するために用いる非線形離散化サンプル点に対応する列の要素を指す。すなわち、上述の「逆伸縮で対応する要素」とは、変換行列Vの各行の要素の中の、当該各行に対応する逆伸縮後周波数に最も近い周波数の非線形離散化サンプル点に対応する列の近傍の要素である。

40

【 0 1 2 0 】

変換行列VをN×M行列として、変換行列Vにおける伸縮曲線は、行を横軸とし、列を縦軸と見做した二次元平面において、左上端の成分に対応する点[1,1]から右下端の成分に対応する点[N,M]に向けて単調減少する曲線となる。言い換えると、変換行列Vは以下のような性質を持つ。

【 0 1 2 1 】

変換行列Vの第i行の重心 g_i ' を、

50

【数 1 0】

$$g'_i = \left(\sum_{k=1}^M k \times V[i, k] \right) / \left(\sum_{k=1}^M V[i, k] \right)$$

と定義する。V[i, k]は変換行列Vの(i, k)要素を表す。行と列のインデックスは1から開始されるものとする。すると、変換行列Vの各行の重心の系列 g'_1, g'_2, \dots, g'_N は $g'_1 < g'_2 < \dots < g'_N$ という関係を満たす。

【0 1 2 2】

ただし、変換行列Vにおける伸縮曲線は、変換行列Uの伸縮曲線とは逆の曲がり方の伸縮曲線（逆伸縮曲線）に沿った成分のみ非零の値を持つような疎行列となる。つまり、変換行列Uの伸縮曲線と変換行列Vの伸縮曲線は、行を横軸とし、列を縦軸と見做した二次元平面において、左上端の成分に対応する点[1, 1]と右下端の成分に対応する点[N, M]とを結ぶ直線に対してほぼ線対称な形状となる。

10

【0 1 2 3】

例えば、変換行列Uが、その上三角行列に含まれる非零成分の数が、その下三角行列に含まれる非零成分の数よりも少ないような疎行列である場合は、変換行列Vは、その上三角行列に含まれる非零成分の数が、その下三角行列に含まれる非零成分の数よりも多いような疎行列となる。このとき、変換行列Vの各行において、当該各行の非零要素のうち対角要素から最も距離の離れている要素と対角要素との距離（この距離は、列のインデックスがどれだけ離れているかを表す。）を各行のインデックスの伸縮距離としたとき、全インデックスのうちの前半分のインデックスに対する伸縮距離の平均は、後ろ半分のインデックスに対する伸縮距離の平均よりも小さい傾向にある。図8に変換行列Vの例を示す。この例ではN=M=16である。

20

【0 1 2 4】

なお、変換行列Vは、変換行列Uと同様に非負値行列であってもよい。この場合も変換行列Vは各行の重心が上記の性質 $g'_1 < g'_2 < \dots < g'_N$ を満たす。

【0 1 2 5】

< 包絡正規化部 2 6 >

包絡正規化部 2 6 には、周波数領域変換部 2 1 が変換したMDCT係数列 $X(0), X(1), \dots, X(N-1)$ 及び逆伸縮変換部 2 5 が変換したパワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ が入力される。

30

【0 1 2 6】

包絡正規化部 2 6 は、パワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ を用いて、周波数領域のサンプル列であるMDCT係数列 $X(0), X(1), \dots, X(N-1)$ を正規化することにより、正規化された周波数領域サンプル列を生成する（ステップ E 6）。正規化された周波数領域サンプル列とは、この例では正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ である。

【0 1 2 7】

生成された正規化された周波数領域サンプル列は、符号化部 2 7 に出力される。

【0 1 2 8】

包絡正規化部 2 6 は、例えば、 $i=0, 1, \dots, N-1$ として、MDCT係数列 $X(0), X(1), \dots, X(N-1)$ の各係数 $X(i)$ をパワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ の各包絡値 $W(i)$ の平方根で除算することにより、正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ の各係数 $X_N(i)$ を生成する。すなわち、 $i=0, 1, \dots, N-1$ として、 $X_N(i) = X(i) / \sqrt{W(i)}$ である。ここで、 x を実数として \sqrt{x} は x の平方根を表す。

40

【0 1 2 9】

なお、パワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ は、伸縮パワースペクトル包絡系列 $-W(0), -W(1), \dots, -W(N-1)$ に由来するものである。

【0 1 3 0】

したがって、包絡正規化部 2 6 は、伸縮パワースペクトル包絡系列 $-W(0), -W(1), \dots, -W(N-1)$ に基づいて周波数領域のサンプル列であるMDCT係数列 $X(0), X(1), \dots, X(N-1)$ を正規化することにより、正規化された周波数領域サンプル列を生成しているとも言える。

50

【 0 1 3 1 】

< 符号化部 2 7 >

符号化部 2 7 には、包絡正規化部 2 6 が生成した正規化された周波数領域サンプル列が
入力される。この例では、正規化された周波数領域サンプル列は、正規化MDCT係数列 $X_N(0)$
 $X_N(1), \dots, X_N(N-1)$ である。

【 0 1 3 2 】

符号化部 2 7 は、正規化された周波数領域サンプル列を符号化して、その正規化された
周波数領域サンプル列に対応する符号を生成する（ステップ E 7）。

【 0 1 3 3 】

生成された符号は、復号装置に出力される。

10

【 0 1 3 4 】

符号化部 2 7 は、例えば従来と同様に正規化された周波数領域サンプル列に対応する符
号を生成する。

【 0 1 3 5 】

すなわち、正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ の各係数を利用（グローバルゲ
イン） g で割り算し、その結果を量子化した整数値による系列である量子化正規化済係数系
列 $X_Q(0), X_Q(1), \dots, X_Q(N-1)$ を符号化して得られる符号を整数信号符号とする。符号化部 2
7 は、この整数信号符号のビット数が、予め配分されたビット数である配分ビット数 B 以
下、かつ、なるべく大きな値となるような利得 g を決定する。そして、符号化部 2 7 は、
この決定された利得 g に対応する利得符号と、この利得符号に対応する利得 g で正規化MDCT
係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ の各係数を割り算した結果を量子化した整数値による系列
である量子化正規化済係数系列 $X_Q(0), X_Q(1), \dots, X_Q(N-1)$ を符号化して得られる整数信号符
号とを生成する。この場合、利得符号と整数信号符号とが、正規化された周波数領域サン
プル列に対応する符号となる。

20

【 0 1 3 6 】

〔変換行列 U と変換行列 V の求め方の例〕

線形離散化サンプル点列から非線形離散化サンプル点列への線形変換を実現する変換行
列 U と、非線形離散化サンプル点列から線形離散化サンプル点列への線形変換を実現する
変換行列 V とは、例えば予め学習によって求めておくことができる。ここで U, V の要素はす
べて負でないという制約をつける。変換前後のベクトルはパワースペクトルまたはその包
絡であるため、すべて正值であるためである。

30

【 0 1 3 7 】

線形変換を用いてパワースペクトル系列 $Y(0), Y(1), \dots, Y(N-1)$ から伸縮疑似パワースペ
クトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を得る方法は、sinc関数などにより補間することで伸
縮疑似パワースペクトルを求める場合と比較して意図しない変換を防ぐことができるとい
う利点がある。

【 0 1 3 8 】

まず、学習データとして、 T 個の線形離散化サンプル点から成る線形離散化サンプル点
列に基づいて表現されるパワースペクトル系列 $(Y_t(0), \dots, Y_t(N-1)) (t=1, 2, \dots, T)$ の集合 Y
と、これをsinc関数などにより補間することで得られる非線形離散化サンプル点列に基づ
いて表現される伸縮疑似パワースペクトル系列 $(\sim Y_t(0), \dots, \sim Y_t(N-1)) (t=1, 2, \dots, T)$ の集合
 $\sim Y$ を用意しておく。

40

【 0 1 3 9 】

そして、予め適当な初期値を設定した帯状の行列 U を用いて、 UY と $\sim Y$ との距離が最小と
なるように、 U の各成分を更新していくことで、変換行列 U を学習する。

【 0 1 4 0 】

距離としては、線形予測係数が包絡の二乗値と元のパワースペクトルとの板倉齋藤距離
が最小となるものであることから、学習時の距離の尺度としても板倉齋藤距離を用いると
よい。すなわち、伸縮変換行列の学習では、以下の式により定義される関数を目的関数と
して学習を行う。 $D_{1s}(a|b)$ は a を基準とする b との板倉齋藤距離を表す。

50

【数 1 1】

$$\min_U D_{IS}(\tilde{Y} | UY)$$

【0 1 4 1】

変換行列Vは、YとVUYとの板倉齋藤距離が最小となるように、すなわち、以下の式により定義される関数を目的関数として学習を行えばよい。

【数 1 2】

$$\min_V D_{IS}(Y | VUY)$$

【0 1 4 2】

この最適化問題は、例えば補助関数法により解くことができる。例えば、所定の条件を満たすまで、以下の更新式によりUとVの要素を更新することで最適解に近づけることができる。

【数 1 3】

$$U_{i,j}^{(p+1)} = U_{i,j}^{(p)} \sqrt{\frac{\sum_k \tilde{Y}_{i,k} Y_{j,k} / Z_{i,k}^2}{\sum_l Y_{j,l} / Z_{i,l}}}, \quad Z = UX$$

$$V_{i,j}^{(p+1)} = V_{i,j}^{(p)} \sqrt{\frac{\sum_k Y_{i,k} \tilde{Y}_{j,k} / S_{i,k}^2}{\sum_l \tilde{Y}_{j,l} / S_{i,l}}}, \quad S = VUX$$

【0 1 4 3】

ここで、 $U_{i,j}^{(p)}$ と $V_{i,j}^{(p)}$ は、それぞれp回目の繰り返しにより得られた変換行列Vの(i, j)成分、変換行列Vの(i, j)成分を表す。また、 $Y_{j,k} = Y_k(j)$ 、すなわち、学習データYのk番目の疑似パワースペクトル系列におけるj番目の要素を表す。同様に $\tilde{Y}_{i,k} = \tilde{Y}_k(i)$ であり、学習データ \tilde{Y} のk番目の疑似パワースペクトル系列におけるi番目の要素を表す。

【0 1 4 4】

この更新式から、変換行列Uや変換行列Vのうち初期値が0の要素は、学習後も0のままであるため、計算をする必要がなく、この制約の中での最適化が可能である。また、UとVは帯状の行列であるので、あらかじめ非ゼロのサンプルの位置を指定した変換により、実際の変換のための演算量を大幅に削減できるとともに、学習においても0成分である多くの要素は学習する必要がないため、低コストで学習を行うことができる。また、初期値を設定する際の帯幅を調整することで、変換や学習の演算量をさらに調節することができる。

【0 1 4 5】

(第一実施形態の復号)

第一実施形態の符号化装置に対応する復号装置の構成例を図9に示す。第一実施形態の復号装置は、図9に示すように、伸縮線形予測係数復号部31と、伸縮パワースペクトル包絡系列生成部32と、逆伸縮変換部33と、復号部34と、包絡逆正規化部35と、時間領域変換部36とを例えば備えている。この復号装置により実現される第一実施形態の復号方法の各処理の例を図10に示す。

【0 1 4 6】

復号装置では、符号化装置による符号化処理に対応する処理で時間領域の復号された音信号が得られる。

【0 1 4 7】

復号装置には、符号化装置が出力した、正規化された周波数領域サンプル列に対応する符号及び伸縮線形予測係数符号が少なくとも入力される。以下、正規化された周波数領域

10

20

30

40

50

サンプル列に対応する符号として、正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ に対応する符号が入力された場合を例に挙げて説明する。

【0148】

以下、図9の各部について説明する。

【0149】

<伸縮線形予測係数復号部31>

伸縮線形予測係数復号部31には、符号化装置が出力した伸縮線形予測係数符号が入力される。

【0150】

伸縮線形予測係数復号部31は、フレームごとに、入力された伸縮線形予測係数符号を例えば従来の復号技術によって復号して復号伸縮線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ を得る(ステップD1)。

10

【0151】

得られた復号伸縮線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ は、伸縮パワースペクトル包絡系列生成部32に出力される。

【0152】

ここで、従来の復号技術とは、例えば、線形予測係数符号が量子化された線形予測係数に対応する符号である場合に線形予測係数符号に対応する量子化された線形予測係数を復号された線形予測係数として得る技術、線形予測係数符号が量子化されたLSPパラメータに対応する符号である場合に線形予測係数符号に対応する量子化されたLSPパラメータを復号されたLSPパラメータとして得て、復号されたLSPパラメータを線形予測係数に変換して復号された線形予測係数を得る技術などである。また、量子化された線形予測係数と量子化されたLSPパラメータは互に変換可能なものであり、入力された予測係数符号と後段での処理において必要な情報に応じて、変換処理を行えばよいのは周知である。以上から、上記の線形予測係数符号の復号処理と必要に応じて行なう上記の変換処理とを包含したものが「従来の復号技術による復号」ということになる。なお、ここでは入力される線形予測係数符号が伸縮線形予測係数符号であるが、処理は従来の復号処理と同様である。

20

【0153】

<伸縮パワースペクトル包絡系列生成部32>

伸縮パワースペクトル包絡系列生成部32には、伸縮線形予測係数復号部31が生成した復号伸縮線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ が入力される。

30

【0154】

伸縮パワースペクトル包絡系列生成部32は、復号伸縮線形予測係数 $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ を用いて、符号化装置の伸縮パワースペクトル包絡系列生成部24と同様の処理により、非線形離散化サンプル点列に基づいて表現される伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ を生成する(ステップD2)。

【0155】

生成された伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ は、逆伸縮変換部33に出力される。

40

【0156】

<逆伸縮変換部33>

逆伸縮変換部33には、伸縮パワースペクトル包絡系列生成部32が生成した伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ が入力される。

【0157】

逆伸縮変換部33は、符号化装置の逆伸縮変換部25と同様の処理により、非線形離散化サンプル点列に基づいて表現される伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ を線形離散化サンプル点列に基づいて表現されるパワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ に変換する(ステップD3)。

【0158】

50

変換されたパワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ は、包絡逆正規化部 3 5 に出力される。

【 0 1 5 9 】

< 復号部 3 4 >

復号部 3 4 には、符号化装置が出力した正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ に対応する符号が入力される。

【 0 1 6 0 】

復号部 3 4 は、フレームごとに、入力された正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ に対応する符号を復号して復号正規化MDCT係数列 $\hat{X}_N(0), \hat{X}_N(1), \dots, \hat{X}_N(N-1)$ を生成する(ステップD 4)。

10

【 0 1 6 1 】

生成された復号正規化MDCT係数列 $\hat{X}_N(0), \hat{X}_N(1), \dots, \hat{X}_N(N-1)$ は、包絡逆正規化部 3 5 に出力される。

【 0 1 6 2 】

例えば、符号化装置でライス符号化を用いた場合には、復号部 3 4 は、ライス符号化に対応した復号処理により符号を復号する。

【 0 1 6 3 】

正規化MDCT係数列 $X_N(0), X_N(1), \dots, X_N(N-1)$ に対応する符号として利得符号及び整数信号符号が入力された場合には、復号部 3 4 は、整数信号符号を復号することにより得られる復号正規化係数列 $X_Q(0), X_Q(1), \dots, X_Q(N-1)$ に利得符号により特定される利得を乗じることにより復号正規化MDCT係数列 $\hat{X}_N(0), \hat{X}_N(1), \dots, \hat{X}_N(N-1)$ を生成する。

20

【 0 1 6 4 】

< 包絡逆正規化部 3 5 >

包絡逆正規化部 3 5 には、逆伸縮変換部 3 3 が変換したパワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ 及び復号部 3 4 が生成した復号正規化MDCT係数列 $\hat{X}_N(0), \hat{X}_N(1), \dots, \hat{X}_N(N-1)$ が入力される。

【 0 1 6 5 】

包絡逆正規化部 3 5 は、パワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ を用いて、正規化された周波数領域のサンプル列である復号正規化MDCT係数列 $\hat{X}_N(0), \hat{X}_N(1), \dots, \hat{X}_N(N-1)$ を逆正規化することにより、復号MDCT係数列 $\hat{X}(0), \hat{X}(1), \dots, \hat{X}(N-1)$ を生成する(ステップD 5)。

30

【 0 1 6 6 】

生成された復号MDCT係数列 $\hat{X}(0), \hat{X}(1), \dots, \hat{X}(N-1)$ は、時間領域変換部 3 6 に出力される。

【 0 1 6 7 】

例えば、包絡逆正規化部 3 5 は、 $i=0, 1, \dots, N-1$ として、復号正規化MDCT係数列 $\hat{X}_N(0), \hat{X}_N(1), \dots, \hat{X}_N(N-1)$ の各係数に $\hat{X}_N(i)$ に、パワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ の各包絡値 $W(i)$ の平方根を乗じることにより復号MDCT係数列 $\hat{X}(0), \hat{X}(1), \dots, \hat{X}(N-1)$ を生成する。すなわち、 $i=0, 1, \dots, N-1$ として、 $\hat{X}(i)=\hat{X}_N(i)*\sqrt{W(i)}$ である。ここで、 x を実数として \sqrt{x} は x の平方根を表す。

40

【 0 1 6 8 】

なお、パワースペクトル包絡系列 $W(0), W(1), \dots, W(N-1)$ は、伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ に由来するものである。

【 0 1 6 9 】

したがって、包絡逆正規化部 3 5 は、伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ に基づいて周波数領域のサンプル列である復号正規化MDCT係数列 $\hat{X}_N(0), \hat{X}_N(1), \dots, \hat{X}_N(N-1)$ を逆正規化することにより、周波数領域サンプル列を生成しているとも言える。

【 0 1 7 0 】

< 時間領域変換部 3 6 >

50

時間領域変換部 3 6 には、包絡逆正規化部 3 5 が生成した復号 MDCT 係数列 $\hat{X}(0), \hat{X}(1), \dots, \hat{X}(N-1)$ が入力される。

【 0 1 7 1 】

時間領域変換部 3 6 は、フレームごとに、包絡逆正規化部 3 5 で得た復号 MDCT 係数列を時間領域に変換してフレーム単位の音信号（復号音信号）を得て、（ステップ D 6）出力する。

【 0 1 7 2 】

[第二実施形態]

（第二実施形態の符号化）

第二実施形態の符号化装置の構成例は、図 2 に示した第一実施形態の符号化装置の構成例と同様である。

10

【 0 1 7 3 】

以下、第一実施形態と異なる部分を中心に説明する。第一実施形態と同様の部分については説明を省略する。

【 0 1 7 4 】

第二実施形態の符号化装置は、MDCT 係数列 $X(0), X(1), \dots, X(N-1)$ を正規化する際に用いるパワースペクトル包絡系列が異なる。すなわち、非線形離散化サンプル点列に対応する伸縮パワースペクトル包絡を平滑化するときの平滑化方法が異なる。言い換えれば、伸縮平滑化パワースペクトル包絡系列生成部 2 4 による伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ の生成方法が異なる。

20

【 0 1 7 5 】

第一実施形態の符号化装置では、伸縮パワースペクトル包絡系列 $\sim W_0(0), \sim W_0(1), \dots, \sim W_0(N-1)$ の一例である伸縮平滑化パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ は、線形離散化サンプル点列に対応するパワースペクトル包絡系列を平滑化する従来と同様の方法で、非線形離散化サンプル点列に対応する伸縮疑似パワースペクトル包絡系列を平滑化することにより生成されている。

【 0 1 7 6 】

このようにして生成された非線形離散化サンプル点列に対応する伸縮パワースペクトル包絡系列を線形離散化サンプル点列に対応するパワースペクトル包絡系列に逆伸縮変換すると、線形離散化サンプル点列中の解像度の高い周波数領域では伸縮パワースペクトル包絡系列に対して施した平滑化の効果が相殺され、線形離散化サンプル点列に対応するパワースペクトル包絡系列に逆伸縮変換したときにピークの形が大きく残ってしまうことがある。その結果、ピークの形が大きく残ったパワースペクトル包絡系列を用いて正規化 MDCT 係数列を求めて符号化することになるので、平滑化効果が十分に得られず、符号化の効率が低下してしまうことがある。

30

【 0 1 7 7 】

このため、第二実施形態の伸縮パワースペクトル包絡系列生成部 2 4 は、逆伸縮変換後のパワースペクトル包絡系列が、線形離散化サンプル点列で均一な解像度で表現されたパワースペクトル包絡を平滑化したときの平滑化パワースペクトル（従来の平滑化パワースペクトル）を近似するものとなるように、非線形離散化サンプル点列の伸縮の度合い $g(k)$ に応じて平滑化の効果を補正する。

40

【 0 1 7 8 】

具体的には、第二実施形態の伸縮パワースペクトル包絡系列生成部 2 4 は、伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ として、式 (3 ') により定義される伸縮平滑化パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ を生成する（ステップ E 4）。生成された伸縮平滑化パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ は、伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ として逆伸縮変換部 2 5 に出力される。

【数 1 4】

$$\tilde{W}_\gamma(k) = \frac{\sigma^2}{2\pi} \frac{1}{\left| 1 + \sum_{n=1}^p \hat{\beta}_n \gamma^{ng(k)} \exp(-j2\pi kn/N) \right|^2}, \quad k=0,1,\dots,N-1 \quad \dots(3')$$

【 0 1 7 9】

式(3')により定義される伸縮平滑化パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ は、 $g(k) (k=0, 1, \dots, N-1)$ により補正された伸縮平滑化パワースペクトル包絡系列であることから、式(3')により定義される伸縮平滑化パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ のことを、補正された伸縮平滑化パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ とも呼ぶ。

10

【 0 1 8 0】

式(3)と式(3')とを比較すると、補正係数 γ に $g(k)$ が乗じられている点異なる。 $g(k)$ は、非線形離散化サンプル点列の線形離散化サンプル点列からの伸縮の度合いに対応する値であり、例えば以下のように定義される。

【数 1 5】

$$g(k) = \begin{cases} f(k)/k & \text{for } 0 < k < N-1 \\ 1 & \text{for } k = 0 \end{cases}$$

【 0 1 8 1】

$f(k)$ は、非線形離散化サンプル点列で k 番目のインデックスに対応するサンプル点の周波数の、線形離散化サンプル点列での相対的な周波数位置を表すものである。したがって、式(3')は、逆伸縮変換の際に間隔を縮める周波数では、補正值 γ を等価的に小さくすることを意味する。なお、非線形離散化サンプル点列から線形離散化サンプル点列への変換、すなわち、逆伸縮変換、を実現する行列を変換行列 V とすれば、 $f(k) (k=0, 1, \dots, N-1)$ は例えば以下のように表される。

20

【数 1 6】

$$\begin{pmatrix} f(0) \\ f(1) \\ \vdots \\ f(N-1) \end{pmatrix} = V \begin{pmatrix} 0 \\ 1 \\ \vdots \\ N-1 \end{pmatrix}$$

30

【 0 1 8 2】

なお、変換行列 V は、変換行列 U と同様に帯状の行列で表現できる。変換行列 V も変換行列 U と同様に、非線形離散化サンプル点列と線形離散化サンプル点列との相関関係から予め学習などにより求めておくことができる。求め方は第一実施形態で説明したものと同様である。なお、離散信号における周波数の非線形変換は不可逆な演算であるため、必ずしも V が U の逆行列の関係にあるものではない。

【 0 1 8 3】

(第二実施形態の復号)

第二実施形態の復号装置の構成例は、図9に示した第一実施形態の復号装置の構成例と同様である。

40

【 0 1 8 4】

第二実施形態の復号装置は、伸縮パワースペクトル包絡系列生成部32が、ステップD2において、伸縮パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ として、式(3')により定義される補正された伸縮平滑化パワースペクトル包絡系列 $\sim W(0), \sim W(1), \dots, \sim W(N-1)$ を生成する部分で第一実施形態の復号装置と異なる。

【 0 1 8 5】

第二実施形態の復号装置は、他の部分については、第一実施形態の復号装置と同様であ

50

る。

【 0 1 8 6 】

[第三実施形態]

(第三実施形態の符号化)

第三実施形態の符号化装置の構成例を図 1 1 に示す。第三実施形態の符号化装置は、図 1 1 に示すように、周波数領域変換部 2 1 と、伸縮疑似パワースペクトル系列生成部 2 2 と、逆伸縮変換部 2 5 と、包絡正規化部 2 6 と、符号化部 2 7 とを、第一実施形態及び第二実施形態の符号化装置と同様に備え、さらに線形予測分析部 3 2 3 と、伸縮パワースペクトル包絡系列生成部 3 2 4 と、伸縮対応平滑化 L S P 線形変換部 3 2 8 を例えば備えている。線形予測分析部 3 2 3 は、線形予測係数算出部 3 2 3 1 と、L S P 計算部 3 2 3 2 と、L S P 符号化部 3 2 3 3 とを例えば備えている。この符号化装置により実現される第三実施形態の符号化方法の各処理の例を図 1 2 に示す。

10

【 0 1 8 7 】

以下、図 1 1 の各部について、第一実施形態及び第二実施形態と異なる部分を中心に説明する。第一実施形態及び第二実施形態と同様の部分については説明を省略する。

【 0 1 8 8 】

< 線形予測分析部 3 2 3 >

線形予測分析部 3 2 3 には、伸縮疑似パワースペクトル系列生成部 2 2 が生成した伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ が入力される。

【 0 1 8 9 】

線形予測分析部 2 3 は、伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を用いて、線形予測係数算出部 3 2 3 1 と L S P 計算部 3 2 3 2 と L S P 符号化部 3 2 3 3 との処理を実行して、伸縮 L S P 符号 CL_f と量子化伸縮 L S P パラメータ $\wedge [1], \wedge [2], \dots, \wedge [p]$ とを生成し (ステップ E 3)、出力する。

20

【 0 1 9 0 】

< 線形予測係数算出部 3 2 3 1 >

線形予測係数算出部 3 2 3 1 には、伸縮疑似パワースペクトル生成部 2 2 が生成した伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ が入力される。

【 0 1 9 1 】

線形予測係数算出部 3 2 3 1 は、伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を用いて、以下の式により定義される $\sim X(0), \sim X(1), \dots, \sim X(N-1)$ を線形予測分析して伸縮線形予測係数 $\wedge_1, \wedge_2, \dots, \wedge_p$ を生成する。

30

【 数 1 7 】

$$\tilde{X}(k) = \sum_{n=0}^{N-1} \tilde{Y}(n) \exp\left(-j \frac{2\pi kn}{N}\right), \quad k = 0, 1, \dots, N-1$$

【 0 1 9 2 】

生成された伸縮線形予測係数 $\wedge_1, \wedge_2, \dots, \wedge_p$ は、L S P 計算部 3 2 3 2 に出力される。

【 0 1 9 3 】

< L S P 計算部 3 2 3 2 >

L S P 計算部 3 2 3 2 は、線形予測係数算出部 3 2 3 1 が生成した伸縮線形予測係数 $\wedge_1, \wedge_2, \dots, \wedge_p$ を受け取り、伸縮線形予測係数 $\wedge_1, \wedge_2, \dots, \wedge_p$ から伸縮 L S P パラメータ $[1], [2], \dots, [p]$ を生成する。

40

【 0 1 9 4 】

生成された伸縮 L S P パラメータ $[1], [2], \dots, [p]$ は、L S P 符号化部 3 2 3 3 に出力される。

【 0 1 9 5 】

< L S P 符号化部 3 2 3 3 >

L S P 符号化部 3 2 3 3 は、L S P 計算部 3 2 3 2 が生成した伸縮 L S P パラメータ $[1], [2], \dots, [p]$ を受け取り、伸縮 L S P パラメータ $[1], [2], \dots, [p]$ を符号化し

50

て、伸縮 L S P 符号 CL_f と、伸縮 L S P 符号に対応する量子化伸縮 L S P パラメータ $\hat{a} [1], \hat{a} [2], \dots, \hat{a} [p]$ とを生成する。なお、量子化伸縮 L S P パラメータは、伸縮 L S P パラメータを量子化したものである。「ITU-T Recommendation G.729」, ITU, 1996 (参考文献 1) では、伸縮 L S P パラメータ $[1], [2], \dots, [p]$ の過去のフレームからの重み付き差分ベクトルを求め、重み付き差分ベクトルを低次側と高次側の 2 つのサブベクトルに分け、各サブベクトルが 2 つの符号帳からのサブベクトルの和となるように符号化する方法で符号化しているが、符号化方法には様々な従来技術がある。したがって、伸縮 L S P パラメータの符号化には、参考文献 1 に記載された方法、多段でベクトル量子化する方法、スカラ量子化する方法、これらを組み合わせた方法、などの様々な周知の符号化方法が採用されることがある。

10

【 0 1 9 6 】

生成された量子化伸縮 L S P パラメータ $\hat{a} [1], \hat{a} [2], \dots, \hat{a} [p]$ は、伸縮対応平滑化 L S P 線形変換部 3 2 8 に出力される。

【 0 1 9 7 】

また、生成された伸縮 L S P 符号 CL_f は、復号装置に送信される。

【 0 1 9 8 】

< 伸縮対応平滑化 L S P 線形変換部 3 2 8 >

伸縮対応平滑化 L S P 線形変換部 3 2 8 では、L S P パラメータの性質を利用して、L S P 符号化部 3 2 3 が生成した量子化伸縮 L S P パラメータ $\hat{a} [1], \hat{a} [2], \dots, \hat{a} [p]$ に近似的な線形変換を施して近似量子化伸縮平滑化 L S P パラメータ $\hat{a} [1], \hat{a} [2], \dots, \hat{a} [p]$ を生成し (ステップ E 8)、出力する。

20

【 0 1 9 9 】

まず、L S P パラメータの性質を説明する。

【 0 2 0 0 】

入力された時間領域の音信号から求まる線形予測係数を $a[i] [i=1, \dots, P]$ とし、周波数方向で均等な間隔のサンプル点を n としたときの、線形予測係数 $a[i] [i=1, \dots, P]$ に対応するパワースペクトル包絡系列を

【 数 1 8 】

$$W(n) = \frac{\sigma^2}{2\pi \left| 1 + \sum_{i=1}^p a[i] \cdot \exp(-j2\pi kn/N) \right|^2} \quad \dots(4)$$

30

とし、線形予測係数 $a[i] [i=1, \dots, P]$ のそれぞれに対して γ^i を乗算することに対応する平滑化後、すなわち、一般の平滑化後のパワースペクトル包絡系列を

【 数 1 9 】

$$W_\gamma(n) = \frac{\sigma^2}{2\pi \left| 1 + \sum_{i=1}^p a[i] \gamma^i \cdot \exp(-j2\pi kn/N) \right|^2} \quad \dots(5)$$

40

とした場合、線形予測係数 $a[i] [i=1, \dots, P]$ に対応する L S P パラメータの性質として、以下が成り立つ。ここで、 $\exp(\cdot)$ はネイピア数を底とする指数関数であり、 j は虚数単位であり、 σ^2 は予測残差エネルギーである。また、 γ は補正係数であり、予め定めた 1 以下の正の整数である。なお、式 (5) は式 (4) の $a[i]$ を $a[i] \times \gamma^i$ に置き換えた式であり、式 (5) で定義される系列 $W(1), W(2), \dots, W(N)$ は、式 (4) で定義される入力された時間領域の音信号のパワースペクトル包絡系列 $W(1), W(2), \dots, W(N)$ の振幅の凹凸に一般の平滑化を施したものに相当する。

【 0 2 0 1 】

線形予測係数 $a[i] [i=1, \dots, P]$ に対応する L S P パラメータを $\hat{a} [i] [i=1, \dots, P]$ として

50

、L S Pパラメータ列 [1], [2], ..., [p]は入力された時間領域の音信号のパワースペクトル包絡と相関性のある周波数領域のパラメータ列である。L S Pパラメータ列の各値は入力された時間領域の音信号のパワースペクトル包絡の極値の周波数位置と相関する。

[i]と [i+1]の間の周波数位置にパワースペクトル包絡の極値が存在し、この極値の周りの接線の傾きが急峻であるほど [i]と [i+1]との間隔(つまり、[i+1]- [i]の値)が小さくなる。すなわち、パワースペクトル包絡の振幅の凹凸が急峻であるほど、各i (i=1,2,...,p-1)について、[i]と [i+1]との間隔が不均一になる。逆に、パワースペクトル包絡の凹凸がほとんどない場合は、各iについて、[i]と [i+1]との間隔が均等間隔に近くなる。

【0202】

10

補正係数 γ が小さいほど、式(5)で定義される一般の平滑化後パワースペクトル包絡系列 $W(1), W(2), \dots, W(N)$ の振幅の凹凸は、式(4)で定義されるパワースペクトル包絡系列 $W(1), W(2), \dots, W(N)$ の振幅の凹凸と比較してなだらかになる。したがって、式(5)で定義される一般の平滑化後のパワースペクトル包絡系列 $W(1), W(2), \dots, W(N)$ に対応するL S Pパラメータ、すなわち、線形予測係数列 $a[1], a[2], \dots, a[p]$ の各係数 $a[i]$ (i=1, ..., p)に補正係数 γ のi乗を乗じた補正済線形予測係数 $a_{\gamma}[i] = a[i] \times \gamma^i$ の系列をL S Pパラメータに変換して得られるL S Pパラメータ、を $\theta_{\gamma}[i]$ (i=1, 2, ..., p)としたとき、補正係数 γ の値が小さいほど $\theta_{\gamma}[i]$ と $\theta_{\gamma}[i+1]$ との間隔が均等間隔に近くなると言える。なお、 $\gamma=0$ の影響がない($\gamma=0$ の)ときは、パワースペクトル包絡が平坦な場合に相当する。

20

【0203】

補正係数 $\gamma=0$ としたときのL S Pパラメータ $\theta_0[1], \theta_0[2], \dots, \theta_0[p]$ は、

【数20】

$$\theta_{\gamma=0}[i] = \frac{i\pi}{p+1}$$

となり、すべてのi=1, ..., p-1について $\theta_0[i]$ と $\theta_0[i+1]$ の間隔が等間隔になる。また、 $\gamma=1$ としたとき、L S Pパラメータ列 $\theta_1[1], \theta_1[2], \dots, \theta_1[p]$ とL S Pパラメータ列 [1], [2], ..., [p]は等価である。なお、L S Pパラメータ $\theta_1[i]$ (i=1, 2, ..., p)は、

$$0 < \theta_1[1] < \theta_1[2] < \dots < \theta_1[p]$$

の性質を満たす。

30

【0204】

図13は、補正係数 γ とL S Pパラメータ $\theta_{\gamma}[i]$ (i=1, 2, ..., p)の関係の一例である。横軸は補正係数 γ の値であり、縦軸は補正係数 γ に対応するL S Pパラメータ $\theta_{\gamma}[i]$ (i=1, 2, ..., p)の値を表す。予測次数p=16として、下から順に $\theta_{\gamma}[1], \theta_{\gamma}[2], \dots, \theta_{\gamma}[16]$ の値を図示したものである。各 $\theta_{\gamma}[i]$ の値は、ある音響信号を線形予測分析して得た線形予測係数列 $a[1], a[2], \dots, a[p]$ の各係数 $a[i]$ (i=1, ..., p)に補正係数 γ のi乗を乗じた補正済線形予測係数 $a_{\gamma}[i] = a[i] \times \gamma^i$ の系列を各 $\theta_{\gamma}[i]$ の値ごとに求め、その補正済線形予測係数列 $a_{\gamma}[1], a_{\gamma}[2], \dots, a_{\gamma}[p]$ をL S Pパラメータに変換して得たものである。なお、 $\gamma=1$ のときの $\theta_1[i]$ は $\theta_0[i]$ と等価である。

40

【0205】

図13に示されているように、 $0 < \gamma < 1$ として、L S Pパラメータ $\theta_{\gamma}[i]$ は、 $\theta_0[i]$ と $\theta_1[i]$ の内分点になる。横軸を補正係数 γ の値とし、縦軸をL S Pパラメータの値とする二次元平面において、各L S Pパラメータ $\theta_{\gamma}[i]$ は、局所的に見れば $\theta_{\gamma}[i]$ の増加または減少に対して線形な関係にある。異なる2つの補正係数 γ_1, γ_2 ($0 < \gamma_1 < \gamma_2 < 1$)として、二次元平面上の点 $(\gamma_1, \theta_{\gamma_1}[i])$ と点 $(\gamma_2, \theta_{\gamma_2}[i])$ を結ぶ直線の傾きの大きさは、L S Pパラメータ列 $\theta_1[1], \theta_1[2], \dots, \theta_1[p]$ 中の $\theta_1[i]$ の前後のL S Pパラメータ(つまり、 $\theta_1[i-1]$ と $\theta_1[i+1]$)と $\theta_1[i]$ との相対的な間隔と相関性がある。具体的には、

【数 2 1】

$$|\theta_{\gamma_1}[i] - \theta_{\gamma_1}[i-1]| > |\theta_{\gamma_1}[i+1] - \theta_{\gamma_1}[i]| \quad \dots (6)$$

である場合、

【数 2 2】

$$|\theta_{\gamma_2}[i+1] - \theta_{\gamma_2}[i]| < |\theta_{\gamma_1}[i+1] - \theta_{\gamma_1}[i]|$$

かつ

$$|\theta_{\gamma_2}[i] - \theta_{\gamma_2}[i-1]| > |\theta_{\gamma_1}[i] - \theta_{\gamma_1}[i-1]| \quad \dots (7)$$

という性質が成り立ち、

【数 2 3】

$$|\theta_{\gamma_1}[i] - \theta_{\gamma_1}[i-1]| < |\theta_{\gamma_1}[i+1] - \theta_{\gamma_1}[i]| \quad \dots (8)$$

である場合、

【数 2 4】

$$|\theta_{\gamma_2}[i+1] - \theta_{\gamma_2}[i]| > |\theta_{\gamma_1}[i+1] - \theta_{\gamma_1}[i]|$$

かつ

$$|\theta_{\gamma_2}[i] - \theta_{\gamma_2}[i-1]| < |\theta_{\gamma_1}[i] - \theta_{\gamma_1}[i-1]| \quad \dots (9)$$

という性質が成り立つ。

【0 2 0 6】

式(6)(7)は、 $\theta_{\gamma_1}[i]$ が $\theta_{\gamma_1}[i+1]$ と $\theta_{\gamma_1}[i-1]$ の midpoint よりも $\theta_{\gamma_1}[i+1]$ 寄りの場合は、 $\theta_{\gamma_2}[i]$ はさらに $\theta_{\gamma_2}[i+1]$ 寄りの値となることを示す(図14参照)。このことは、横軸を θ_{γ_1} の値とし、縦軸をLSPパラメータの値とする二次元平面上における点(0, $\theta_{\gamma_2=0}[i]$)と点(1, $\theta_{\gamma_1}[i]$)を結ぶ直線L1の傾きよりも、点(1, $\theta_{\gamma_1}[i]$)と点(2, $\theta_{\gamma_2}[i]$)を結ぶ直線L2の傾きの方が大きいことを意味する(図15参照)。

【0 2 0 7】

式(8)(9)は、 $\theta_{\gamma_1}[i]$ が $\theta_{\gamma_1}[i+1]$ と $\theta_{\gamma_1}[i-1]$ の midpoint よりも $\theta_{\gamma_1}[i-1]$ 寄りのときは、 $\theta_{\gamma_2}[i]$ はさらに $\theta_{\gamma_2}[i-1]$ 寄りの値となることを示す。このことは、横軸を θ_{γ_1} の値とし、縦軸をLSPパラメータの値とする二次元平面上における点(0, $\theta_{\gamma_2=0}[i]$)と点(1, $\theta_{\gamma_1}[i]$)を結ぶ直線の傾きよりも、点(1, $\theta_{\gamma_1}[i]$)と点(2, $\theta_{\gamma_2}[i]$)を結ぶ直線の傾きの方が小さいことを意味する。

【0 2 0 8】

以上の性質に基づけば、 $\theta_{\gamma_1}[1], \theta_{\gamma_1}[2], \dots, \theta_{\gamma_1}[p]$ と $\theta_{\gamma_2}[1], \theta_{\gamma_2}[2], \dots, \theta_{\gamma_2}[p]$ の関係は、 $\theta_{\gamma_1} = (\theta_{\gamma_1}[1], \theta_{\gamma_1}[2], \dots, \theta_{\gamma_1}[p])^T$ とし、 $\theta_{\gamma_2} = (\theta_{\gamma_2}[1], \theta_{\gamma_2}[2], \dots, \theta_{\gamma_2}[p])^T$ とし、 $\theta_{\gamma_2=0} = (\theta_{\gamma_2=0}[1], \theta_{\gamma_2=0}[2], \dots, \theta_{\gamma_2=0}[p])^T$ とし、式(10)でモデル化することができる。

【数 2 5】

$$\theta_{\gamma_2} \approx K(\theta_{\gamma_1} - \theta_{\gamma_2=0})(\gamma_2 - \gamma_1) + \theta_{\gamma_1} \quad \dots (10)$$

ただし、Kは式(11)で定義される $p \times p$ 行列である。

10

20

30

40

【数 2 6】

$$K = \begin{pmatrix} x_1 & y_1 & & & & 0 \\ z_2 & x_2 & y_2 & & & \\ & z_3 & x_3 & y_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \\ 0 & & & & z_p & x_p \end{pmatrix} \quad \dots(11)$$

【 0 2 0 9】

10

ここでは、 $0 < \gamma_1 < \gamma_2$ 、かつ、 $\gamma_1 < \gamma_2$ である。式(6)~(9)では $\gamma_1 < \gamma_2$ と仮定して関係性を記述したが、式(10)のモデルでは γ_1 と γ_2 の大小関係に制限はなく、 $\gamma_1 < \gamma_2$ であっても $\gamma_1 > \gamma_2$ であってもよい。

【 0 2 1 0】

行列Kは対角成分とその近傍の要素のみ非零の値を持つ帯行列であり、対角成分に対応するLSPパラメータとそれに隣接するLSPパラメータとの間に成り立つ上述の相関関係を表現する行列である。なお、式(11)では帯幅3の帯行列を例示したが、帯幅は3に限定されない。

【 0 2 1 1】

20

ここで、

【数 2 7】

$$\tilde{\theta}_{\gamma_2} = K(\theta_{\gamma_1} - \theta_{\gamma=0})(\gamma_2 - \gamma_1) + \theta_{\gamma_1} \quad \dots(10a)$$

とすれば、

$$\tilde{\theta}_{\gamma_2} = (\tilde{\theta}_{\gamma_2}[1], \tilde{\theta}_{\gamma_2}[2], \dots, \tilde{\theta}_{\gamma_2}[p])^T$$

は θ_{γ_2} の近似値である。

【 0 2 1 2】

式(10a)を展開すると以下の式(12)が得られる。

【数 2 8】

30

$$\tilde{\theta}_{\gamma_2}[i] = z_i(\theta_{\gamma_1}[i-1] - \theta_{\gamma=0}[i-1]) + y_i(\theta_{\gamma_1}[i+1] - \theta_{\gamma=0}[i+1]) + x_i(\theta_{\gamma_1}[i] - \theta_{\gamma=0}[i]) + \theta_{\gamma_1}[i] \quad \dots(12)$$

ただし、 $i=2, \dots, p-1$ とする。

【 0 2 1 3】

横軸を γ の値とし、縦軸をLSPパラメータの値とする二次元平面上の点($\gamma_1, \theta_{\gamma_1}[i]$)と点($0, \theta_{\gamma=0}[i]$)を結ぶ直線L1の延長上の γ_2 に対応する縦軸の値、つまり、 $\theta_{\gamma_1}[i]$ と $\theta_{\gamma=0}[i]$ を結ぶ直線L1の傾きから直線近似したときの γ_2 に対応する縦軸の値を $\tilde{\theta}_{\gamma_2}[i]$ とする(図15参照)。すると、

【数 2 9】

40

$$\tilde{\theta}_{\gamma_2}[i] = \frac{\theta_{\gamma_1}[i] - \theta_{\gamma=0}[i]}{\gamma_1}(\gamma_2 - \gamma_1) + \theta_{\gamma_1}[i]$$

が成り立つ。 $\gamma_1 > \gamma_2$ ならば直線補間、 $\gamma_1 < \gamma_2$ ならば直線外挿を意味する。

【 0 2 1 4】

式(11)において、

【数 3 0】

$$x_i = \frac{1}{\gamma_1}, y_i = 0, z_i = 0$$

とすれば、 $\tilde{\theta}_{\gamma_2}(i) = \theta_{\gamma_1}[i]$ となり、式(10a)のモデルにより得られる $\tilde{\theta}_{\gamma_2}[i]$ 50

]は、二次元平面上の点(1, 1[i])と点(0, =0[i])を結ぶ直線により直線近似した場合の 2に対応する L S P パラメータの値の推定値 2[i]と一致する。

【 0 2 1 5 】

u_i, v_i を 1 以下の正の値として、上述の式 (1 1) において、

【数 3 1】

$$x_i = u_i + v_i + \frac{\gamma_2 - \gamma_1}{\gamma_1}, y_i = -v_i, z_i = -u_i \quad \dots(13)$$

とすれば、式 (1 2) は以下のように書き換えることができる。

【数 3 2】

$$\begin{aligned} \tilde{\theta}_{\gamma_2}[i] &= u_i(\theta_{\gamma_1}[i] - \theta_{\gamma=0}[i] - (\theta_{\gamma_1}[i-1] - \theta_{\gamma=0}[i-1])) \\ &\quad + v_i(\theta_{\gamma_1}[i] - \theta_{\gamma=0}[i] - (\theta_{\gamma_1}[i+1] - \theta_{\gamma=0}[i+1])) \\ &\quad + \frac{\gamma_2 - \gamma_1}{\gamma_1}(\theta_{\gamma_1}[i] - \theta_{\gamma=0}[i]) + \theta_{\gamma_1}[i] \\ &= u_i(\theta_{\gamma_1}[i] - \theta_{\gamma_1}[i-1] - (\theta_{\gamma=0}[i] - \theta_{\gamma=0}[i-1])) \\ &\quad + v_i(\theta_{\gamma_1}[i] - \theta_{\gamma_1}[i+1] - (\theta_{\gamma=0}[i] - \theta_{\gamma=0}[i+1])) + \bar{\theta}_{\gamma_2}[i] \\ &= u_i\left(\theta_{\gamma_1}[i] - \theta_{\gamma_1}[i-1] - \frac{\pi}{p+1}\right) - v_i\left(\theta_{\gamma_1}[i+1] - \theta_{\gamma_1}[i] - \frac{\pi}{p+1}\right) + \bar{\theta}_{\gamma_2}[i] \quad \dots(14) \end{aligned}$$

【 0 2 1 6 】

式 (1 4) は、L S P パラメータ列 1[1], 1[2], ..., 1[p] 中の i 番目の L S P パラメータ 1[i] の前後の L S P パラメータの値との差 (すなわち、1[i] - 1[i-1] と 1[i+1] - 1[i]) の重み付けで 2[i] の値を補正し、~ 2[i] を得ることを意味する。つまり、上述の式 (6) ~ (9) のような相関性が式 (1 0 a) の行列 K の帯部分の要素 (非零要素) に反映されていることになる。

【 0 2 1 7 】

なお、式 (1 0 a) により得られる ~ 2[1], ~ 2[2], ..., ~ 2[p] は線形予測係数列 a[1] x (2), a[2] x (2)², ..., a[p] x (2)^p を L S P パラメータに変換したときの L S P パラメータの値 2[1], 2[2], ..., 2[p] の近似値 (推定値) である。

【 0 2 1 8 】

また、特に 2 > 1 の場合には、式 (1 3) (1 4) に示されているように、式 (1 1) の行列 K は対角成分が正の値を持ち、その近傍の要素が負の値を持つ傾向がある。逆に、1 > 2 の場合には、行列 K は対角成分が負の値を持ち、その近傍の要素が正の値を持つ傾向がある。

【 0 2 1 9 】

行列 K は予め設定しておく行列であり、例えば、学習データを用いて予め学習したものをを用いる。行列 K の学習方法については後述する。

【 0 2 2 0 】

ここで、周波数方向で不均等な間隔のサンプル点についてのパワースペクトル包絡系列は、上述の式 (2) で定義される伸縮非平滑化パワースペクトル包絡系列である。また、伸縮非平滑化パワースペクトル包絡系列に対応する平滑化後のパワースペクトル包絡系列は、上述の式 (3 ') 等で定義される伸縮平滑化パワースペクトル包絡系列である。上述の式 (3 ') で補正係数 が ng(k) 乗されていることから分かる通り、周波数方向での伸縮度合いに応じて、伸縮非平滑化パワースペクトル包絡系列の伸縮非平滑化パワースペクトル包絡系列に対する凸凹の鈍りの度合い、すなわち、平滑化度合い、が変わる。例えば対数関数を用いて伸縮を行った場合、低域は引き伸ばされることから、式 (3 ') より補正係数 に対する補正が強くなり、平滑化度合いが上がる。一方、高域は縮められることから、補正係数 に対する補正が弱まり、平滑化度合いが下がる。

【 0 2 2 1 】

この に対する補正は、L S P パラメータの各値が入力された時間領域の音信号のパワースペクトル包絡の極値の周波数位置と相関する事実を用いることにより、上記の L S P

線形変換の枠組みで近似することができる。先の対数関数による伸縮の例においては、行列K内で、平滑化度合いの上がる低域と相関を持つ伸縮LSPパラメータ、つまり低次の伸縮LSPパラメータに対応する成分の絶対値を大きくすればよい。逆に、平滑化度合いの下がる高域と相関を持つ伸縮LSPパラメータ、つまり高次の伸縮LSPパラメータに対応する成分の絶対値を小さくすることで平滑化度合いの変化を近似することができる。

【0222】

一方、行列K内で、平滑化度合いの上がる低域と相関を持つ伸縮LSPパラメータ、つまり低次の伸縮LSPパラメータに対応する成分の絶対値を小さくする、もしくは、平滑化度合いの下がる高域と相関を持つ伸縮LSPパラメータ、つまり高次の伸縮LSPパラメータに対応する成分の絶対値を大きくすることで、低次よりも高次の方がパラメータ値の変更が小さくなるような線形変換としてもよい。

10

【0223】

量子化された伸縮LSPパラメータに対しても、同様の性質が成り立つ。つまり、式(10)におけるLSPパラメータ列のベクトル $\hat{\gamma}_1$ と $\hat{\gamma}_2$ を、それぞれ量子化された伸縮LSPパラメータ列のベクトル $\hat{\gamma}_1$ と $\hat{\gamma}_2$ に置き換えることができる。具体的には、

$$\begin{aligned} \hat{\gamma}_1 &= (\hat{\gamma}_1[1], \hat{\gamma}_1[2], \dots, \hat{\gamma}_1[p])^T \text{とし、} \\ \hat{\gamma}_2 &= (\hat{\gamma}_2[1], \hat{\gamma}_2[2], \dots, \hat{\gamma}_2[p])^T \text{として、} \end{aligned}$$

【数33】

$$\hat{\Phi}_{\gamma_2} \approx K(\hat{\Phi}_{\gamma_1} - \hat{\Phi}_{\gamma=0})(\gamma_2 - \gamma_1) + \hat{\Phi}_{\gamma_1} \quad \dots(10b)$$

20

が成り立つ。

【0224】

Kが帯行列であるため、式(10)(10a)(10b)の演算に要する計算コストは非常に小さい。

【0225】

伸縮対応平滑化LSP線形変換部328は、例えば、 $\gamma_1=1$ 、 $\gamma_2=$ として(すなわち、 γ_1 の値を1とし、 γ_2 の値を固定値であるとして)、式(10b)に基づいて、量子化伸縮LSPパラメータ $\hat{\gamma}_1[1], \hat{\gamma}_1[2], \dots, \hat{\gamma}_1[p]$ から近似量子化伸縮平滑化LSPパラメータ $\hat{\gamma}_2[1], \hat{\gamma}_2[2], \dots, \hat{\gamma}_2[p]$ を生成し、出力する。

30

【0226】

[行列Kの学習方法]

伸縮対応平滑化LSP線形変換部328で用いる行列Kは以下のような方法により予め求めておき、符号化装置および復号装置内の記憶部に格納しておく。

【0227】

(step1) 予め用意したR個のフレーム単位の音声・音響信号のサンプルデータについて、各サンプルデータを周波数領域変換部21、伸縮疑似パワースペクトル系列生成部22、及び線形予測分析部323と同様の処理をして伸縮線形予測係数を得る。r番目(1 ≤ r ≤ R)のサンプルデータについて得た伸縮線形予測係数を $(r)[1], (r)[2], \dots, (r)[p]$ とし、以下では「r番目のサンプルデータに対応する伸縮線形予測係数列 $(r)[1], (r)[2], \dots, (r)[p]$ 」と呼ぶ。

40

【0228】

(step2) 各rについて、伸縮線形予測係数 $(r)[1], (r)[2], \dots, (r)[p]$ から伸縮LSPパラメータ $(r)[1], (r)[2], \dots, (r)[p]$ を求める。また、伸縮LSPパラメータ $(r)[1], (r)[2], \dots, (r)[p]$ から、LSP符号化部323と同様の方法で、量子化伸縮LSPパラメータ $\hat{\gamma}_1^{(r)}[1], \hat{\gamma}_1^{(r)}[2], \dots, \hat{\gamma}_1^{(r)}[p]$ を得る。以下では、

$$\hat{\gamma}_1^{(r)} = (\hat{\gamma}_1^{(r)}[1], \hat{\gamma}_1^{(r)}[2], \dots, \hat{\gamma}_1^{(r)}[p])^T$$

とする。

【0229】

(step3) 各rについて、伸縮線形予測係数 $(r)[1], (r)[2], \dots, (r)[p]$ と $g(k)$ と

50

を用いて式(3')により伸縮平滑化パワースペクトル包絡系列 $\sim W^{(r)}(0), \sim W^{(r)}(1), \dots, \sim W^{(r)}(N-1)$ を得る。次に、各 r について、伸縮平滑化パワースペクトル包絡系列 $\sim W^{(r)}(0), \sim W^{(r)}(1), \dots, \sim W^{(r)}(N-1)$ から、伸縮と平滑化がされた線形予測係数を得る。伸縮平滑化パワースペクトル包絡系列 $\sim W^{(r)}(0), \sim W^{(r)}(1), \dots, \sim W^{(r)}(N-1)$ から伸縮と平滑化がされた線形予測係数を得る方法は、線形予測係数算出部3231が行う伸縮疑似パワースペクトル系列 $\sim Y(0), \sim Y(1), \dots, \sim Y(N-1)$ を用いて伸縮線形予測係数 $\alpha_1, \alpha_2, \dots, \alpha_p$ を得る方法と同様である。次に、各 r について、伸縮と平滑化がされた線形予測係数から、LSP計算部3232と同様の方法でLSPパラメータに変換して伸縮平滑化LSPパラメータ $\hat{\alpha}^{(r)}[1], \hat{\alpha}^{(r)}[2], \dots, \hat{\alpha}^{(r)}[p]$ を求める。さらに、各 r について、伸縮平滑化LSPパラメータ $\hat{\alpha}^{(r)}[1], \hat{\alpha}^{(r)}[2], \dots, \hat{\alpha}^{(r)}[p]$ から、LSP符号化部3233と同様の方法で、量子化伸縮平滑化LSPパラメータ $\hat{\alpha}^{(r)}[1], \hat{\alpha}^{(r)}[2], \dots, \hat{\alpha}^{(r)}[p]$ を得る。以下では、

$$\hat{\alpha}^{(r)}_2 = (\hat{\alpha}^{(r)}[1], \hat{\alpha}^{(r)}[2], \dots, \hat{\alpha}^{(r)}[p])^T$$

とする。

【0230】

step1~3により、M組の量子化された伸縮LSPパラメータ列の組 $(\hat{\alpha}^{(r)}_1, \hat{\alpha}^{(r)}_2)$ が得られる。この集合を学習用データ集合Qとする。

$$Q = \{(\hat{\alpha}^{(r)}_1, \hat{\alpha}^{(r)}_2) \mid r=1, \dots, R\}$$

である。なお、学習用データ集合Qを生成する際に用いる補正係数 γ の値は全て共通(固定値)とする。

【0231】

(step4)学習用データQに含まれる各LSPパラメータ列の組 $(\hat{\alpha}^{(r)}_1, \hat{\alpha}^{(r)}_2)$ について、 $\gamma_1=1, \gamma_2=\gamma, \hat{\alpha}^{(r)}_1=\hat{\alpha}^{(r)}_1, \hat{\alpha}^{(r)}_2=\hat{\alpha}^{(r)}_2$ として式(10b)のモデルに代入し、自乗誤差基準で行列Kの係数を学習する。すなわち、行列Kの帯部分の成分を上から順に並べたベクトルを

【数34】

$$B = \begin{pmatrix} x_1 \\ y_1 \\ z_2 \\ x_2 \\ y_2 \\ z_3 \\ \vdots \\ x_p \end{pmatrix}$$

として、

【数35】

$$B = \frac{1}{(1-\gamma)} \left(\sum_{r=1}^R J_r^T J_r \right)^{-1} \sum_{r=1}^R J_r^T (\hat{\Phi}^{(r)} \gamma_1 - \hat{\Phi}^{(r)} \gamma_2)$$

により、Bを得る。ここで、

10

20

30

40

< 伸縮パワースペクトル包絡系列生成部 3 2 4 >

伸縮パワースペクトル包絡系列計算部 3 2 4 は、伸縮対応平滑化 L S P 線形変換部 3 2 8 から出力された近似量子化伸縮平滑化 L S P パラメータ $\hat{w}^{[1]}, \hat{w}^{[2]}, \dots, \hat{w}^{[p]}$ を用いて、式 (1 5) により伸縮パワースペクトル包絡系列 $\hat{w}(1), \hat{w}(2), \dots, \hat{w}(N)$ を求めて出力する (ステップ E 4) 。

【 数 3 7 】

$$\hat{w}(k) = \sqrt{\frac{\delta^2}{2\pi} \frac{1}{|A(\exp(j\omega_k))|^2}}$$

$$|A(\exp(j\omega_k))|^2 = \begin{cases} 2^{p-1} \left[(1 - \cos\omega_k) \prod_{n=1}^{p/2} (\cos\hat{\phi}_\gamma[2n] - \cos\omega_k)^2 + (1 + \cos\omega_k) \prod_{n=1}^{p/2} (\cos\hat{\phi}_\gamma[2n-1] - \cos\omega_k)^2 \right] & (p: \text{odd}) \\ 2^{p-1} \left[(1 - \cos\omega_k)(1 + \cos\omega_k) \prod_{n=1}^{(p-1)/2} (\cos\hat{\phi}_\gamma[2n] - \cos\omega_k)^2 + \prod_{n=1}^{(p+1)/2} (\cos\hat{\phi}_\gamma[2n-1] - \cos\omega_k)^2 \right] & (p: \text{even}) \end{cases} \quad 10$$

$$\omega_k = -\frac{2\pi k}{N} \quad \dots(15)$$

【 0 2 3 7 】

(第三実施形態の復号)

第三実施形態の符号化装置に対応する復号装置の構成例を図 1 6 に示す。第三実施形態の復号装置は、図 1 6 に示すように、逆伸縮変換部 3 3 と、復号部 3 4 と、包絡逆正規化部 3 5 と、時間領域変換部 3 6 とを、第一実施形態及び第二実施形態の復号装置と同様に備え、さらに伸縮線形予測係数復号部 4 3 1 と、伸縮パワースペクトル包絡系列生成部 4 3 2 と、伸縮対応平滑化 L S P 線形変換部 4 3 8 とを例えば備えている。この復号装置により実現される第三実施形態の復号方法の各処理の例を図 1 7 に示す。

【 0 2 3 8 】

以下、図 1 6 の各部について、第一実施形態及び第二実施形態と異なる部分を中心に説明する。第一実施形態及び第二実施形態と同様の部分については説明を省略する。

【 0 2 3 9 】

< 伸縮線形予測係数復号部 4 3 1 >

伸縮線形予測係数復号部 4 3 1 は、伸縮 L S P 符号 CL_f を受け取り、伸縮 L S P 符号 CL_f を復号して復号伸縮 L S P パラメータ $\hat{w}^{[1]}, \hat{w}^{[2]}, \dots, \hat{w}^{[p]}$ を得て出力する (ステップ D 1) 。符号化装置が出力した伸縮 L S P 符号 CL_f が誤りなく復号装置に入力されていれば、得られる復号伸縮 L S P パラメータは、符号化装置の L S P 符号化部 3 2 3 3 で得られた量子化伸縮 L S P パラメータと同じになる。

【 0 2 4 0 】

< 伸縮対応平滑化 L S P 線形変換部 4 3 8 >

伸縮対応平滑化 L S P 線形変換部 4 3 8 には、伸縮線形予測係数復号部 4 3 1 で得られた復号伸縮 L S P パラメータ $\hat{w}^{[1]}, \hat{w}^{[2]}, \dots, \hat{w}^{[p]}$ が入力される。

【 0 2 4 1 】

伸縮対応平滑化 L S P 線形変換部 4 3 8 は、伸縮対応平滑化 L S P 線形変換部 3 2 8 と同様の処理により復号伸縮 L S P パラメータ $\hat{w}^{[1]}, \hat{w}^{[2]}, \dots, \hat{w}^{[p]}$ から近似復号伸縮平滑化 L S P パラメータ $\hat{w}^{[1]}, \hat{w}^{[2]}, \dots, \hat{w}^{[p]}$ を求めて出力する (ステップ D 7) 。

【 0 2 4 2 】

< 伸縮パワースペクトル包絡系列生成部 4 3 2 >

伸縮パワースペクトル包絡系列計算部 4 3 2 は、伸縮パワースペクトル包絡系列計算部 3 2 4 と同様に、近似復号伸縮平滑化 L S P パラメータ $\hat{w}^{[1]}, \hat{w}^{[2]}, \dots, \hat{w}^{[p]}$ を用いて、上記の式 (1 5) により、伸縮パワースペクトル包絡系列 $\hat{w}(1), \hat{w}(2), \dots, \hat{w}(N)$ を求めて出力する (ステップ D 2) 。

【 0 2 4 3 】

< 第三実施形態の効果 >

第三実施形態の符号化装置及び復号装置は伸縮パワースペクトル包絡系列生成部におい

て、第一実施形態の符号化装置や復号装置よりも少ない演算量で伸縮パワースペクトル包絡系列を求めることができる。

【0244】

[第四実施形態]

第三実施形態の符号化装置が備える伸縮対応平滑化LSP線形変換部328、第三実施形態の復号装置が備える伸縮対応平滑化LSP線形変換部438を、独立した周波数領域パラメータ変換装置として構成することも可能である。

【0245】

以下では、第三実施形態の符号化装置が備える伸縮対応平滑化LSP線形変換部328、第三実施形態の復号装置が備える伸縮対応平滑化LSP線形変換部438を、独立した周波数領域パラメータ変換装置として構成する例について説明する。

【0246】

<周波数領域パラメータ変換装置>

第四実施形態の周波数領域パラメータ変換装置10は、図18に示すように、パラメータ列変換部20を例えば含み、周波数領域パラメータ [1], [2], ..., [p]を入力とし、変換後周波数領域パラメータ \sim [1], \sim [2], ..., \sim [p]を出力する。

【0247】

入力される周波数領域パラメータ [1], [2], ..., [p]は、所定の時間区間の音信号に対応する隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做したときの時間領域の信号に対応する線形予測係数に由来する周波数領域パラメータ列である。周波数領域パラメータ [1], [2], ..., [p]は、例えば、従来の符号化方法で用いたLSPパラメータ列 [1], [2], ..., [p]であってもよいし、量子化済LSPパラメータ列 \wedge [1], \wedge [2], ..., \wedge [p]であってもよい。さらに、例えば、ISPパラメータ列のような、LSPパラメータと等価な周波数領域パラメータであってもよい。ここで、線形予測係数 $a[1], a[2], \dots, a[p]$ に由来する周波数領域パラメータ列とは、LSPパラメータ列に代表されるような、線形予測係数 $a[1], a[2], \dots, a[p]$ に由来する周波数領域パラメータ [1], [2], ..., [p]の全てが0から までの間に存在し、かつ、全ての線形予測係数 $a[1], a[2], \dots, a[p]$ が0である場合には周波数領域パラメータ [1], [2], ..., [p]が0から までの間に均等間隔に存在する周波数領域パラメータ列であって、線形予測係数 $a[1], a[2], \dots, a[p]$ の予測次数と同じ個数で表されるものである。または、線形予測係数 $a[1], a[2], \dots, a[p]$ に由来する周波数領域パラメータ列とは、ISPパラメータ列に代表されるような、線形予測係数 $a[1], a[2], \dots, a[p+1]$ に由来する周波数領域パラメータ [1], [2], ..., [p+1]のうちの1次からP次までの周波数領域パラメータ [1], [2], ..., [p]の全てが0から までの間に存在し、かつ、全ての線形予測係数が0である場合には周波数領域パラメータ [1], [2], ..., [p]が0から までの間に均等間隔に存在する周波数領域パラメータ列であって、線形予測係数 $a[1], a[2], \dots, a[p+1]$ の予測次数と同じ個数で表されるものである。すなわち、所定の時間区間の音信号に対応する隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列をサンプル点の周波数の間隔が均等であると見做したときの時間領域の信号のスペクトル包絡に対応する周波数領域パラメータ列である。

なお、LSPはLSFなどと呼ばれることもあり、ISPはISFなどと呼ばれることもある。

【0248】

周波数領域パラメータ列 [1], [2], ..., [p]における各 [i] ($i=1, 2, \dots, p$)は、 α_i を1以下の正の定数とし、 $a_{i+1}[i]=a[i] \times (\alpha_i)^i$ として、 $a_{i+1}[1], a_{i+1}[2], \dots, a_{i+1}[p]$ と等価な周波数領域のパラメータまたはその量子化値であってもよい。また、周波数領域パラメータ列 [1], [2], ..., [p]における各 [i] ($i=1, 2, \dots, p$)は、 β_i を1以下の正の定数とし音信号に対応する隣接するサンプル点間の周波数の間隔が均等な周波数領域の係数列を線形離散化サンプル点列とし、音信号に対応する隣接するサンプル点間の周波数の間隔が不均等な周波数領域の係数列を非線形離散化サンプル点列とし、線形離散化サン

10

20

30

40

50

ル点列と非線形離散化サンプル点列のサンプル点数をNとし、非線形離散化サンプル点列でk番目のインデックスに対応するサンプル点の周波数の、上記線形離散化サンプル点列での相対的な周波数位置をf(k)とし、

【数38】

$$g(k) = \begin{cases} f(k)/k & \text{for } 0 < k < N-1 \\ 1 & \text{for } k = 0 \end{cases}$$

とし、 $a_{-1}[i] = a[i] \times (1)^{ig(k)}$ として、 $a_{-1}[1], a_{-1}[2], \dots, a_{-1}[p]$ と等価な周波数領域のパラメータまたはその量子化値であってもよい。

【0249】

パラメータ列変換部20は、伸縮対応平滑化LSP線形変換部328、438と同様に、LSPパラメータの性質を利用して、周波数領域パラメータ列[1],[2]...,[p-1]に近似的な線形変換を施して変換後周波数領域パラメータ列~[1],~[2],...~,~[p]を生成する。パラメータ列変換部20は、例えば、各*i*=1,2,...,pについて、以下のいずれかの方法により、変換後周波数領域パラメータ~[i]の値を求める。

【0250】

1. [i]と[i]に近接する1つまたは複数の周波数領域パラメータとの値の関係に基づく線形変換により変換後周波数領域パラメータ~[i]の値を求める。例えば、周波数領域パラメータ列[i]よりも変換後周波数領域パラメータ列~[i]のほうが、パラメータ値の間隔が均等間隔に近くなるか遠くなり、かつ、低い周波数に対応するパラメータの方が高い周波数に対応するパラメータよりも周波数領域パラメータ[i]と変換後周波数領域パラメータ~[i]との差が小さくなるように、線形変換する。または、周波数領域パラメータ列[i]よりも変換後周波数領域パラメータ列~[i]のほうが、パラメータ値の間隔が均等間隔に近くなるか遠くなり、かつ、高い周波数に対応するパラメータの方が低い周波数に対応するパラメータよりも周波数領域パラメータ[i]と変換後周波数領域パラメータ~[i]との差が小さくなるように、線形変換する。周波数領域パラメータ列[i]よりも変換後周波数領域パラメータ列~[i]のほうが、パラメータ値の間隔が均等間隔に近くなるようにする線形変換は、周波数領域においてパワースペクトル包絡の振幅の凹凸を鈍らせる処理(パワースペクトル包絡を平滑化する処理)に相当する。また、周波数領域パラメータ列[i]よりも変換後周波数領域パラメータ列~[i]のほうが、パラメータ値の間隔が均等間隔から遠くなるようにする線形変換は、周波数領域においてパワースペクトル包絡の振幅の凹凸を強調する処理(パワースペクトル包絡を逆平滑化する処理)に相当する。また、低い周波数に対応するパラメータの方が高い周波数に対応するパラメータよりも周波数領域パラメータ[i]と変換後周波数領域パラメータ~[i]との差が小さくなるようにする線形変換は、所定の時間区間の音信号に対応する周波数領域の係数列において高周波数領域ほど隣接するサンプル点間の周波数の間隔が狭い場合に相当する。また、高い周波数に対応するパラメータの方が低い周波数に対応するパラメータよりも周波数領域パラメータ[i]と変換後周波数領域パラメータ~[i]との差が小さくなるようにする線形変換は、所定の時間区間の音信号に対応する周波数領域の係数列において低周波数領域ほど隣接するサンプル点間の周波数の間隔が狭い場合に相当する。なお、上記の周波数領域パラメータ列[1],[2]...,[p-1]と変換後周波数領域パラメータ列~[1],~[2],...~,~[p]においては、低い周波数に対応するパラメータは*i*の値が小さいパラメータであり、高い周波数に対応するパラメータ*i*の値が大きいパラメータである。

【0251】

2. [i]が[i+1]と[i-1]との中点よりも[i+1]に近い場合には、~[i]が~[i+1]と~[i-1]との中点よりも~[i+1]に近く、かつ、[i+1]-[i]よりも~[i+1]-~[i]の方が値が小さく、かつ、iの値が小さいほど~[i]-[i]の絶対値が小さくなるように~[i]を求める。また、[i]が[i+1]と[i-1]との中点よりも[i-1]に近い場合には、~[i]が~[i+1]と~[i-1]との中点よりも~[i-1]に近く、かつ、[i]-[i-1]よりも~[i]-~[i-1]の方が値が小さく、かつ、iの値が小さいほど[i]-~[i]の絶対値

10

20

30

40

50

が小さくなるように $\sim [i]$ を求める。これは、所定の時間区間の音信号に対応する周波数領域の係数列において高周波数領域ほど隣接するサンプル点間の周波数の間隔が狭い場合の、周波数領域においてパワースペクトル包絡の振幅の凹凸を強調する処理（パワースペクトル包絡を逆平滑化する処理）に相当する。もしくは、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i+1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i+1]$ に近く、かつ、 $[i+1]-[i]$ よりも $\sim [i+1]-\sim [i]$ の方が値が小さく、かつ、 i の値が大きいほど $\sim [i]-[i]$ の絶対値が小さくなるように $\sim [i]$ を求める。また、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i-1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i-1]$ に近く、かつ、 $[i]-[i-1]$ よりも $\sim [i]-\sim [i-1]$ の方が値が小さく、かつ、 i の値が大きいほど $[i]-\sim [i]$ の絶対値が小さくように $\sim [i]$ を求める。これは、所定の時間区間の音信号に対応する周波数領域の係数列において低周波数領域ほど隣接するサンプル点間の周波数の間隔が狭い場合の、周波数領域においてパワースペクトル包絡の振幅の凹凸を強調する処理（パワースペクトル包絡を逆平滑化する処理）に相当する。

10

【 0 2 5 2 】

3. $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i+1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i+1]$ に近く、かつ、 $[i+1]-[i]$ よりも $\sim [i+1]-\sim [i]$ の方が値が大きく、かつ、 i の値が小さいほど $\sim [i]-[i]$ の絶対値が小さくなるように $\sim [i]$ を求める。また、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i-1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i-1]$ に近く、かつ、 $[i]-[i-1]$ よりも $\sim [i]-\sim [i-1]$ の方が値が大きく、かつ、 i の値が小さいほど $[i]-\sim [i]$ の絶対値が小さくなるように $\sim [i]$ を求める。これは、所定の時間区間の音信号に対応する周波数領域の係数列において高周波数領域ほど隣接するサンプル点間の周波数の間隔が狭い場合の、周波数領域においてパワースペクトル包絡の振幅の凹凸を鈍らせる処理（パワースペクトル包絡を平滑化する処理）に相当する。もしくは、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i+1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i+1]$ に近く、かつ、 $[i+1]-[i]$ よりも $\sim [i+1]-\sim [i]$ の方が値が大きく、かつ、 i の値が大きいほど $\sim [i]-[i]$ の絶対値が小さくなるように $\sim [i]$ を求める。また、 $[i]$ が $[i+1]$ と $[i-1]$ との中点よりも $[i-1]$ に近い場合には、 $\sim [i]$ が $\sim [i+1]$ と $\sim [i-1]$ との中点よりも $\sim [i-1]$ に近く、かつ、 $[i]-[i-1]$ よりも $\sim [i]-\sim [i-1]$ の方が値が大きく、かつ、 i の値が大きいほど $[i]-\sim [i]$ の絶対値が小さくなるように $\sim [i]$ を求める。これは、所定の時間区間の音信号に対応する周波数領域の係数列において低周波数領域ほど隣接するサンプル点間の周波数の間隔が狭い場合の、周波数領域においてパワースペクトル包絡の振幅の凹凸を鈍らせる処理（パワースペクトル包絡を平滑化する処理）に相当する。

20

30

【 0 2 5 3 】

例えば、周波数領域パラメータ列がLSPパラメータ列である場合、パラメータ列変換部20は、下記の式(16)により、変換後周波数領域パラメータ $\sim [1], \sim [2], \dots, \sim [p]$ を求めて出力する。

40

【 数 3 9 】

$$\begin{pmatrix} \tilde{\omega}[1] \\ \tilde{\omega}[2] \\ \vdots \\ \tilde{\omega}[p] \end{pmatrix} = K \begin{pmatrix} \omega[1] - \frac{\pi}{p+1} \\ \omega[2] - \frac{2\pi}{p+1} \\ \vdots \\ \omega[p] - \frac{p\pi}{p+1} \end{pmatrix} (\gamma_2 - \gamma_1) + \begin{pmatrix} \omega[1] \\ \omega[2] \\ \vdots \\ \omega[p] \end{pmatrix} \quad \dots(16)$$

【 0 2 5 4 】

50

ここで、 α_1 と α_2 は1以下の正の係数である。式(16)は、LSPパラメータをモデル化した式(10)において、 $\mathbf{a}_1 = ([1], [2], \dots, [p])^T$ とし、 $\mathbf{a}_2 = (\sim[1], \sim[2], \dots, \sim[p])^T$ とし、

【数40】

$$\Theta_{\gamma=0} = \left(\frac{\pi}{p+1}, \frac{2\pi}{p+1}, \dots, \frac{p\pi}{p+1} \right)$$

とすることで、導出することができる。この場合、周波数領域パラメータ $[1], [2], \dots, [p]$ は、線形予測係数 $a[1], a[2], \dots, a[p]$ の各係数 $a[i]$ に係数 α_1 の i 乗を乗じることにより補正した係数列である

$$a[1] \times (\alpha_1), a[2] \times (\alpha_1)^2, \dots, a[p] \times (\alpha_1)^p$$

と等価な周波数領域のパラメータ列、もしくは、その量子化値である。また、変換後周波数領域パラメータ $\sim[1], \sim[2], \dots, \sim[p]$ は、線形予測係数 $a[1], a[2], \dots, a[p]$ の各係数 $a[i]$ に係数 α_2 の i 乗を乗じることにより補正した係数列である

$$a[1] \times (\alpha_2), a[2] \times (\alpha_2)^2, \dots, a[p] \times (\alpha_2)^p$$

と等価な周波数領域のパラメータ列を近似する系列となる。

【0255】

なお、周波数領域パラメータ列がLSPパラメータ列である場合には、1次から $p+1$ 次に対応する $p+1$ 個のパラメータのうち1次から p 次までの p 個のパラメータが、LSPパラメータ列の1次から p 次までの p 個のパラメータと性質が一致するので、上記の説明を $p+1$ 次の場合のLSPパラメータ列の場合の説明に読み替えることができる。

【0256】

< 第四実施形態の効果 >

第四実施形態の周波数領域パラメータ変換装置は、第三実施形態の符号化装置や復号装置と同様に、第一実施形態や第二実施形態の符号化装置や復号装置のような周波数領域パラメータから線形予測係数を介して変換後周波数領域パラメータを求める場合よりも少ない演算量で、周波数領域パラメータから変換後周波数領域パラメータを求めることができる。

【0257】

[変形例等]

上記周波数領域パラメータ変換方法、符号化方法、復号方法、これらの装置において説明した処理は、記載の順にしたがって時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的あるいは個別に実行されてもよい。

【0258】

また、周波数領域パラメータ変換方法による各ステップをコンピュータによって実現する場合、周波数領域パラメータ変換方法の各ステップの処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、周波数領域パラメータ変換方法の各ステップがコンピュータ上で実現される。

【0259】

同様に、符号化方法による各ステップをコンピュータによって実現する場合、符号化方法の各ステップの処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、符号化方法の各ステップがコンピュータ上で実現される。

【0260】

同様に、復号方法による各ステップをコンピュータによって実現する場合、復号方法の各ステップの処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、復号方法の各ステップがコンピュータ上で実現される。

【0261】

この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気

10

20

30

40

50

記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよい。

【0262】

また、符号化方法及び復号方法の各ステップは、コンピュータ上で所定のプログラムを実行させることにより構成することにしてもよいし、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。

【0263】

その他、この発明の趣旨を逸脱しない範囲で適宜変更が可能であることはいうまでもない。

【符号の説明】

【0264】

1 1	周波数領域変換部	
1 2	線形予測分析部	
1 3	パワースペクトル包絡系列生成部	
1 4	包絡正規化部	
1 5	符号化部	
2 1	周波数領域変換部	
2 2	伸縮疑似パワースペクトル系列生成部	
2 3	線形予測分析部	
2 4	伸縮パワースペクトル包絡系列生成部	
2 5	逆伸縮変換部	20
2 6	包絡正規化部	
2 7	符号化部	
3 1	伸縮線形予測係数復号部	
3 2	伸縮パワースペクトル包絡系列生成部	
3 3	逆伸縮変換部	
3 4	復号部	
3 5	包絡逆正規化部	
3 6	時間領域変換部	
3 2 3	線形予測分析部	
3 2 3 1	線形予測係数算出部	30
3 2 3 2	L S P 計算部	
3 2 3 3	L S P 符号化部	
3 2 4	伸縮パワースペクトル包絡系列生成部	
3 2 8	伸縮対応平滑化 L S P 線形変換部	
4 3 1	伸縮線形予測係数復号部	
4 3 2	伸縮パワースペクトル包絡系列生成部	
4 3 8	伸縮対応平滑化 L S P 線形変換部	
1 0	周波数領域パラメータ変換装置	
2 0	パラメータ列変換部	

【 図 1 】

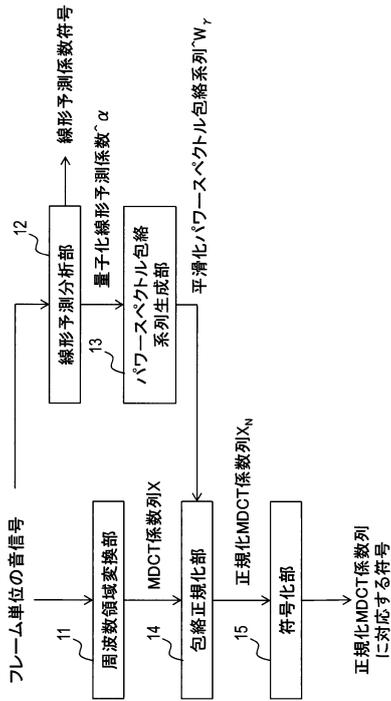


図1

【 図 2 】

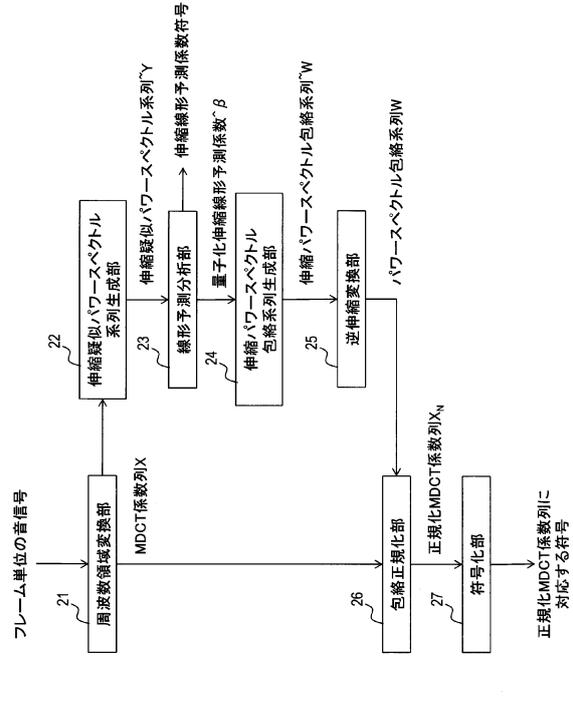


図2

【 図 3 】

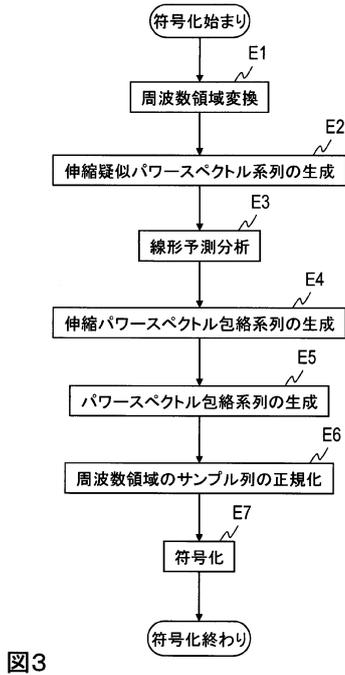


図3

【 図 4 】

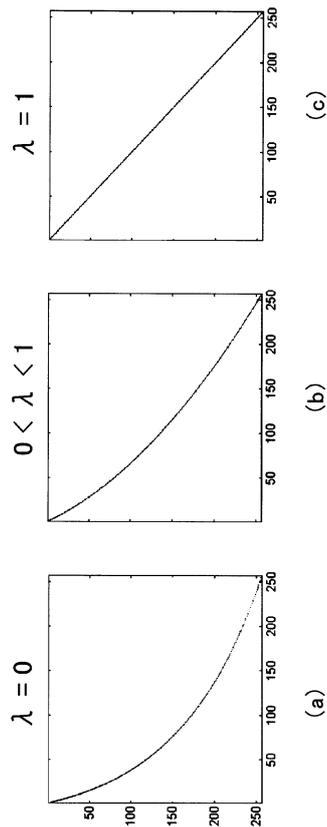


図4

【 図 9 】

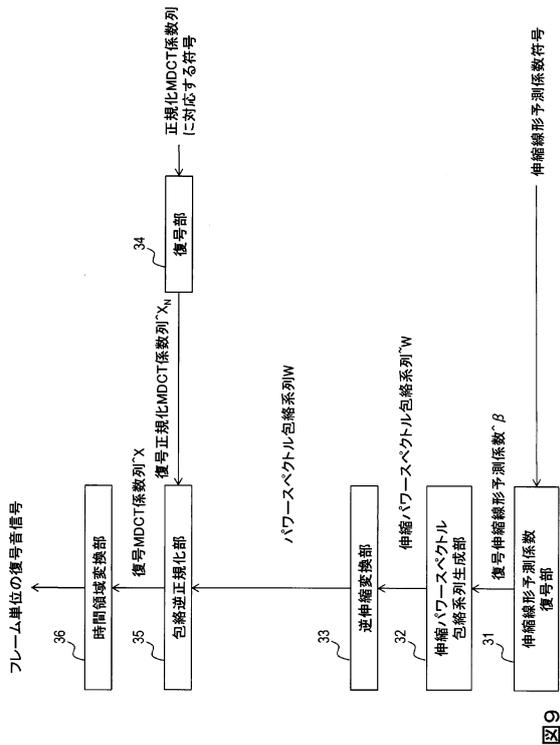


図 9

【 図 10 】

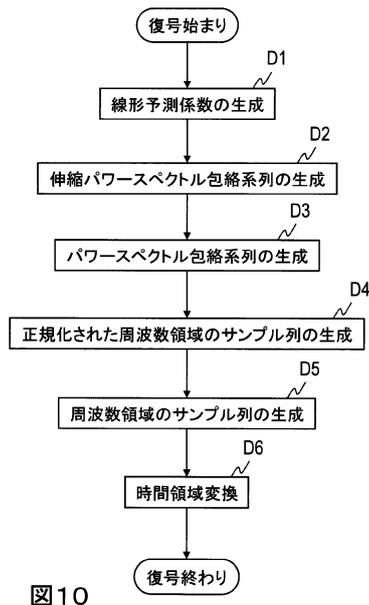


図 10

【 図 11 】

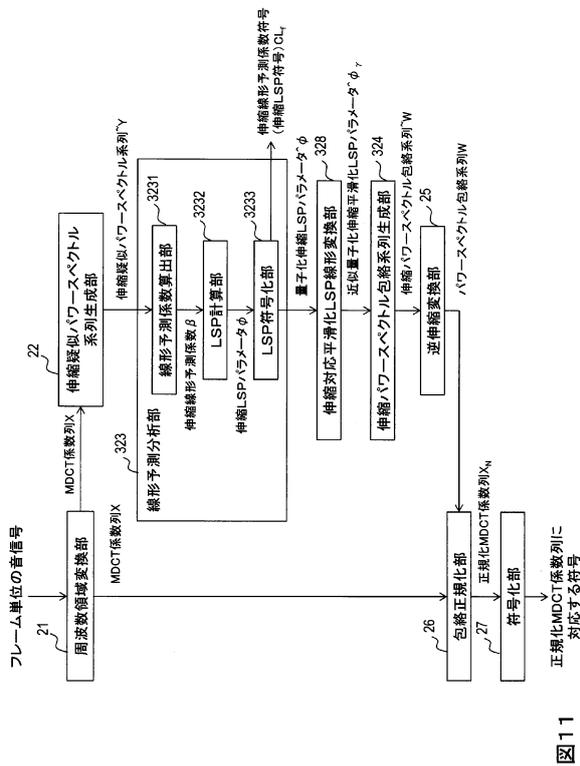


図 11

【 図 12 】

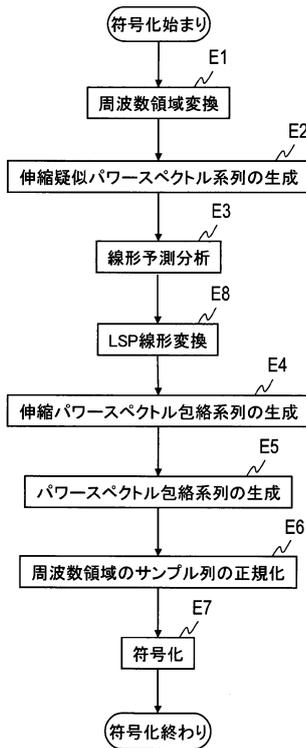


図 12

【図13】

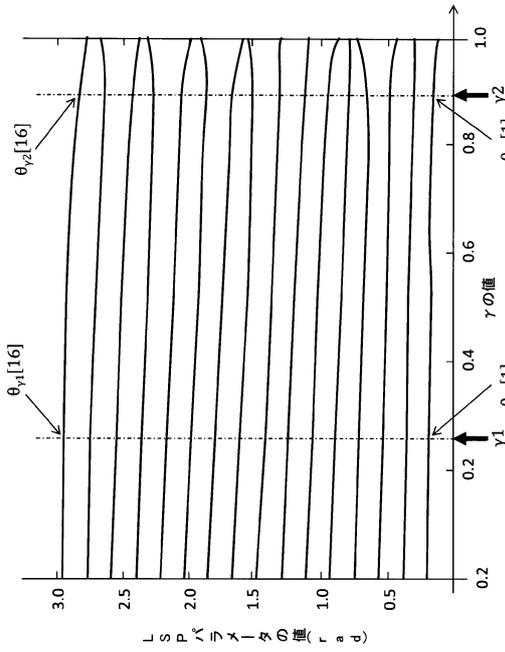


図13

【図14】

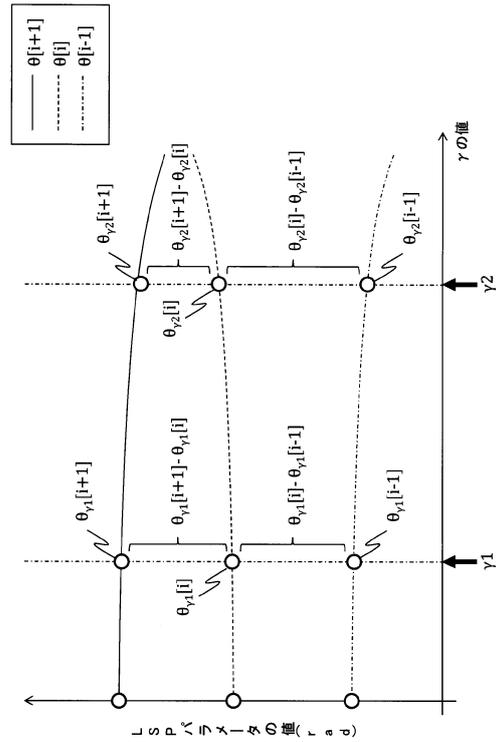


図14

【図15】

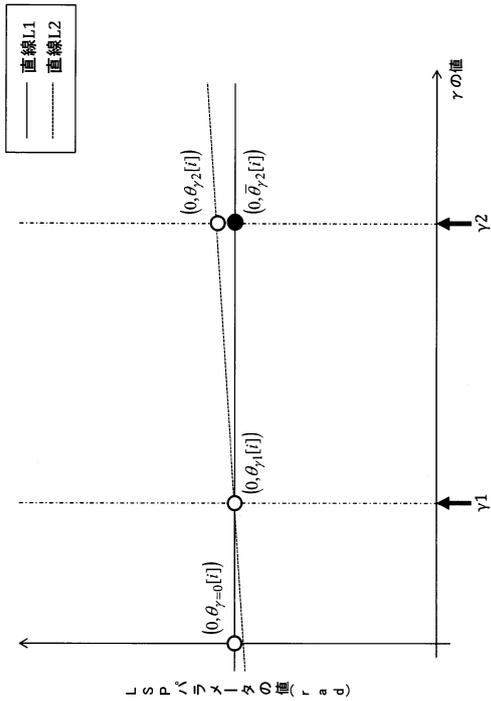


図15

【図16】

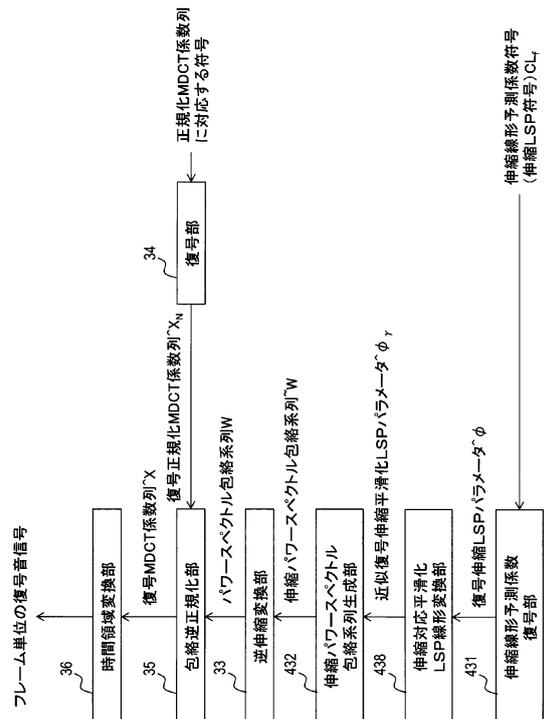


図16

【 図 17 】

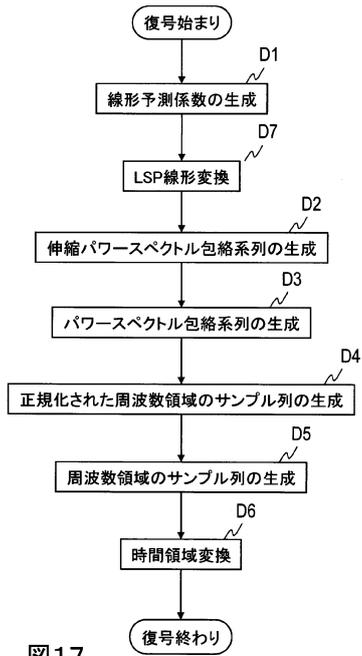


図 17

【 図 18 】

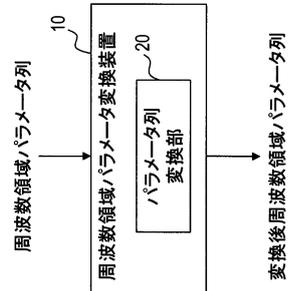


図 18

フロントページの続き

特許法第30条第2項適用 (1) 発行日 平成26年3月3日 刊行物名 日本音響学会 2014年春季研究発表会「講演論文集」講演要旨・講演論文CD-ROM 発表題目 「音響信号符号化のための周波数軸を伸縮するスペクトル包絡表現」(1-6-8) 開催日 平成26年3月10日~平成26年3月12日 集會名 「日本音響学会 2014年春季研究発表会」 開催場所 日本大学 公開者 杉浦亮介・鎌本優・原田登・亀岡弘和・守谷健弘 (2) 発行日 2014年5月17日 刊行物名 情報処理学会研究報告音楽情報科学研究会「音学シンポジウム2014」 発表題目 (39)「音響符号化のための周波数解像度の伸縮されたスペクトル包絡の表現法」 開催日 平成26年5月24日~平成26年5月25日 集會名 情報処理学会音楽情報科学研究会「音学シンポジウム2014」 開催場所 日本大学 公開者 杉浦亮介・鎌本優・原田登・亀岡弘和・守谷健弘

(72)発明者 鎌本 優

東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

(72)発明者 原田 登

東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

(72)発明者 亀岡 弘和

東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

(72)発明者 杉浦 亮介

東京都文京区本郷七丁目3番1号 国立大学法人東京大学内

審査官 鈴木 圭一郎

(56)参考文献 特開2006-235243(JP,A)

国際公開第2012/046685(WO,A1)

特開2009-251029(JP,A)

間野一則, 音声の高効率符号化, 信号処理, 日本, 信号処理研究会, 2008年 3月 4日, Vol.2 No.6

Kazuhito Koishida, A WIDEBAND CELP SPEECH CODER AT 16 KBIT/S BASED ON MEL_GENERALIZED CEPSTRAL ANALYSIS, Acoustics, Speech and Signal Processing, 1998. Proceedings on the 1998 IEEE International Conference on, 米国, IEEE, 1998年 3月15日, p161-164

(58)調査した分野(Int.Cl., DB名)

G10L19/00-19/26

G10L21/00-21/18