

(19) 日本国特許庁(JP)

(12) 公 開 特 許 公 報(A)

(11) 特許出願公開番号
特開2010-78812
(P2010-78812A)

(43) 公開日 平成22年4月8日(2010.4.8)

(51) Int.Cl.
G 1 O L 21/02 (2006.01)

F I
G 1 O L 21/02 3 O 1 A
G 1 O L 21/02 3 O 2 B

テーマコード (参考)

審査請求 未請求 請求項の数 8 O L (全 34 頁)

(21) 出願番号	特願2008-246015 (P2008-246015)	(71) 出願人	000005223
(22) 出願日	平成20年9月25日 (2008. 9. 25)		富士通株式会社
			神奈川県川崎市中原区上小田中4丁目1番1号
		(74) 代理人	100099759
			弁理士 青木 篤
		(74) 代理人	100119987
			弁理士 伊坪 公一
		(74) 代理人	100081330
			弁理士 樋口 外治
		(74) 代理人	100141254
			弁理士 榎原 正巳
		(74) 代理人	100113826
			弁理士 倉地 保幸
		最終頁に続く	

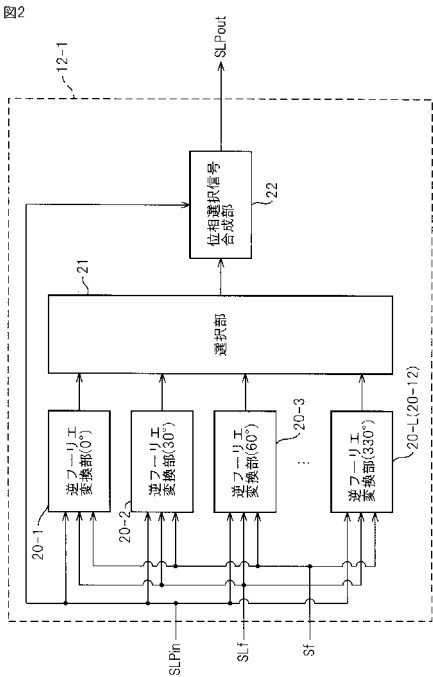
(54) 【発明の名称】 音声信号処理装置及び音声信号処理方法

(57) 【要約】

【課題】 ユーザの聴覚に感知される音声品質を劣化させることなく、ユーザに聞き易い信号になるように音声信号を処理する。

【解決手段】 音声信号処理装置1は、デジタル音声信号を所定長毎に分割した音声フレーム信号の周波数成分に、互いに異なる位相シフトを与えることにより得られる複数の異なる音声フレーム信号の、それぞれの最大振幅値を決定する最大振幅値決定手段(20-1~20-12、21、60-1~60-T、61)と、複数の異なる音声フレーム信号のうち最大振幅値が最も小さいものを選択する選択手段(21、61)を備える。

【選択図】 図2



【特許請求の範囲】**【請求項 1】**

ディジタル音声信号を所定長毎に分割した音声フレーム信号の周波数成分に、互いに異なる位相シフトを与えることにより得られる複数の異なる音声フレーム信号の、それぞれの最大振幅値を決定する最大振幅値決定手段と、

前記複数の異なる音声フレーム信号のうち最大振幅値が最も小さいものを選択する選択手段と、を備える音声信号処理装置。

【請求項 2】

前記最大振幅値決定手段は、

前記音声フレーム信号の各周波数成分を決定する周波数成分決定手段と、

10

前記各周波数成分にそれぞれ与えられる位相シフト量の複数の組み合わせを決定する組み合わせ決定手段と、を備え、

前記選択手段は、前記組み合わせ決定手段により決定された前記複数の組み合わせのうち、前記音声フレーム信号の最大振幅値を最も小さくする組み合わせを選択する請求項 1 に記載の音声信号処理装置。

【請求項 3】

前記組み合わせ決定手段は、前記周波数成分決定手段により決定されたいずれか 1 つの周波数成分に、互いに異なるシフト量の位相シフトを与えて複数の候補信号を生成する候補生成手段を備え、

前記選択手段は、前記複数の候補信号のうち最大振幅値が最も小さい候補信号に与えられたシフト量を選択する候補選択手段を備える、請求項 2 に記載の音声信号処理装置。

20

【請求項 4】

前記組み合わせ決定手段は、前記周波数成分決定手段により決定される全ての周波数成分の各々に対して、シフト量が異なる複数の位相シフトの各々を与えて複数の音声フレーム信号を生成する候補生成手段を備え、

前記選択手段は、前記複数の音声フレーム信号のうち最大振幅値が最も小さいものを選択する請求項 2 に記載の音声信号処理装置。

【請求項 5】

前記最大振幅値決定手段は、前記音声フレーム信号をフィルタリングする、異なる周波数一位相特性を有する複数のオールパスフィルタを備え、

30

前記選択手段は、前記複数のオールパスフィルタによりフィルタリングされた前記音声フレーム信号のうち、最大振幅値を最も小さいものを選択する請求項 1 に記載の音声信号処理装置。

【請求項 6】

ディジタル音声信号を所定長毎に分割した音声フレーム信号の周波数成分に位相シフトを与えることにより前記音声フレーム信号の最大振幅値を減少化させる最大値減少化手段と、

最大振幅値が減少化された後の音声フレーム信号の最大振幅値に応じて決定される信号増幅率で、前記最大振幅値が減少化された後の音声フレーム信号を増幅する信号増幅手段と、

40

を備える音声信号処理装置。

【請求項 7】

ディジタル音声信号を所定長の音声フレーム信号に分割し、

分割された前記音声フレーム信号の周波数成分に互いに異なる位相シフトを与えることにより得られる複数の異なる音声フレーム信号の、それぞれの最大振幅値を決定し、

前記複数の異なる音声フレーム信号のうち最大振幅値が最も小さいものを選択する、音声信号処理方法。

【請求項 8】

ディジタル音声信号を所定長の音声フレーム信号に分割し、

分割された前記音声フレーム信号の周波数成分に位相シフトを与えることにより前記音

50

声フレーム信号の最大振幅値を減少化させ、

最大振幅値が減少化された後の音声フレーム信号の最大振幅値に応じて決定される信号増幅率で、前記最大振幅値が減少化された後の音声フレーム信号を増幅する、

音声信号処理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、入力又は受信した音声信号を処理する音声信号処理装置及び音声信号処理方法に関する。

【背景技術】

【0002】

たとえば携帯電話による通話中のように、スピーカから出力された音声周囲の雑音のためによく聞こえない状況がよく生じうる。このような状況下で、出力された音声をユーザに聞き易くするためのいくつかの提案が考えられる。

【0003】

たとえば、出力される音声信号のスペクトル分析を行い、特定の重要な周波数成分、たとえばフォルマント周波数などの周波数成分を強調することが考えられる。また、出力される音声と背景雑音との S/N 比を算出し、ある値以上の S/N 比が得られるように音声信号のレベルを増幅することが考えられる。さらに、出力される音声信号の元信号のレベルに応じて適応的に音声信号の増幅率を制御するコンパンダ回路も考案されている。コンパンダ回路は、小さな元信号は大きく増幅し大きな元信号は小さく増幅することにより、増幅後の信号が増幅回路の許容最大出力レベルを超えないように信号を増幅する。

【0004】

なお、送話器、受話器、送話器から入る周囲雑音の周波数特性を解析する周波数解析手段、および周波数解析手段の解析結果に基づいて受話器に出力される受話音声の周波数特性を変換する周波数特性変換手段を備え、周波数解析手段は、周囲雑音の大きい高雑音周波数帯域を検出し、この解析結果に基づいて周波数特性変換手段は高雑音周波数帯域以外の受話音声帯域を強調する音声制御装置が開示されている。

また、送話器と受話器を備え無線信号による音声通話が可能なものであって、送話器から入る周囲雑音の周波数特性を解析する周波数解析手段、および音声通話時に周波数解析手段の解析結果に基づいて、無線信号による受話音声の周波数特性を変換する周波数特性変換手段をさらに備えた携帯電話が開示されている。

【0005】

【特許文献1】特開2002-223268号公報

【発明の開示】

【発明が解決しようとする課題】

【0006】

従来の方法では、周囲の雑音レベルが非常に大きい場合にはユーザの聞き易さを改善できる程度に限界がある。たとえば、出力音声と背景雑音との S/N 比を算出して所望の S/N 比を実現するように音声信号のレベルを増幅する従来の方法では、増幅後の出力音声レベルが増幅回路の最大許容値を超えると波形にクリッピング歪みが生じて音声品質が劣化する。また、コンパンダ回路を用いた方法でも波形に歪みが生じ音声品質が劣化する。

【0007】

このような従来の問題点に鑑み、開示する装置及び方法は、ユーザの聴覚により感知される音声品質を劣化させることなく、ユーザに聞き易い信号になるように、入力又は受信された音声信号を処理することを目的とする。

【課題を解決するための手段】

【0008】

ある実施例の形態による音声信号処理装置は、ディジタル音声信号を所定長毎に分割した音声フレーム信号の周波数成分に、互いに異なる位相シフトを与えることにより得られ

10

20

30

40

50

る複数の異なる音声フレーム信号の、それぞれの最大振幅値を決定する最大振幅値決定手段と、複数の異なる音声フレーム信号のうち最大振幅値が最も小さいものを選択する選択手段と、を備える。

【0009】

他の実施例の形態による音声信号処理装置は、ディジタル音声信号を所定長毎に分割した音声フレーム信号の周波数成分に位相シフトを与えることにより音声フレーム信号の最大振幅値を減少化させる最大値減少化手段と、最大振幅値が減少化された後の音声フレーム信号の最大振幅値に応じて決定される信号増幅率で、最大振幅値が減少化された後の音声フレーム信号を増幅する信号増幅手段と、を備える。

【発明の効果】

10

【0010】

開示の装置及び方法によれば、最大振幅値が減少するように音声信号が処理されるので、増幅段においてクリッピング歪みを発生させずに増幅可能な最大増幅率を増加させることができる。その結果、ユーザの聴覚により感知される音声品質を劣化させることなく、ユーザに聞き易い信号になるように、入力又は受信された音声信号を処理することが可能になる。

【発明を実施するための最良の形態】

【0011】

以下、添付する図面を参照して実施例を説明する。図1は、開示の音声処理装置の第1実施例の概略構成図である。音声処理装置1は、フレーム分割部2と、最大値減少化処理部3と、増幅率決定部4と、増幅部5と、フレーム記憶部6と、フレーム接続部7を備える。

20

【0012】

フレーム分割部2は、入力されたディジタル形式の音声信号を、所定長毎の音声フレーム信号へ分割する。

最大値減少化処理部3は、フレーム分割部2から順次出力される各音声フレーム信号の周波数成分に位相シフトを与えることにより音声フレーム信号の最大振幅値を減少化させる。

【0013】

増幅率決定部4は、最大値減少化処理部3によって最大振幅値が減少化された音声フレーム信号の最大振幅値に応じて、この音声フレーム信号を増幅すべき信号増幅率を決定する。増幅部5は、最大値減少化処理部3によって最大振幅値が減少化された音声フレーム信号を、増幅率決定部4により決定された信号増幅率で増幅する。

30

【0014】

フレーム記憶部6は、増幅部5によって増幅された音声フレーム信号の最後のサンプルから少なくともR個のサンプルを、次の音声フレーム信号が増幅部5から出力されるまで保持する。フレーム接続部7は、増幅部5から出力された音声フレーム信号と、この音声フレーム信号の前のフレームの音声フレーム信号を接続する。フレーム接続部7によるフレームの接続処理については後述する。

【0015】

40

最大値減少化処理部3は、フーリエ変換部10と、周波数選択部11と、直列接続されたM段の位相選択部12-1、12-2、…12-Mと、逆フーリエ変換部13を備える。フーリエ変換部10は、フレーム分割部2から順次供給される音声フレーム信号をフーリエ変換して、音声フレーム信号の各周波数成分を示す周波数領域信号を生成する。この周波数領域信号は、周波数選択部11、位相選択部12-1～12-M及び逆フーリエ変換部13へ出力される。位相選択部12-1～12-Mは、周波数領域信号を入力Sfとして入力する。

【0016】

周波数選択部11は、フーリエ変換部10から入力した各周波数成分のスペクトル強度にしたがって、スペクトル強度が最も強い周波数を指示する信号、2番目に強い周波数を

50

指示する信号、…M番目に強い周波数を指示する信号を出力する。これらスペクトル強度が最も強い周波数を指示する信号、2番目に強い周波数を指示する信号、…M番目に強い周波数を指示する信号は、入力SLfとして位相選択部12-1、12-2、…12-Mへそれぞれ入力される。

【0017】

位相選択部12-1～12-Mは、入力Sfとして与えられた各周波数成分のうち、入力SLfによって指定された周波数fの周波数成分へ、複数の異なるシフト量の位相シフトを与えて時間領域信号に逆フーリエ変換したとき、音声フレーム信号の最大振幅値が最も小さくなる位相シフト量を、周波数fの周波数成分に与えるべき位相シフト量として選択する。

10

【0018】

位相選択部12-1～12-Mは、各々選択した位相シフト量を示す位相選択信号を、出力SLPoutとして出力する。前段の位相選択部12-1～12-(M-1)から出力される位相選択信号は、入力SLPinとして後段の位相選択部12-2～12-Mへ入力される。

【0019】

周波数f_iの周波数成分に与える位相シフト量を選択した前段の位相選択部12-iから位相選択信号を入力した後段の位相選択部12-(i+1)は(i=1～M-1)、入力SLfによって指定された周波数f(i+1)の周波数成分に与えるべき位相シフト量を選択し、選択された位相シフト量を、前段の位相選択部12-iから入力した位相選択信号に追加してから、さらに後段の位相選択部12-(i+2)へ出力する。

20

【0020】

また、入力SLfによって指定された周波数f_iの周波数成分に与えるべき位相シフト量を選択するとき(i=2～M)、各位相選択部12-iは、周波数f_i以外の他の各周波数成分には前段の位相選択部12-(i-1)から入力した位相選択信号によって指定される各シフト量の位相シフトをそれぞれ与える。

【0021】

すなわち、各位相選択部12-i(i=2～M)は、周波数f_i以外の他の各周波数成分には前段の位相選択部12-(i-1)から入力した位相選択信号によって指定される各シフト量の位相シフトをそれぞれ与え、かつ周波数f_iの周波数成分には複数の異なるシフト量の位相シフトΔθ₁～Δθ_Lを与えて時間領域信号に逆フーリエ変換したときに、音声フレーム信号の最大振幅値が最も小さくなるような位相シフト量を、位相シフト量Δθ₁～Δθ_Lの中から選択する。1段目の位相選択部12-1の入力SPLinには、全ての周波数成分について位相シフト量を指定しない位相選択信号が入力される。

30

【0022】

最終段の位相選択部12-Mの出力SPLoutからは、各段の位相選択部12-1～12-Mによってそれぞれ選択された、スペクトル強度が最も強い周波数、2番目に強い周波数、…M番目に強い周波数へそれぞれ与える各位相シフト量を指示する位相選択信号が結合されて出力され、逆フーリエ変換部13に出力される。

【0023】

逆フーリエ変換部13は、フーリエ変換部10から与えられる周波数領域信号の各周波数成分に、位相選択部12-Mから与えられる位相選択信号によって指定される各位相シフトをそれぞれ与えて、周波数領域信号を逆フーリエ変換した音声フレーム信号を生成する。逆フーリエ変換部13は、音声フレーム信号を増幅率決定部4及び増幅部5へ出力する。

40

【0024】

図2は、図1に示す位相選択部12-1の構成例を示す図である。他の位相選択部12-2～12-Mも同様の構成を有する。位相選択部12-1は、L個の逆フーリエ変換部20-1～20-Lと、選択部21と、位相選択信号合成部22とを備える。

【0025】

50

L 個の逆フーリエ変換部 $20-j$ ($j = 1, 2, \dots, L$) は、入力 S_f である周波数領域信号に含まれる各周波数成分のうち、入力 SL_f により指定された周波数 f の周波数成分には、シフト量 ($360 / L \times (j - 1)$) 度の位相シフトを与え、それ以外の各周波数成分には、入力 SL_{Pin} である位相選択信号によって指定される各シフト量の位相シフトを与えてから、逆フーリエ変換を行って音声フレーム信号を生成する。

【0026】

本実施例では自然数 $L = 12$ の場合の構成例である。すなわち位相選択部 $12-1$ は、 12 個の逆フーリエ変換部 $20-1 \sim 20-12$ を備える。そして、逆フーリエ変換部 $20-1$ は入力 SL_f により指定された周波数 f の周波数成分に 0 度の位相シフトを与え、逆フーリエ変換部 $20-2$ は周波数 f の周波数成分に 30 度の位相シフトを与え、逆フーリエ変換部 $20-3$ は周波数 f の周波数成分に 60 度の位相シフトを与え、逆フーリエ変換部 $20-12$ は周波数 f の周波数成分に 330 度の位相シフトを与える。自然数 L は他の 2 以上の自然数を使用してもよい。

【0027】

選択部 21 は、逆フーリエ変換部 $20-1 \sim 20-12$ により生成された各音声フレーム信号のうち、最大振幅値が最小である音声フレーム信号を選択する。選択部 21 は、選択された音声フレーム信号の周波数 f の周波数成分に与えられた位相シフト量を示す位相選択信号を出力する。

【0028】

位相選択信号合成部 22 は、入力 SL_{Pin} である位相選択信号へ、選択部 21 が出力した位相選択信号を、周波数 f の周波数成分に与えるべき位相シフト量として挿入することによって、入力 SL_{Pin} として入力した位相選択信号と、選択部 21 が出力した位相選択信号とを合成する。位相選択信号合成部 22 は、合成した位相選択信号を出力 SL_{Pout} として出力する。

【0029】

なお、逆フーリエ変換部 $20-1 \sim 20-12$ 及び選択部 21 は、特許請求の範囲に記載される最大振幅値決定手段に相当する。また、選択部 21 は、特許請求の範囲に記載される選択手段に相当する。

【0030】

フーリエ変換部 10 は、特許請求の範囲に記載される周波数成分決定手段に相当する。各段の位相選択部 $12-1 \sim 12-M$ の逆フーリエ変換部 $20-1 \sim 20-12$ は、特許請求の範囲に記載される組み合わせ決定手段に相当する。

【0031】

逆フーリエ変換部 $20-1 \sim 20-12$ は、特許請求の範囲に記載される候補生成手段に相当し、逆フーリエ変換部 $20-1 \sim 20-12$ から出力される音声フレーム信号は、特許請求の範囲に記載される候補信号に相当する。選択部 21 は特許請求の範囲に記載される候補選択手段に相当する。

【0032】

図 3 は、開示の音声処理方法の実施例の全体フローチャートである。ステップ $S1$ において、図 1 に示すフレーム分割部 2 は、入力されたデジタル形式の音声信号を、所定長毎の音声フレーム信号へ分割する。ステップ $S2$ において最大値減少化処理部 3 は、音声フレーム信号の最大振幅値を減少化させる。

【0033】

図 4 は、図 1 に示す最大値減少化処理部 3 による音声信号の最大値の減少化処理の第 1 例を示すフローチャートである。ステップ $S10$ において、図 1 に示すフーリエ変換部 10 は、音声フレーム信号をフーリエ変換して、音声フレーム信号の各周波数成分を示す周波数領域信号を生成する。

【0034】

ステップ $S11$ において周波数選択部 11 は、フーリエ変換部 10 から入力した周波数領域信号によって示される各周波数成分のスペクトル強度にしたがって、第 1 番目～第 M

10

20

30

40

50

番目に強いスペクトル強度を有する周波数 f_i ($i = 1 \sim M$) を決定する。周波数選択部 11 は、第 1 番目～第 M 番目に強いスペクトル強度を有する周波数 $f_i \sim f_M$ をそれぞれ指示する信号を、入力 SLf として、位相選択部 12-1、12-2、 $\dots$ 12-M へそれぞれ入力する。

【0035】

ステップ S12 において各位相選択部 12- i ($i = 1 \sim M$) を参照するインデックス変数 i の値を「1」に初期化する。

【0036】

ステップ S13 において、 i 段目の位相選択部 12- i は、第 i 番目にスペクトル強度が強い周波数 f_i を指示する信号を入力 SLf として受信する。

10

【0037】

位相選択部 12- i の逆フーリエ変換部 20- j ($j = 1 \sim 12$) は、フーリエ変換部 10 から与えられた各周波数成分のうち、入力 SLf によって指定された周波数 f_i 以外の他の各周波数成分には、前段の位相選択部 12- $(i-1)$ から入力した位相選択信号によって指定される各シフト量の位相シフトをそれぞれ与え、かつ周波数 f_i の周波数成分にはそれぞれ $(360/L \times (j-1))$ 度の位相シフトを与えて時間領域信号に逆フーリエ変換する。

【0038】

ステップ S14 において、位相選択部 12- i の選択部 21 は、逆フーリエ変換部 20-1～20-12 により生成された各音声フレーム信号のうち、最大振幅値が最小である音声フレーム信号を選択する。選択部 21 は、逆フーリエ変換部 20-1～20-12 が生成した音声フレーム信号のうちの、選択された音声フレーム信号の周波数成分 f_i に与えられた位相シフト量を示す位相選択信号を出力する。位相選択信号合成部 22 は、入力 $SLPin$ として入力した位相選択信号と、選択部 21 が出力した位相選択信号とを合成する。位相選択信号合成部 22 は、合成した位相選択信号を出力 $SLPout$ として出力する。

20

【0039】

ステップ S15 においてインデックス変数 i の値を 1 つ増加する。ステップ S16 において、インデックス変数 i の値が「 M 」以下であるとき、すなわち、まだ位相選択処理が済んでいない位相選択部の段が残っている場合には、処理はステップ S13 へ戻り、ステップ S13～S16 が反復される。

30

【0040】

ステップ S16 の判定において、インデックス変数 i の値が「 M 」以下でないとき、処理はステップ S17 へ移る。ステップ S17 において、図 1 に示す逆フーリエ変換部 13 は、フーリエ変換部 10 から与えられる各周波数成分に、最終段の位相選択部 12-M から与えられる位相選択信号によって指定される各位相シフトをそれぞれ与えて、周波数領域信号を逆フーリエ変換した音声フレーム信号を生成する。

【0041】

図 5 の (A) 及び図 5 の (B) は、最大値減少化処理部 3 による低減化処理の前後の音声フレーム信号の波形の模式図である。最大値減少化処理部 3 の各段の位相選択部 12-1～12-M の逆フーリエ変換部 20-1～20-12 によってそれぞれ生成される音声フレーム信号の波形は、音声フレーム信号の周波数成分に位相シフトが加えられているために、元の音声フレーム信号の波形と異なる波形になる。

40

【0042】

位相選択部 12-1～12-M の選択部 21 は、これら異なる波形の音声フレーム信号のうち最大振幅値が最も小さい音声フレーム信号を選択する。したがって選択部 21 によって選択される音声フレーム信号の最大振幅値は、元の音声フレームの最大振幅値以下となる。たとえば、元の音声フレーム信号の最大振幅値が複数の周波数成分の比較的振幅が大きい部分の重なり合いにより生じている場合には、各周波数成分に異なる位相シフトを与えることにより最大振幅値を減少できる。

【0043】

50

このため、最大値減少化処理部 3 による低減化処理後の音声フレーム信号、すなわち図 5 の (B) に示される音声フレーム信号の最大振幅値 S_{max2} は、図 5 の (A) に示す元の音声フレーム信号の最大振幅値 S_{max1} よりも小さくなる。

【0044】

ここで人間の聴覚には、各周波数成分における位相特性がある程度ずれても殆ど感知できないという性質がある。したがって、最大値減少化処理部 3 は、人間の聴覚に感知される音声品質を劣化させることなく、音声フレーム信号の最大振幅値を減少化することができる。

【0045】

図 3 に示すステップ S 3 において増幅率決定部 4 は、最大値減少化処理部 3 から出力された音声フレーム信号の最大振幅値に応じて、この音声フレーム信号を増幅すべき信号増幅率 A を決定する。ステップ S 4 において増幅部 5 は、最大値減少化処理部 3 から出力された音声フレーム信号を、増幅率決定部 4 により決定された信号増幅率 A で増幅する。

【0046】

図 6 は、図 1 に示す増幅率決定部 4 による信号増幅率 A の決定処理の例を説明する説明図である。図 6 に示す波形は、最大値減少化処理部 3 から出力された音声フレーム信号の信号波形である。たとえば増幅率決定部 4 は、後段の増幅部 5 による増幅後の音声フレーム信号が増幅部 5 の許容最大出力振幅値 S_{th} を超えない最大の増幅率を、信号増幅率 A として決定してよい。

【0047】

たとえば増幅率決定部 4 は、最大値減少化処理部 3 から出力された音声フレーム信号の最大振幅値が S_{max} であるとき、 $A = S_{th} / S_{max}$ を信号増幅率 A として決定してよい。増幅率決定部 4 がこのように信号増幅率 A を決定することにより、音声フレーム信号は、増幅部 5 においてクリッピング歪みを生じずに増幅される。

【0048】

このように増幅率決定部 4 及び増幅部 5 は、増幅前の最大振幅値が小さいほど、音声フレーム信号をより大きな信号増幅率で増幅できる。本実施例では、最大値減少化処理部 3 によって音声フレーム信号の最大振幅値が減少化されるので、より大きな増幅率で音声信号を増幅することができ、人間の聴覚に感知される音声品質を劣化させることなく背景雑音が大きい環境におけるユーザの聞き易さを改善することができる。

【0049】

ステップ S 5 においてフレーム接続部 7 は、増幅部 5 から出力された音声フレーム信号と、この音声フレーム信号の前のフレームの音声フレーム信号を接続する。

【0050】

最大値減少化処理部 3 による音声信号処理を行う前は、連続する 2 つのフレームのうち前フレームの最後のサンプル値と後フレームの最初のサンプル値の値はほぼ同じである。

【0051】

しかしながら、最大値減少化処理部 3 によって各周波数成分に位相シフトが与えられると、各音声フレーム信号毎に波形が変化し、その結果、連続する 2 つのフレームの前フレームの最後のサンプル値と、後フレームの最初のサンプル値の間のギャップが大きくなる可能性がある。

【0052】

フレーム接続部 7 は、前フレームの最後のサンプル値 S_b と後フレームの最初のサンプル値 S_a の間に目標値を定め、前フレームの最後の R 個のサンプルと後のフレームの S 個のサンプルとを目標値に向かって漸近させることにより、これら 2 つのフレームをスムーズに接続する接続処理を実行する。図 7 は、図 1 に示すフレーム接続部 7 による音声フレーム信号の接続処理の例を示すフローチャートである。

【0053】

ステップ S 20 においてフレーム接続部 7 は、前フレームの最後のサンプルの値 S_b の符号と、後フレームの最初のサンプルの値 S_a の符号とが異なるか否かを判定する。 S_b

10

20

30

40

50

の符号と S a の符号が同じ場合には、フレーム接続部 7 は処理をステップ S 2 2 に移す。

【0054】

S b の符号と S a の符号が異なる場合には、ステップ S 2 1 においてフレーム接続部 7 は後フレームの各サンプルの符号を反転させる。これにより、S b と S a の値を近づけることができ、よりスムーズに前フレームと後フレームを接続できるようになる。

【0055】

ステップ S 2 2 においてフレーム接続部 7 は、前フレームの最後のサンプル値 S b と後フレームの最初のサンプル値 S a の間に目標値 S m を定める。目標値 S m は、たとえば S b と S a の中間値でよい。図 8 の (A) に、前フレームの最後の R 個のサンプル時刻 S b (P - R + 1)、~ S b (P - 2)、S b (P - 1)、S b (P) におけるサンプルと、
10
後フレームの S 個のサンプル時刻 S a (1)、S a (2)、S a (3)、~ S a (S) におけるサンプルと、目標値 S m を示す。

【0056】

ステップ S 2 3 においてフレーム接続部 7 は、前フレームの最後の R 個のサンプルを目標値 S m に向かって漸近させる。具体的には、前フレームの最後の R 個のサンプル時刻 S b (P - R + j) のサンプルの値を、それぞれ $(1 + (S m / S b - 1) \times j / R)$ 倍する ($j = 1 \sim R$)。この乗算処理によって、前フレームの最後の R 個のサンプルには、フレームの最後に近づくにつれて値 $1 \sim S m / S b$ へと変化する係数が乗算され、これらサンプルの値は目標値 S m へ徐々に近づく。図 8 の (B) には、ステップ S 2 3 に示す乗算
20
処理が施された前フレームが示されている。

【0057】

ステップ S 2 4 においてフレーム接続部 7 は、後フレームの最初の S 個のサンプルを目標値 S m に向かって漸近させる。具体的には、後フレームの最初の S 個のサンプル時刻 S a (j) のサンプルの値を、それぞれ $(S m / S a + (1 - S m / S a) \times (j - 1) / S)$ 倍する ($j = 1 \sim S$)。この乗算処理によって、後フレームの最後の S 個のサンプルには、フレームの始めに近づくにつれて値 $1 \sim S m / S b$ へと変化する係数が乗算され、これらサンプルの値は目標値 S m へ徐々に近づく。図 8 の (B) には、ステップ S 2 3 に示す乗算処理が施された後フレームが示されている。

【0058】

図 9 は、開示の音声処理装置の第 2 実施例の概略構成図である。図 9 に示す音声処理装置 1 は、図 1 に示す構成と類似する構成を有しており、図 1 に示す構成要素と同様の構成要素には同じ参照符号を使用し、また同一の機能については説明を省略する。
30

【0059】

本構成例の音声処理装置 1 は、音声フレーム信号の信号増幅率を決定する際の目標値である目標増幅率 A t を決定する目標増幅率決定部 8 を備える。目標増幅率決定部 8 は、たとえば前フレームの音声フレーム信号を増幅する際に増幅率決定部 4 が決定した信号増幅率を目標増幅率 A t としてよい。または目標増幅率決定部 8 は、たとえば音声処理装置 1 が作動を開始した際の初めのフレームの音声フレーム信号を増幅する際に増幅率決定部 4 が決定した信号増幅率を目標増幅率 A t としてよい。

【0060】

最大値減少化処理部 3 は、図 2 を参照して説明した位相選択部 1 2 - 1 と同様の位相選択部を直列に (M - 1) 段接続した位相選択部 1 2 - 1、1 2 - 2、... 1 2 - (M - 1) と、位相選択部 1 2 - (M - 1) の後段に接続される最終段の位相選択部 1 4 を備える。
40

【0061】

図 10 は、図 9 に示す位相選択部 1 4 の構成例を示す図である。位相選択部 1 4 には、フーリエ変換部 10 から出力される周波数領域信号が入力 S f として、周波数選択部 11 から出力される M 番目に強い周波数を指示する信号が入力 S L f として、位相選択部 1 2 - (M - 1) から出力 S L P out として出力される位相選択信号が入力 S L P in として入力される。

【0062】

位相選択部 1 4 は、図 2 を参照して説明した逆フーリエ変換部 20 - 1 ~ 20 - L と同
50

様に動作する逆フーリエ変換部 30-1 ~ 30-L と、図 2 を参照して説明した位相選択信号合成部 22 と同様に動作する位相選択信号合成部 32 と、選択部 31 を備える。本実施例では自然数 $L = 12$ の場合の構成例である。自然数 L は他の 2 以上の自然数を使用してもよい。

【0063】

また位相選択部 14 には、目標増幅率決定部 8 が決定した目標増幅率 A_t と、フレーム記憶部 6 に記憶された前フレームの最後のサンプル値 S_b が入力される。選択部 31 は、逆フーリエ変換部 30-1 ~ 30-12 が生成する音声フレーム信号を入力する。

【0064】

選択部 31 は、逆フーリエ変換部 30-1 ~ 30-12 により生成された各音声フレーム信号の最大振幅値に基づいて、これら各音声フレーム信号に与えられた位相シフトのうち、所定の選択要件を満足する位相シフト量があるか否かを判定する。

【0065】

ここで、ある位相シフト量が選択されるための所定の選択要件とは、音声フレーム信号に対して、入力 SLf によって指定された周波数 f の周波数成分にその位相シフト量が与えられ、周波数 f 以外の他の各周波数成分には前段までの位相選択部によって指定された各シフト量の位相シフトをそれぞれ与えられたとき、下記の条件 (1) ~ (3) が満たす信号増幅率 A が存在することである。

【0066】

(1) 信号増幅率 A が、目標増幅率 A_t から所定の許容範囲内に存在する。所定の許容範囲とは $A_t \times (1 - b\%) \sim A_t \times (1 + b\%)$ である。ここに b は所定の定数である。

(2) 増幅部 5 が、信号波形にクリッピング歪みを生じることなく音声フレーム信号を信号増幅率 A で増幅できる。

(3) 信号増幅率 A で音声フレーム信号を増幅したときに、音声フレーム信号の最初のサンプル値 S_a が前フレームの最初のサンプル値 S_b から所定の許容範囲内に収まる。所定の許容範囲とは、 $S_b \times (1 - Q\%) \sim S_b \times (1 + Q\%)$ である。ここに Q は所定の定数である。

【0067】

選択部 31 は、所定の選択要件を満足する位相シフト量が与えられた音声フレーム信号のうち、最小の最大振幅値を有する音声フレーム信号に与えられた位相シフト量を選択する。選択部 31 は、選択された音声フレーム信号に与えられた位相シフト量を示す位相選択信号を位相選択信号合成部 32 へ出力する。

【0068】

選択部 31 がこのような位相シフト量を選択することによって、現在処理中の音声フレーム信号に与えられる信号増幅率と、前フレームに与えられた信号増幅率との差を所定の範囲内に収めることができる。このためユーザが音量の変化を感知しにくくなる。

【0069】

またこのような位相シフト量が選択されることによって、現在処理中の音声フレーム信号の最初のサンプル値 S_a と、前フレームの最後のサンプル値 S_b との差を所定の範囲内に収めることができる。このためユーザがフレームの間のつなぎ目を感知しにくくなる。

【0070】

位相選択信号合成部 32 は、入力 $SLPin$ として入力した位相選択信号と、選択部 31 が出力した位相選択信号とを合成し、合成した位相選択信号を出力 $SLPout$ として逆フーリエ変換部 13 へ出力する。

【0071】

図 11 は、図 9 に示す最大値減少化処理部 3 により実行される音声信号の最大値の減少化処理の第 2 例を示すフローチャートである。

【0072】

ステップ S30 ~ ステップ S36 では、図 4 に示したステップ S10 ~ S16 にて第 1

10

20

30

40

50

番目～第（ $M-1$ ）番目の各周波数の周波数成分に与える位相が選択されたのと同様に、第1番目～第（ $M-1$ ）番目の各周波数の周波数成分に与える位相が選択される。

【0073】

ステップS37において、M段目の位相選択部14は、第M番目のスペクトル強度を有する周波数 f_M を指示する信号を入力 SLf として受信する。

【0074】

位相選択部14の逆フーリエ変換部30-j（ $j=1\sim 12$ ）は、フーリエ変換部10から与えられた各周波数成分のうち、周波数 f_M 以外の他の各周波数成分には、前段の位相選択部12-（ $M-1$ ）から入力した位相選択信号によって指定される各シフト量の位相シフトをそれぞれ与え、かつ周波数 f_M の周波数成分にはそれぞれ（ $360/L \times (j-1)$ ）度の位相シフトを与えて時間領域信号に逆フーリエ変換する。

10

【0075】

ステップS38において、位相選択部14の選択部31は、逆フーリエ変換部30-1～30-12により生成された各音声フレーム信号に与えられた位相シフトの中に、上述した所定の選択要件を満たす位相シフトがあるか否かを判定する。

【0076】

図12は、ある位相シフトが所定の選択要件を満足するか否かを判定する判定処理のフローチャートである。ステップS50において選択部31は、前フレームの最後のサンプルの値 S_b の符号と、位相シフトを与えた現在のフレームの最初のサンプルの値 S_a' の符号とが異なるか否かを判定する。 S_b の符号と S_a' の符号が同じ場合には、選択部31は処理をステップS52に移す。

20

【0077】

S_b の符号と S_a' の符号が異なる場合には、ステップS51においてフレーム接続部7は現在のフレームの各サンプルの符号を反転させる。これにより S_b と S_a' の値の差が小さくなる。

【0078】

ステップS52において選択部31は、既知の増幅部5の許容最大出力振幅値 S_{th} と音声フレーム信号の最大振幅値 S_{max} とに基づいて、音声フレーム信号の最大振幅値 S_{max} が、所定値（ $S_{th}/(A_t \times (1-b\%))$ ）より大きいかなんかを判定する。この判定によって、選択部31は、増幅後の音声フレーム信号にクリッピング歪みを生じない最大増幅率（ S_{th}/S_{max} ）が、所定の許容範囲の下限（ $A_t \times (1-b\%)$ ）よりも小さいかなんかを判定する。

30

【0079】

$S_{max} > (S_{th}/(A_t \times (1-b\%)))$ であるとき選択部31は、処理をS53へ移行する。 $S_{max} > (S_{th}/(A_t \times (1-b\%)))$ でないとき選択部31は、処理をS54へ移行する。ステップS53において選択部31は、位相シフトが所定の選択要件を満たさないと決定して判定処理を終了する。

【0080】

ステップS54において選択部31は、 $S_{max} \leq (S_{th}/(A_t \times (1+b\%)))$ であるか否かを判定することにより、増幅後の音声フレーム信号にクリッピング歪みを生じない最大増幅率（ S_{th}/S_{max} ）が、所定の許容範囲の上限（ $A_t \times (1+b\%)$ ）以上であるか否かを判定する。

40

【0081】

$S_{max} \leq (S_{th}/(A_t \times (1+b\%)))$ である場合には、選択部31は処理をステップS55に移す。ステップS55において選択部31は、増幅部5にて使用可能な信号増幅率の上限値 A_{max} を（ $A_t \times (1+b\%)$ ）に定め、下限値 A_{min} を（ $A_t \times (1-b\%)$ ）に定める。その後、選択部31は処理をステップS57に移す。

【0082】

ステップS54の判定において $S_{max} \leq (S_{th}/(A_t \times (1+b\%)))$ でない場合には、選択部31は処理をステップS56に移す。ステップS56において選択部31

50

は、上限値 A_{max} を最大増幅率 (S_{th}/S_{max}) に定め、下限値 A_{min} を ($A_t \times (1 - b\%)$) に定める。その後、選択部 31 は処理をステップ S57 に移す。

【0083】

ステップ S57 において選択部 31 は、ステップ S55 又は S56 において下限値及び上限値が定められた範囲の信号増幅率 $A_{min} \sim A_{max}$ によって現在の音声フレーム信号が増幅されたときの最初のサンプル値の範囲を決定する。増幅前の現在の音声フレーム信号の最初のサンプル値を $S_{a'}$ とすると、増幅後の現在の音声フレーム信号の最初のサンプル値の範囲は $S_{a'} \times A_{min} \sim S_{a'} \times A_{max}$ である。

【0084】

選択部 31 は、増幅後の現在の音声フレーム信号の最初のサンプル値 S_a に許される所定の許容範囲 $S_b \times (1 - Q\%) \sim S_b \times (1 + Q\%)$ と、 $S_{a'} \times A_{min} \sim S_{a'} \times A_{max}$ とが重複しないか否かを判定する。これらの範囲が重複しないとき上記の所定の選択要件 (3) を満たす信号増幅率が存在しないため、選択部 31 は、処理をステップ S53 へ移し、位相シフトが所定の選択要件を満たさないと決定して判定処理を終了する。

【0085】

図 13 の (A) 及び図 13 の (B) は、範囲 $S_b \times (1 - Q\%) \sim S_b \times (1 + Q\%)$ と範囲 $S_{a'} \times A_{min} \sim S_{a'} \times A_{max}$ が重複部分 R を有する 2 つの態様を示し、図 13 の (C) 及び図 13 の (D) は、範囲 $S_{a'} \times A_{min} \sim S_{a'} \times A_{max}$ に重複部分がない 2 つの態様を示す。これらの図から明らかなように、($S_{a'} \times A_{min} > S_b \times (1 + Q\%)$) であるとき、または ($S_b \times (1 - Q\%) > S_{a'} \times A_{max}$) であるとき、2 つの範囲には重複部分がない。

【0086】

そこで選択部 31 は、($S_{a'} \times A_{min} > S_b \times (1 + Q\%)$) であるか、または ($S_b \times (1 - Q\%) > S_{a'} \times A_{max}$) であるか否かを判定することにより、範囲 $S_b \times (1 - Q\%) \sim S_b \times (1 + Q\%)$ と範囲 $S_{a'} \times A_{min} \sim S_{a'} \times A_{max}$ とが重複しないか否かを判定する。これらの範囲が重複するとき、選択部 31 は処理をステップ S58 へ移す。ステップ S58 において選択部 31 は、位相シフトが所定の選択要件を満たすと決定して判定処理を終了する。

【0087】

図 11 のステップ S38 の判定において、所定の選択要件を満たす位相シフトがある場合には、選択部 31 は処理をステップ S39 へ移し、所定の選択要件を満たす位相シフトがない場合には、選択部 31 は処理をステップ S40 へ移す。

【0088】

ステップ S39 において選択部 31 は、所定の選択要件を満足する位相シフト量が与えられた音声フレーム信号のうち最小の最大振幅値を有する音声フレーム信号に与えられた位相シフト量を選択することにより、所定の選択要件を満たす位相シフトのうちから周波数 f_M の周波数成分に与える位相シフト量を選択する。選択部 31 は、選択した位相シフト量を示す位相選択信号を出力する。位相選択信号合成部 32 は、入力 $SLPin$ として入力した位相選択信号と、選択部 31 が出力した位相選択信号とを合成する。位相選択信号合成部 32 は、合成した位相選択信号を出力 $SLPout$ として出力する。その後、処理は S41 へ移る。

【0089】

ステップ S40 において選択部 31 は、所定の優先順序付け基準に従って、逆フーリエ変換部 30-1 ~ 30-12 により生成された各音声フレーム信号に与えられた位相シフト量のうち最も優先度が高い位相シフト量を、周波数 f_M の周波数成分に与える位相シフト量として選択する。優先順序付け基準として、各位相シフト量を与えた場合の、(1) 各音声フレーム信号の最大振幅値の大小、(2) 増幅部 5 において各音声フレーム信号にクリッピング歪みを生じさせずに増幅できる増幅率の範囲と目標増幅率 A との間の距離の大小、(3) 増幅部 5 において各音声フレーム信号にクリッピング歪みを生じない範囲で増幅させたときの各音声フレーム信号の最初のサンプル値と、その直前のフレームの最後

10

20

30

40

50

のサンプル値との差の大小、などを使用してよい。

【0090】

選択部31は、選択した位相シフト量を示す位相選択信号を出力する。位相選択信号合成部32は、入力SLPinとして入力した位相選択信号と、選択部31が出力した位相選択信号とを合成する。位相選択信号合成部32は、合成した位相選択信号を出力SLPoutとして出力する。その後、処理はS41へ移る。

【0091】

ステップS41において、図9に示す逆フーリエ変換部13は、フーリエ変換部10から与えられる各周波数成分に、位相選択部14から与えられる位相選択信号によって指定される各位相シフトをそれぞれ与えて、周波数領域信号を逆フーリエ変換した音声フレーム信号を生成する。

10

【0092】

図14は、図9に示す増幅率決定部4による信号増幅率の決定処理の第1例を示すフローチャートである。ステップS60において増幅率決定部4は、前フレームの最初のサンプルの値Sbの符号と、位相シフトを与えた現在のフレームの最後のサンプルの値Sa'の符号とが異なるか否かを判定する。Sbの符号とSa'の符号が同じ場合には処理はステップS62に移る。Sbの符号とSa'の符号が異なる場合には、ステップS61においてフレーム接続部7は現在のフレームの各サンプルの符号を反転させる。

【0093】

ステップS62において増幅率決定部4は、音声フレーム信号の最大振幅値Smaxが、所定値($S_{th} / (A_t \times (1 - b\%))$)より大きいか否かを判定する。Smax > ($S_{th} / (A_t \times (1 - b\%))$)であるとき増幅率決定部4は、処理をS63へ移行する。Smax > ($S_{th} / (A_t \times (1 - b\%))$)でないとき増幅率決定部4は、処理をS64へ移行する。

20

【0094】

Smax > ($S_{th} / (A_t \times (1 - b\%))$)であるとき、音声フレーム信号にクリッピング歪みを生じない最大振幅値Smaxであっても信号増幅率の許容範囲の下限值($A_t \times (1 - b\%)$)より小さい。したがってステップS63において増幅率決定部4は、信号増幅率Aを($A_t \times (1 - b\%)$)に決定して処理を終了する。

【0095】

ステップS64において増幅率決定部4は、 $S_{max} \leq (S_{th} / (A_t \times (1 + b\%)))$ であるか否かを判定する。 $S_{max} \leq (S_{th} / (A_t \times (1 + b\%)))$ であるとき増幅率決定部4は、処理をステップS65に移す。ステップS65において増幅率決定部4は、増幅部5にて使用可能な信号増幅率の上限値Amaxを($A_t \times (1 + b\%)$)に定め、下限値Aminを($A_t \times (1 - b\%)$)に定める。その後、増幅率決定部4は処理をステップS67に移す。

30

【0096】

ステップS64の判定において $S_{max} \leq (S_{th} / (A_t \times (1 + b\%)))$ でないとき増幅率決定部4は、処理をステップS66に移す。ステップS66において増幅率決定部4は、上限値Amaxを最大増幅率(S_{th} / S_{max})に定め、下限値Aminを($A_t \times (1 - b\%)$)に定める。その後、増幅率決定部4は処理をステップS67に移す。

40

【0097】

ステップS67において増幅率決定部4は、ステップS65又はS66において定められた範囲の信号増幅率Amin ~ Amaxによって現在の音声フレーム信号が増幅されたときの最初のサンプル値の範囲Sa' × Amin ~ Sa' × Amaxと、増幅後の現在の音声フレーム信号の最初のサンプル値Saに許される所定の許容範囲Sb × (1 - Q%) ~ Sb × (1 + Q%)と、が重複しないか否かを判定する。

【0098】

これらの範囲が重複しないとき、増幅率決定部4は処理をステップS68に移行し、こ

50

これらの範囲が重複するとき、増幅率決定部 4 は処理をステップ S 6 9 に移行する。ステップ S 6 8 において増幅率決定部 4 は、増幅率 $A_{min} \sim A_{max}$ のうち最も目標増幅率 A_t に近い増幅率を信号増幅率 A に選択し、処理を終了する。

【0099】

ステップ S 6 9 において増幅率決定部 4 は、増幅率 $A_{min} \sim A_{max}$ のうち、増幅前の現在のフレームの最初のサンプル値 S_a' を増幅した値が、前フレームの最後のサンプル値 S_b に最も近くなる増幅率を選択する。増幅率決定部 4 がこのような増幅率を選択することで、増幅後の現在のフレームの最初のサンプル値 S_a が前フレームの最後のサンプル値 S_b に最も近くなるような信号増幅率が選択され、フレーム間のサンプル値のギャップを低減できる。

10

【0100】

たとえば図 15 の (A) に示すように、増幅された現在の音声フレーム信号の最初のサンプル値の範囲 $S_a' \times A_{min} \sim S_a' \times A_{max}$ が、前フレームの最後のサンプル値 S_b よりも小さい範囲であるときは、増幅率決定部 4 は最大の増幅率 A_{max} を選択する。また、図 15 の (B) に示すように、増幅された現在の音声フレーム信号の最初のサンプル値の範囲 $S_a' \times A_{min} \sim S_a' \times A_{max}$ が、前フレームの最後のサンプル値 S_b よりも大きい範囲であるときは、増幅率決定部 4 は、最小の増幅率 A_{min} を選択する。

【0101】

図 15 の (C) に示すように、前フレームの最後のサンプル値 S_b が、増幅された現在の音声フレーム信号の最初のサンプル値の範囲 $S_a' \times A_{min} \sim S_a' \times A_{max}$ の範囲内にあるときは、増幅率決定部 4 は増幅率 (S_b / S_a') を選択する。

20

【0102】

図 16 は、図 9 に示す増幅率決定部 4 による信号増幅率の決定処理の第 2 例を示すフローチャートである。ステップ S 6 0 ～ S 6 8 までは図 14 を参照して説明した決定処理と同様である。ステップ S 6 7 における判定において、増幅後の最初のサンプル値の範囲 $S_a' \times A_{min} \sim S_a' \times A_{max}$ と、所定の許容範囲 $S_b \times (1 - Q\%) \sim S_b \times (1 + Q\%)$ とが重複するとき、増幅率決定部 4 は処理をステップ S 7 0 に移す。

【0103】

ステップ S 7 0 において増幅率決定部 4 は、範囲 $S_a' \times A_{min} \sim S_a' \times A_{max}$ と、範囲 $S_b \times (1 - Q\%) \sim S_b \times (1 + Q\%)$ との間の重複範囲 $S_{a1} \sim S_{a2}$ を決定する。増幅率決定部 4 により決定される重複範囲 $S_{a1} \sim S_{a2}$ の例を図 17 に示す。

30

【0104】

ステップ S 7 1 において増幅率決定部 4 は、値 $S_{a1} / S_a' \sim S_{a2} / S_a'$ のうち最も目標増幅率 A_t に近い値を信号増幅率として選択する。増幅率決定部 4 がこのような値を信号増幅率として選択することで、上記の所定の選択要件を満たしつつ、かつ現在のフレームの信号増幅率と前フレームの信号増幅率とのギャップを低減することができる。

【0105】

図 18 は、開示の音声処理装置の第 3 実施例の概略構成図である。図 18 に示す音声処理装置 1 は、図 9 に示す構成と類似する構成を有しており、図 9 に示す構成要素と同様の構成要素には同じ参照符号を使用し、また同一の機能については説明を省略する。

40

【0106】

最大値減少化処理部 3 は、図 2 を参照して説明した位相選択部 12-1 と同様の位相選択部を直列に M 段接続した位相選択部 12-1、12-2、… 12-M と、位相選択部 12-M の後段に接続され、直列に N 段接続される位相選択部 15-1 ～ 15-N を備える。

【0107】

図 19 は、図 18 に示す位相選択部 15-1 の構成例を示す図である。他の位相選択部 15-2 ～ 15-N も同様の構成を有する。位相選択部 15-i ($i = 1 \sim N$) には、フーリエ変換部 10 から出力される周波数領域信号が入力 S_f として入力される。また、位相

50

選択部 15-i には、周波数選択部 11 から出力される ($M+i$) 番目のスペクトル強度を有する周波数を指示する信号が入力SLfとして入力される。さらに位相選択部 15-i には、前段の位相選択部である、位相選択部 12-M 又は位相選択部 15-($i-1$) から出力SLPoutとして出力される位相選択信号が入力SLPinとして入力される。

【0108】

位相選択部 15-1 は、図 2 を参照して説明した逆フーリエ変換部 20-1 ~ 20-L と同様に動作する逆フーリエ変換部 40-1 ~ 40-L と、図 2 を参照して説明した位相選択信号合成部 22 と同様に動作する位相選択信号合成部 42 と、選択部 41 を備える。本実施例では自然数 $L=12$ の場合の構成例である。自然数 L は他の 2 以上の自然数を使用してもよい。

10

【0109】

また位相選択部 15-1 ~ 15-N には、目標増幅率決定部 8 が決定した目標増幅率 A_t と、フレーム記憶部 6 に記憶された前フレームの最後のサンプル値 S_b が入力される。選択部 41 は、逆フーリエ変換部 40-1 ~ 40-12 が生成する音声フレーム信号を入力する。

【0110】

選択部 41 は、図 12 を参照して説明した判定処理を行い、逆フーリエ変換部 40-1 ~ 40-12 により生成された各音声フレーム信号に与えられた位相シフトの中に、上述した所定の選択要件を満たす位相シフトがあるか否かを判定する。選択部 41 は、所定の選択要件を満たす位相シフトがある場合には値「1」を、その他の場合には値「0」を有する判定結果信号を出力Routとして出力する。

20

【0111】

位相選択部 15-i ($i=1 \sim N$) は、前段の位相選択部から出力Routとして出力された判定結果信号を入力Rinとして入力する。入力Rinとして入力された判定結果信号は、逆フーリエ変換部 40-1 ~ 40-12 及び選択部 41 へ入力される。

【0112】

逆フーリエ変換部 40-1 ~ 40-12 及び選択部 41 は、入力された判定結果信号の値が「1」であるとき、すなわち前段の位相選択部 15-($i-1$) にて選択要件を満たす位相シフトが見つかった場合には処理を停止し、このとき選択部 41 は、出力Routの値を「1」に設定する。但し、第 ($M+1$) 段目の位相選択部 15-1 の入力Rinには値「0」が入力される。

30

【0113】

位相選択部 15-i ($i=1 \sim N$) の選択部 41 は、所定の選択要件を満足する位相シフト量が与えられた音声フレーム信号のうち、最小の最大振幅値を有する音声フレーム信号の周波数 $f(M+i)$ の周波数成分に与えられた位相シフト量を選択する。選択部 41 は、選択された音声フレーム信号に与えられた位相シフト量を示す位相選択信号を位相選択信号合成部 42 へ出力する。位相選択信号合成部 42 は、入力SLPinとして入力した位相選択信号と、選択部 41 が出力した位相選択信号とを合成し、合成した位相選択信号を出力SLPoutとして出力する。

【0114】

前段の位相選択部 15-1 ~ 15-($N-1$) から出力される位相選択信号は、後段の位相選択部 15-2 ~ 15-N へ入力SLPinとして入力される。また各位相選択部 15-1 ~ 15-N から出力される位相選択信号はセレクタ 9 へも入力される。

40

【0115】

図 18 に示すとおり、セレクタ 9 は、各位相選択部 15-i ($i=1 \sim N$) から出力される判定結果信号をセレクト信号に使用し、値「1」である判定結果信号を出力した各位相選択部 15-i のうち、最前段に配置されている位相選択部から出力SLPoutとして出力される位相選択信号を選択して、逆フーリエ変換部 13 へ入力する。

【0116】

図 20 は、図 18 に示す最大値減少化処理部 3 により実行される音声信号の最大値の減

50

少化処理の第3例を示すフローチャートである。ステップS80～ステップS86では、図4に示したステップS10～S16にて第1番目～第M番目の各周波数の周波数成分に与える位相が選択されたのと同様に、第1番目～第M番目の各周波数の周波数成分に与える位相が選択される。但しステップS81において周波数選択部11は、第1番目～第(M+N)番目に強いスペクトル強度を有する周波数 f_i ($i=1\sim M+N$)を決定する。

【0117】

ステップS87において各位相選択部15-i ($i=1\sim N$)を参照するインデックス変数 i の値を「1」に初期化する。

【0118】

ステップS88において、(M+i)段目の位相選択部15-iは、第(M+i)番目のスペクトル強度を有する周波数 $f(M+i)$ を指示する信号を入力SLfとして受信する。

10

【0119】

位相選択部15-iの逆フーリエ変換部40-j ($j=1\sim 12$)は、フーリエ変換部10から与えられた各周波数成分のうち、入力SLfによって指定された周波数 $f(M+i)$ 以外の他の各周波数成分には、前段の位相選択部15-(i-1)から入力した位相選択信号によって指定される各シフト量の位相シフトをそれぞれ与え、かつ周波数 $f(M+i)$ の周波数成分にはそれぞれ($360/L \times (j-1)$)度の位相シフトを与えて時間領域信号に逆フーリエ変換する。

【0120】

20

ステップS89において、位相選択部15-iの選択部41は、逆フーリエ変換部40-1～40-12により生成された各音声フレーム信号に与えられた位相シフトの中に、上述した所定の選択要件を満たす位相シフトがあるか否かを判定する。ある位相シフトが所定の選択要件を満足するか否かを判定する判定処理は、図12を参照して示した処理と同様でよい。

【0121】

ステップS89の判定において、所定の選択要件を満たす位相シフトがあるとき選択部41は処理をステップS90へ移し、所定の選択要件を満たす位相シフトがないとき選択部41は処理をステップS91へ移す。

【0122】

30

ステップS90において選択部41は、図11に示すステップS39と同様にして周波数 $f(M+i)$ の周波数成分に与える位相シフト量を選択する。選択部41は、選択した位相シフト量を示す位相選択信号を出力する。位相選択信号合成部42は、入力SLPinとして入力した位相選択信号と、選択部41が出力した位相選択信号とを合成する。位相選択信号合成部42は、合成した位相選択信号を出力SLPoutとして出力する。その後、処理はS95へ移る。

【0123】

ステップS91において選択部41は、逆フーリエ変換部40-1～40-12により生成された各音声フレーム信号のうち、最大振幅値が最小である音声フレーム信号を選択する。選択部41は、逆フーリエ変換部40-1～40-12のうち、選択された音声フレーム信号の周波数成分 $f(M+i)$ に与えられた位相シフト量を示す位相選択信号を出力する。位相選択信号合成部42は、入力SLPinとして入力した位相選択信号と、選択部41が出力した位相選択信号とを合成する。位相選択信号合成部42は、合成した位相選択信号を出力SLPoutとして出力する。

40

【0124】

ステップS92において、インデックス変数 i の値を1つ増加する。ステップS93において、インデックス変数 i の値が「N」以下であるとき、すなわち、まだ位相選択処理が済んでいない位相選択部の段が残っている場合には、処理はステップS88へ戻り、ステップS88～S93が反復される。

【0125】

50

ステップ S 9 3 の判定において、インデックス変数 i の値が「N」以下でないとき、処理はステップ S 9 4 へ移る。ステップ S 9 4 において選択部 4 1 は、図 1 1 に示すステップ S 4 0 と同様にして周波数 f ($M+N$) の周波数成分に与える位相シフト量を選択する。その後、処理は S 9 5 へ移る。

【0126】

ステップ S 9 5 において図 1 8 に示すセクタ 9 は、各位相選択部 1 5 - i ($i = 1 \sim N$) から出力される判定結果信号をセレクト信号に使用して、各位相選択部 1 5 - i から出力される位相選択信号からいずれか 1 つを選択して逆フーリエ変換部 1 3 へ入力する。逆フーリエ変換部 1 3 は、入力された位相選択信号によって指定される位相シフトを、フーリエ変換部 1 0 から与えられる各周波数成分に与え、周波数領域信号を逆フーリエ変換した音声フレーム信号を生成する。

10

【0127】

本実施例により、所定の選択要件を満たす位相シフトを決定するのが容易な音声フレーム信号については、比較的少ない段数の位相選択部によってより少ない計算量で位相シフトを決定することができ、一方で所定の選択要件を満たす位相シフトを決定するのが難しい音声フレーム信号については、位相選択部の段数を動的に増やしてより適切な位相シフトを決定することができる。

【0128】

図 2 1 は、開示の音声処理装置の第 4 実施例の概略構成図である。図 2 1 に示す音声処理装置 1 は、図 1 8 に示す構成と類似する構成を有しており、図 1 8 に示す構成要素と同様の構成要素には同じ参照符号を使用し、また同一の機能については説明を省略する。

20

【0129】

フーリエ変換部 1 0 は、音声フレーム信号をフーリエ変換して、音声フレーム信号の M 個の周波数 f_i ($i = 1 \sim M$) の各周波数成分を示す周波数領域信号を生成する。周波数選択部 1 6 は、フーリエ変換部 1 0 から与えられる各周波数成分のスペクトル強度にしたがって、スペクトル強度が強い順に、各周波数 f_i を指示する信号を入力 SLf として位相選択部 1 5 - 1 へ順次入力する。

【0130】

最大値減少化処理部 3 は、図 1 8 に示す位相選択部 1 5 - 1 を備える。位相選択部 1 5 - 1 は、それぞれ出力 $SLPout$ 及び出力 $Rout$ として出力した位相選択信号及び判定結果信号を、入力 $SLPin$ 及び入力 Rin としてフィードバックする。

30

【0131】

位相選択部 1 5 - 1 は、第 i 番目のスペクトル強度の周波数 f_i の周波数成分に与える位相シフト量を選択したときに出力 $SLPout$ として出力した位相選択信号を、第 $(i+1)$ 番目のスペクトル強度の周波数 $f_{(i+1)}$ の周波数成分に与える位相シフト量を決定する際の入力 $SLPin$ としてフィードバックする。

【0132】

また、位相選択部 1 5 - 1 は、周波数 f_i の周波数成分に与える位相シフト量を選択したときに出力 $Rout$ として出力した判定結果信号を、周波数 $f_{(i+1)}$ の周波数成分に与える位相シフト量を決定する際の入力 Rin としてフィードバックする。

40

【0133】

最大値減少化処理部 3 はスイッチ 1 7 を備える。スイッチ 1 7 は、第 1 番目の周波数 f_1 の周波数成分に与える位相シフト量の選択の際に、入力 Rin には「0」を入力し、入力 $SLPin$ には、全ての周波数成分について位相シフト量を指定しない位相選択信号を入力する。

【0134】

位相選択部 1 5 - 1 から出力 $SLPout$ 及び出力 $Rout$ として出力された位相選択信号及び判定結果信号は、逆フーリエ変換部 1 3 に入力される。逆フーリエ変換部 1 3 は、判定結果信号の値が「1」になったときに入力された位相選択信号によって指定される位相シフトを、フーリエ変換部 1 0 から与えられる各周波数成分に与え、周波数領域信号を逆フーリ

50

エ変換した音声フレーム信号を生成する。

【0135】

本実施例の最大値減少化処理部3は、上述の所定の選択要件を満たす位相シフト量が見つかるか、またはフーリエ変換部10が生成したM個の周波数の周波数成分 $f_1 \sim f_M$ のすべてについて位相シフト量を決定し終わるまで、周波数 $f_1 \sim f_M$ の周波数成分に与えるべき各位相シフト量を、1段の位相選択部15-1によって選択することができる。

【0136】

図22は、図21に示す最大値減少化処理部3により実行される音声信号の最大値の減少化処理の第4例を示すフローチャートである。ステップS100において図21に示すフーリエ変換部10は、音声フレーム信号をフーリエ変換して、M個の周波数 f_i ($i = 1 \sim M$) について、音声フレーム信号の各周波数成分を示す周波数領域信号を生成する。

10

【0137】

ステップS101において周波数選択部16は、各周波数 f_i の周波数成分のスペクトル強度にしたがって、各周波数 f_i を指示する信号をスペクトル強度が強い順に位相選択部15-1へ入力する順番を決定する。ステップS102において、第1番目から第M番目のスペクトル強度の周波数 f_i を参照するインデックス変数 i の値を「1」に初期化する。

【0138】

ステップS103において、位相選択部15-1は、フーリエ変換部10から与えられた周波数領域信号に含まれる第 i 番目にスペクトル強度が強い周波数 f_i を指示する信号を入力SLfとして受信する。

20

【0139】

位相選択部15-1の逆フーリエ変換部40-j ($j = 1 \sim 12$) は、入力SLfによって指定された周波数 f_i 以外の他の各周波数成分には、周波数 $f_{(i-1)}$ に与える位相シフト量を選択したときに出力SLPoutとして出力した位相選択信号によって指定される各シフト量の位相シフトをそれぞれ与え、かつ周波数 f_i の周波数成分にはそれぞれ $(360/L \times (j-1))$ 度の位相シフトを与えて時間領域信号に逆フーリエ変換する。

【0140】

ステップS104において、位相選択部15-1の選択部41は、逆フーリエ変換部40-1~40-12により生成された各音声フレーム信号に与えられた位相シフトの中に、上述した所定の選択要件を満たす位相シフトがあるか否かを判定する。ある位相シフトが所定の選択要件を満足するか否かを判定する判定処理は、図12を参照して示した処理と同様でよい。

30

【0141】

ステップS104の判定において、所定の選択要件を満たす位相シフトがあるとき選択部41は処理をステップS105へ移し、所定の選択要件を満たす位相シフトがないとき選択部41は処理をステップS106へ移す。ステップS105において選択部41は、図11に示すステップS39と同様にして周波数 f_i の周波数成分に与える位相シフト量を選択する。

【0142】

選択部41は、選択した位相シフト量を示す位相選択信号を出力する。位相選択信号合成部42は、入力SLPinとして入力した位相選択信号と、選択部41が出力した位相選択信号とを合成する。位相選択信号合成部42は、合成した位相選択信号を出力SLPoutとして出力する。また選択部41は、判定結果信号の値を「0」から「1」へ変化させる。その後、処理はステップS110へ移る。

40

【0143】

ステップS106において、選択部41は、逆フーリエ変換部40-1~40-12により生成された各音声フレーム信号のうち、最大振幅値が最小である音声フレーム信号を選択する。選択部41は、逆フーリエ変換部40-1~40-12のうち、選択された音声フレーム信号の周波数成分 f_i に与えられた位相シフト量を示す位相選択信号を出力す

50

る。位相選択信号合成部 4 2 は、入力SLPinとして入力した位相選択信号と、選択部 4 1 が出力した位相選択信号とを合成する。位相選択信号合成部 4 2 は、合成した位相選択信号を出力SLPoutとして出力する。

【0144】

ステップ S 1 0 7 において、インデックス変数 i の値を 1 つ増加する。ステップ S 1 0 8 において、インデックス変数 i の値が「M」以下であるとき、すなわち、まだ位相選択処理が済んでいない周波数が残っている場合には、処理はステップ S 1 0 3 へ戻り、ステップ S 1 0 3 ~ S 1 0 8 が反復される。

【0145】

ステップ S 1 0 8 の判定において、インデックス変数 i の値が「M」以下でないとき、処理はステップ S 1 0 9 へ移る。ステップ S 1 0 9 において選択部 4 1 は、図 1 1 に示すステップ S 4 0 と同様にして周波数 f_M の周波数成分に与える位相シフト量を選択する。

【0146】

選択部 4 1 は、選択した位相シフト量を示す位相選択信号を出力する。位相選択信号合成部 4 2 は、入力SLPinとして入力した位相選択信号と、選択部 4 1 が出力した位相選択信号とを合成する。位相選択信号合成部 4 2 は、合成した位相選択信号を出力SLPoutとして出力する。また選択部 4 1 は、判定結果信号の値を「0」から「1」へ変化させる。その後処理は、ステップ S 1 1 0 へ移る。

【0147】

ステップ S 1 1 0 において図 2 1 に示す逆フーリエ変換部 1 3 は、判定結果信号の値が「1」になったときに入力された位相選択信号によって指定される位相シフトを、フーリエ変換部 1 0 から与えられる各周波数成分に与え、周波数領域信号を逆フーリエ変換した音声フレーム信号を生成する。

【0148】

図 2 3 は、開示の音声処理装置の第 5 実施例の概略構成図である。図 2 3 に示す音声処理装置 1 は、図 1 8 に示す構成と類似する構成を有しており、図 1 8 に示す構成要素と同様の構成要素には同じ参照符号を使用し、また同一の機能については説明を省略する。

【0149】

最大値減少化処理部 3 は、フーリエ変換部 1 0 と、逆フーリエ変換部 5 0 と、音声信号選択部 5 1 とを備える。フーリエ変換部 1 0 は、音声フレーム信号をフーリエ変換して、 K 個の周波数 f_i ($i = 1 \sim K$) について、音声フレーム信号の各周波数成分を示す周波数領域信号を生成する。

【0150】

逆フーリエ変換部 5 0 は、 K 個の周波数 f_i ($i = 1 \sim K$) の全ての周波数成分のそれぞれに、複数種類のシフト量 $\Delta\theta_j = (360 / L \times (j - 1))$ ($j = 1 \sim L$) 度の位相シフトのうちのいずれかを与える位相シフトの全ての組み合わせについて、それぞれの組み合わせの位相シフトを各周波数成分に与えて逆フーリエ変換を行うことにより、 L^K 個の音声フレーム信号を生成する。

【0151】

すなわち逆フーリエ変換部 5 0 は、各周波数 f_i の周波数成分にそれぞれ与えるべき位相シフト量が以下の表 1 によって与えられる、 L^K 組の位相シフトの組み合わせ $PS - 1 \sim PS - L^K$ の各々を、各周波数成分に与えて逆フーリエ変換を行うことによって、 L^K 個の音声フレーム信号を生成する。

【0152】

10

20

30

40

【表 1】

表 1

位相シフト量の 組み合わせ	周波数 f_i ($i=1\sim K$) に与える位相シフト量				
	f_1	f_2	...	$f_{(K-1)}$	f_K
PS-1	$\Delta\theta_1$	$\Delta\theta_1$	...	$\Delta\theta_1$	$\Delta\theta_1$
PS-2	$\Delta\theta_2$	$\Delta\theta_1$	...	$\Delta\theta_1$	$\Delta\theta_1$
⋮	⋮	⋮		⋮	⋮
PS-L	$\Delta\theta_L$	$\Delta\theta_1$	...	$\Delta\theta_1$	$\Delta\theta_1$
PS-(L+1)	$\Delta\theta_1$	$\Delta\theta_2$	...	$\Delta\theta_1$	$\Delta\theta_1$
PS-(L+1)	$\Delta\theta_2$	$\Delta\theta_2$	...	$\Delta\theta_1$	$\Delta\theta_1$
⋮	⋮	⋮		⋮	⋮
PS-(2×L)	$\Delta\theta_L$	$\Delta\theta_2$	...	$\Delta\theta_1$	$\Delta\theta_1$
PS-(2×L+1)	$\Delta\theta_1$	$\Delta\theta_3$	...	$\Delta\theta_1$	$\Delta\theta_1$
⋮	⋮	⋮		⋮	⋮
PS-L ²	$\Delta\theta_L$	$\Delta\theta_L$	...	$\Delta\theta_1$	$\Delta\theta_1$
⋮	⋮	⋮		⋮	⋮
PS-(L ^K -L)	$\Delta\theta_L$	$\Delta\theta_{(L-1)}$	...	$\Delta\theta_L$	$\Delta\theta_L$
PS-(L ^K -L+1)	$\Delta\theta_1$	$\Delta\theta_L$	...	$\Delta\theta_L$	$\Delta\theta_L$
⋮	⋮	⋮		⋮	⋮
PS-L ^K	$\Delta\theta_L$	$\Delta\theta_L$	...	$\Delta\theta_L$	$\Delta\theta_L$

【0153】

音声信号選択部 51 には、目標増幅率決定部 8 が決定した目標増幅率 A_t と、フレーム記憶部 6 に記憶された前フレームの最後のサンプル値 S_b が入力される。音声信号選択部 51 は、逆フーリエ変換部 50 により生成された各音声フレーム信号の最大振幅値に基づいて、これら各音声フレーム信号に与えられた位相シフトのうち、所定の選択要件を満足する音声フレーム信号があるか否かを判定する。

【0154】

ここである音声フレーム信号が選択されるための所定の選択要件は、図 9～図 12 を参照して説明した、位相シフト量が選択されるための所定の選択要件と同様の条件である。すなわち、ある音声フレーム信号に対して、上記の条件 (1)～(3) が満たす信号増幅率 A が存在することである。

【0155】

音声信号選択部 51 は、所定の選択要件を満足する音声フレーム信号のうち、最小の最大振幅値を有する音声フレーム信号を選択して増幅率決定部 4 及び増幅部 5 へ出力する。

【0156】

図 24 は、図 23 に示す最大値減少化処理部 3 により実行される音声信号の最大値の減少化処理の第 5 例を示すフローチャートである。ステップ S120 において、フーリエ変換部 10 は、音声フレーム信号をフーリエ変換して、 K 個の周波数 f_i ($i=1\sim K$) について、音声フレーム信号の各周波数成分を示す周波数領域信号を生成する。

【0157】

ステップ S 1 2 1 において逆フーリエ変換部 5 0 は、周波数 f_i ($i = 1 \sim K$) の全ての周波数成分のそれぞれに、複数種類のシフト量 $\Delta \theta_j = (360 / L \times (j - 1))$ ($j = 1 \sim L$) 度の位相シフトのうちのいずれかを与える位相シフトの全ての組み合わせ $PS - 1 \sim PS - L^K$ を各周波数成分に与えて逆フーリエ変換を行って音声フレーム信号を生成する。

【0158】

ステップ S 1 2 2 において音声信号選択部 5 1 は、逆フーリエ変換部 5 0 により生成された各音声フレーム信号の中に、上述した所定の選択要件を満たす音声フレーム信号があるか否かを判定する。ある音声フレーム信号が所定の選択要件を満足するか否かを判定する判定処理は、図 1 2 を参照して示した処理と同様の処理でよい。

10

【0159】

ステップ S 1 2 2 の判定において、所定の選択要件を満たす音声フレーム信号があるとき音声信号選択部 5 1 は処理をステップ S 1 2 3 へ移し、所定の選択要件を満たす音声フレーム信号がないとき音声信号選択部 5 1 は処理をステップ S 1 2 4 へ移す。

【0160】

ステップ S 1 2 3 において音声信号選択部 5 1 は、所定の選択要件を満足する音声フレーム信号のうち最小の最大振幅値を有する音声フレーム信号を選択し、処理を終了する。ステップ S 1 2 4 において音声信号選択部 5 1 は、所定の優先順序付け基準に従って、逆フーリエ変換部 5 0 により生成された各音声フレーム信号のうち最も優先度が高い音声フレーム信号を選択する。優先順序付け基準として、(1) 各音声フレーム信号の最大振幅値の大小、(2) 増幅部 5 において各音声フレーム信号にクリッピング歪みを生じさせずに増幅できる増幅率の範囲と目標増幅率 A との間の距離の大小、(3) 増幅部 5 において各音声フレーム信号にクリッピング歪みを生じない範囲で増幅させたときの各音声フレーム信号の最初のサンプル値と、その直前のフレームの最後のサンプル値との差の大小、などを使用してよい。

20

【0161】

本実施例では、全ての位相シフト量の組み合わせ $PS - 1 \sim PS - L^K$ を与えた場合の音声フレーム信号について比較を行うので、より最適な音声フレーム信号を選択できる。

【0162】

図 2 5 は、開示の音声処理装置の第 6 実施例の概略構成図である。図 2 5 に示す音声処理装置 1 は、図 2 3 に示す構成と類似する構成を有しており、図 2 3 に示す構成要素と同様の構成要素には同じ参照符号を使用し、また同一の機能については説明を省略する。

30

【0163】

最大値減少化処理部 3 は、異なる周波数一位相特性を有する複数のオールパスフィルタ $60 - 1 \sim 60 - T$ と、音声信号選択部 6 1 とを備える。フレーム分割部 2 から出力される音声フレーム信号は、並列に配置されたオールパスフィルタ $60 - 1 \sim 60 - T$ によってフィルタリングされる。

【0164】

図 2 6 は、オールパスフィルタ $60 - 1 \sim 60 - T$ の周波数一位相特性の例を示す特性図である。オールパスフィルタ $60 - 1 \sim 60 - T$ は、入力信号の周波数成分に応じて、それぞれの周波数成分に異なるシフト量 $\Delta \theta$ の位相シフトを与えるフィルタである。図示の $C1 \sim C3$ はそれぞれ異なる周波数一位相特性を示し、これらの特性 $C1 \sim C3$ の間において、各周波数における位相シフト量が異なっている。

40

【0165】

オールパスフィルタ $60 - 1 \sim 60 - T$ として、 $C1 \sim C3$ に示すような異なる周波数一位相特性を有するフィルタを使用することにより、ユーザの聴覚により感知される音声品質を劣化させずに、異なる波形を有する、すなわち異なる最大振幅値を有する音声信号を生成することができる。したがって、オールパスフィルタ $60 - 1 \sim 60 - T$ は、上述した図 2 3 において複数の異なるシフト量の位相シフトを与えて逆フーリエ変換を行う逆フーリエ変換部 5 0 の代わりに使用することができる。

50

【0166】

図27は、(A)～(D)はオールパスフィルタの第1～4構成例を示す構成図であり、図28は、(A)及び(B)はオールパスフィルタの第5及び6構成例を示す構成図である。各図において要素60及び61はそれぞれ増幅率b1及びb2で信号増幅を行う増幅器を示し、要素70～73は1サンプル分の遅延を与える遅延素子を示し、要素80～82は加算器を示す。増幅率b1及びb2を変えて各例のオールパスフィルタを構成することによって、異なる周波数一位相特性を有するオールパスフィルタを実現することが可能である。

【0167】

各オールパスフィルタ60-1～60-Tの各々によってフィルタリングされた音声フレーム信号は、図25の音声信号選択部61に入力される。音声信号選択部61には、目標増幅率決定部8が決定した目標増幅率A_tと、フレーム記憶部6に記憶された前フレームの最後のサンプル値S_bが入力される。音声信号選択部61は、オールパスフィルタ60-1～60-Tによりそれぞれフィルタリングされた各音声フレーム信号の最大振幅値に基づいて、これら各音声フレーム信号に与えられた位相シフトのうち、上述の所定の選択要件を満足する音声フレーム信号があるか否かを判定する。

【0168】

音声信号選択部61は、所定の選択要件を満足する音声フレーム信号のうち、最小の最大振幅値を有する音声フレーム信号を選択して増幅率決定部4及び増幅部5へ出力する。

【0169】

図29は、図25に示す最大値減少化処理部3により実行される音声信号の最大値の減少化処理の第6例を示すフローチャートである。ステップS130において、フレーム分割部2から出力される音声フレーム信号を、各オールパスフィルタ60-1～60-Tにてフィルタリングする。

【0170】

ステップS131において音声信号選択部61は、各オールパスフィルタ60-1～60-Tによってフィルタリングした後の各音声フレーム信号の中に、上述した所定の選択要件を満たす音声フレーム信号があるか否かを判定する。ある音声フレーム信号が所定の選択要件を満足するか否かを判定する判定処理は、図12を参照して示した処理と同様の処理でよい。

【0171】

ステップS131の判定において、所定の選択要件を満たす音声フレーム信号があるとき音声信号選択部61は処理をステップS132へ移し、所定の選択要件を満たす音声フレーム信号がないとき音声信号選択部61は処理をステップS133へ移す。

【0172】

ステップS132において音声信号選択部61は、所定の選択要件を満足する音声フレーム信号のうち最小の最大振幅値を有する音声フレーム信号を選択し、処理を終了する。ステップS133において音声信号選択部61は、図24のステップS124と同様の処理により各音声フレーム信号を選択し、処理を終了する。

【0173】

本構成例によれば、フーリエ変換及び逆フーリエ変換を行うことなく、簡易な構成で最大値減少化処理部3を実現することができる。

【0174】

以上の実施例を含む実施形態に関し、更に以下の付記を開示する。

(付記1)

ディジタル音声信号を所定長毎に分割した音声フレーム信号の周波数成分に、互いに異なる位相シフトを与えることにより得られる複数の異なる音声フレーム信号の、それぞれの最大振幅値を決定する最大振幅値決定手段と、

前記複数の異なる音声フレーム信号のうち最大振幅値が最も小さいものを選択する選択手段と、を備える音声信号処理装置。

10

20

30

40

50

【0175】

(付記2)

前記最大振幅値決定手段は、

前記音声フレーム信号の各周波数成分を決定する周波数成分決定手段と、

前記各周波数成分にそれぞれ与えられる位相シフト量の複数の組み合わせを決定する組み合わせ決定手段と、を備え、

前記選択手段は、前記組み合わせ決定手段により決定された前記複数の組み合わせのうち、前記音声フレーム信号の最大振幅値を最も小さくする組み合わせを選択する付記1に記載の音声信号処理装置。

【0176】

(付記3)

前記組み合わせ決定手段は、前記周波数成分決定手段により決定されたいずれか1つの周波数成分に、互いに異なるシフト量の位相シフトを与えて複数の候補信号を生成する候補生成手段を備え、

前記選択手段は、前記複数の候補信号のうち最大振幅値が最も小さい候補信号に与えられたシフト量を選択する候補選択手段を備える、付記2に記載の音声信号処理装置。

【0177】

(付記4)

前記候補生成手段及び前記候補選択手段は、複数の前記周波数成分についてスペクトル強度が大きい順に前記複数の候補信号の生成と前記シフト量の選択を行い、

前記候補生成手段は、各周波数成分について前記複数の候補信号を作成するとき、当該周波数成分よりも以前にシフト量が選択された他の周波数成分に対して、それぞれの周波数成分について選択されたシフト量を与えて前記複数の候補信号を作成する付記3に記載の音声信号処理装置。

【0178】

(付記5)

前記候補生成手段及び前記候補選択手段は、前記周波数成分決定手段が決定した各周波数成分のうちから順次選択された周波数成分について、前記複数の候補信号の生成と前記シフト量の選択を行い、

前記候補生成手段は、各周波数成分について前記複数の候補信号を作成するとき、当該周波数成分よりも以前にシフト量が選択された他の周波数成分に対して、それぞれの周波数成分について選択されたシフト量を与えて前記複数の候補信号を作成し、

前記候補生成手段及び前記候補選択手段は、各周波数成分に対してそれぞれ選択されたシフト量を与えることにより生成される候補信号の最大振幅値が所定の閾値より小さくなったとき、前記候補信号の生成とシフト量の選択を停止する付記3に記載の音声信号処理装置。

【0179】

(付記6)

前記候補選択手段は、最大振幅値に基づいて、所定の増幅器において所定の許容範囲内の信号増幅率にて各候補信号を増幅できるか否かをそれぞれ判定し、増幅できると判定された各候補信号にそれぞれ与えられたシフト量の中から、前記シフト量の選択を行う付記3～5のいずれか一項に記載の音声信号処理装置。

【0180】

(付記7)

前記音声信号処理装置は、現在の音声フレーム信号の1つ前に処理された前フレームの少なくとも最後のサンプルを記憶するフレーム記憶手段を備え、

前記候補選択手段は、前記候補信号の最大振幅値に応じて信号増幅率を決定する増幅率決定手段と、決定された前記信号増幅率で増幅されたときの前記候補信号の最初のサンプル値を決定するサンプル値決定部と、を備え、

前記候補選択手段は、前記サンプル値決定部によって決定されたサンプル値が、前フレ

10

20

30

40

50

ームの最後のサンプル値から所定の許容範囲内に収まる各候補信号にそれぞれ与えられたシフト量の中から、前記シフト量の選択を行う付記 3～5 のいずれか一項に記載の音声信号処理装置。

【0181】

(付記 8)

前記組み合わせ決定手段は、前記周波数成分決定手段により決定される全ての周波数成分の各々に対して、シフト量が異なる複数の位相シフトの各々を与えて複数の音声フレーム信号を生成する候補生成手段を備え、

前記選択手段は、前記複数の音声フレーム信号のうち最大振幅値が最も小さいものを選択する付記 2 に記載の音声信号処理装置。

10

【0182】

(付記 9)

前記最大振幅値決定手段は、前記音声フレーム信号をフィルタリングする、異なる周波数一位相特性を有する複数のオールパスフィルタを備え、

前記選択手段は、前記複数のオールパスフィルタによりフィルタリングされた前記音声フレーム信号のうち、最大振幅値を最も小さいものを選択する付記 1 に記載の音声信号処理装置。

【0183】

(付記 10)

前記選択手段は、最大振幅値に基づいて、所定の増幅器において所定の許容範囲内の信号増幅率にて前記音声フレーム信号を増幅できるか否かをそれぞれ判定し、増幅できると判定された音声フレーム信号の中から、音声フレーム信号の選択を行う付記 8 又は 9 に記載の音声信号処理装置。

20

【0184】

(付記 11)

前記音声信号処理装置は、現在の音声フレーム信号の 1 つ前に処理された前フレームの少なくとも最後のサンプルを記憶するフレーム記憶手段を備え、

前記選択手段は、前記音声フレーム信号の最大振幅値に応じて信号増幅率を決定する増幅率決定手段と、決定された前記信号増幅率で増幅されたときの前記音声フレーム信号の最初のサンプル値を決定するサンプル値決定部と、を備え、

30

前記選択手段は、前記サンプル値決定部によって決定されたサンプル値が、前フレームの最後のサンプル値から所定の許容範囲内に収まる音声フレーム信号のうちから、音声フレーム信号の選択を行う付記 8 又は 9 に記載の音声信号処理装置。

【0185】

(付記 12)

前記選択手段により選択された前記音声フレーム信号の最大振幅値に応じた信号増幅率で前記音声フレーム信号を増幅する信号増幅手段と、

現在の音声フレーム信号の 1 つ前に処理された前フレームへ、前記信号増幅手段により増幅された前記音声フレーム信号を接続するフレーム接続手段と、をさらに備え、

該フレーム接続手段は、前記音声フレーム信号の最初のサンプル値と前記前フレームの最後のサンプル値との間に存在する目標値を選択し、前記音声フレーム信号の初めの複数サンプルの値と前記前フレームの最後の複数サンプルの値とを前記目標値に向かって漸近させる付記 1～11 に記載の音声信号処理装置。

40

【0186】

(付記 13)

ディジタル音声信号を所定長毎に分割した音声フレーム信号の周波数成分に位相シフトを与えることにより前記音声フレーム信号の最大振幅値を減少化させる最大値減少化手段と、

最大振幅値が減少化された後の音声フレーム信号の最大振幅値に応じて決定される信号増幅率で、前記最大振幅値が減少化された後の音声フレーム信号を増幅する信号増幅手段

50

と、

を備える音声信号処理装置。

【0187】

(付記14)

ディジタル音声信号を所定長の音声フレーム信号に分割し、
分割された前記音声フレーム信号の周波数成分に互いに異なる位相シフトを与えること
により得られる複数の異なる音声フレーム信号の、それぞれの最大振幅値を決定し、
前記複数の異なる音声フレーム信号のうち最大振幅値が最も小さいものを選択する、
音声信号処理方法。

【0188】

10

(付記15)

ディジタル音声信号を所定長の音声フレーム信号に分割し、
分割された前記音声フレーム信号の周波数成分に位相シフトを与えることにより前記音
声フレーム信号の最大振幅値を減少化させ、
最大振幅値が減少化された後の音声フレーム信号の最大振幅値に応じて決定される信号
増幅率で、前記最大振幅値が減少化された後の音声フレーム信号を増幅する、
音声信号処理方法。

【図面の簡単な説明】

【0189】

20

【図1】開示の音声処理装置の第1実施例の概略構成図である。

【図2】図1に示す位相選択部の構成例を示す図である。

【図3】開示の音声処理方法の実施例の全体フローチャートである。

【図4】音声信号の最大値の減少化処理の第1例を示すフローチャートである。

【図5】(A)及び(B)は音声信号の最大値の低減化処理の前後の音声フレーム信号の
波形の模式図である。

【図6】図1に示す増幅率決定部による信号増幅率の決定処理の例を説明する説明図であ
る。

【図7】図1に示すフレーム接続部による音声フレーム信号の接続処理の例を示すフロー
チャートである。

【図8】(A)及び(B)は、図1に示すフレーム接続部による音声フレーム信号の接続
処理の例の説明図である。

30

【図9】開示の音声処理装置の第2実施例の概略構成図である。

【図10】図9に示す位相選択部の構成例を示す図である。

【図11】音声信号の最大値の減少化処理の第2例を示すフローチャートである。

【図12】ある位相シフトが所定の選択要件を満足するか否かを判定する判定処理のフロ
ーチャートである。

【図13】(A)～(D)は、図12に示す判定処理の説明図である。

【図14】図9に示す増幅率決定部による信号増幅率の決定処理の第1例を示すフローチ
ャートである。

【図15】(A)～(C)は、図9に示す増幅率決定部による信号増幅率の決定処理の第
1例の説明図である。

40

【図16】図9に示す増幅率決定部による信号増幅率の決定処理の第2例を示すフローチ
ャートである。

【図17】図9に示す増幅率決定部による信号増幅率の決定処理の第2例の説明図である
。

【図18】開示の音声処理装置の第3実施例の概略構成図である。

【図19】図18に示す位相選択部の構成例を示す図である。

【図20】音声信号の最大値の減少化処理の第3例を示すフローチャートである。

【図21】開示の音声処理装置の第4実施例の概略構成図である。

【図22】音声信号の最大値の減少化処理の第4例を示すフローチャートである。

50

【図 2 3】開示の音声処理装置の第 5 実施例の概略構成図である。

【図 2 4】音声信号の最大値の減少化処理の第 5 例を示すフローチャートである。

【図 2 5】開示の音声処理装置の第 6 実施例の概略構成図である。

【図 2 6】 オールパスフィルタの周波数一位相特性を示す特性図である。

【図 27】 (A) ~ (D) はオールパスフィルタの第 1 ~ 4 構成例を示す構成図である。

【図 28】(A) 及び (B) はオールパスフィルタの第 5 及び 6 構成例を示す構成図である。

【図 2 9】音声信号の最大値の減少化処理の第 6 例を示すフローチャートである。

【符号の説明】

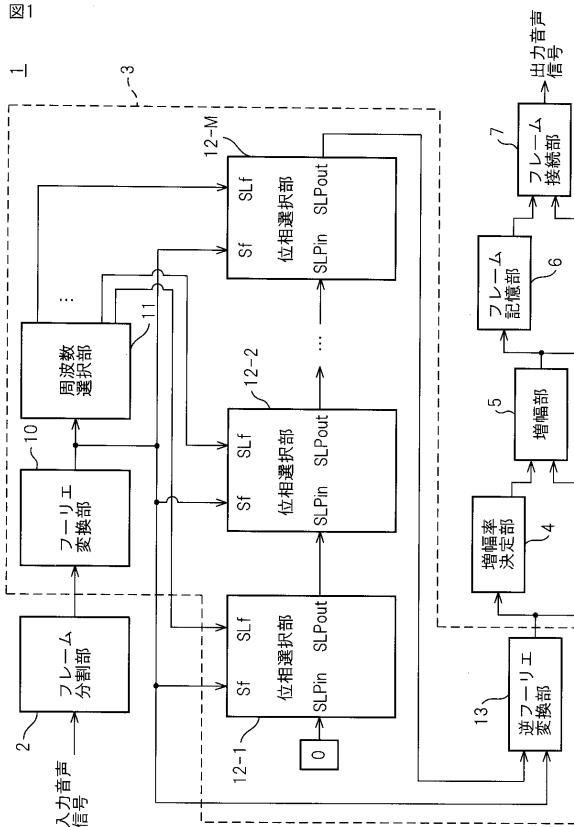
【 0 1 9 0 】

- 1 音声信号処理装置
3 最大値減少化処理部

10

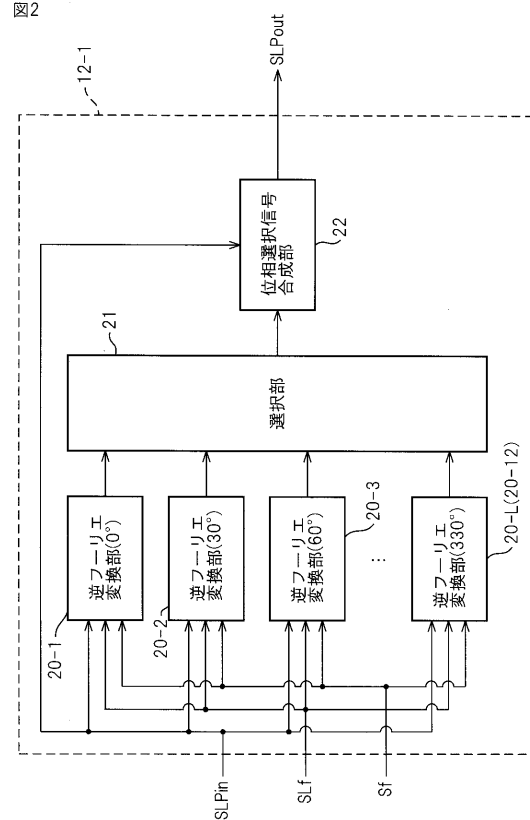
【 図 1 】

图1



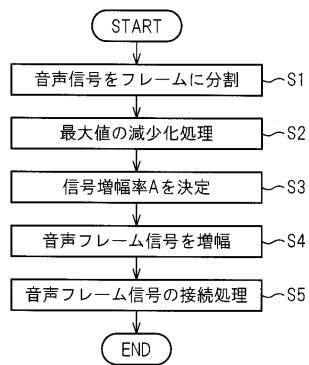
【図 2】

图2



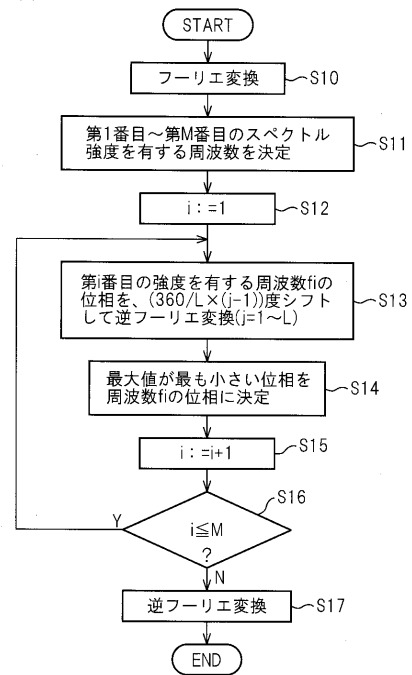
【図 3】

図3



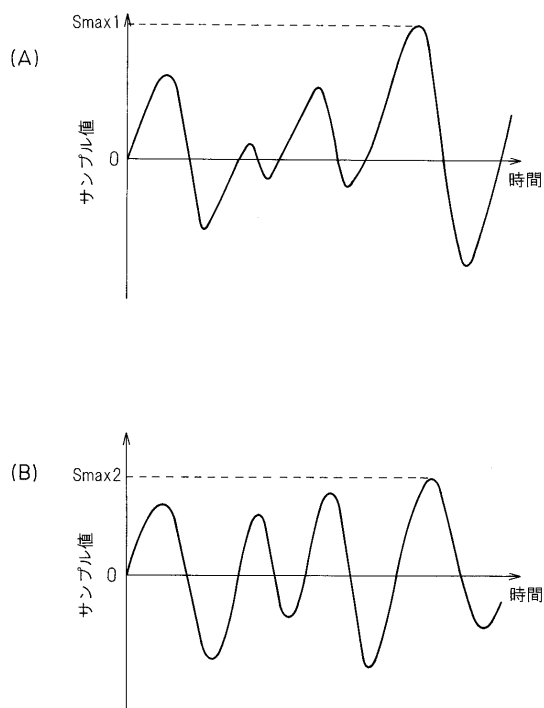
【図 4】

図4



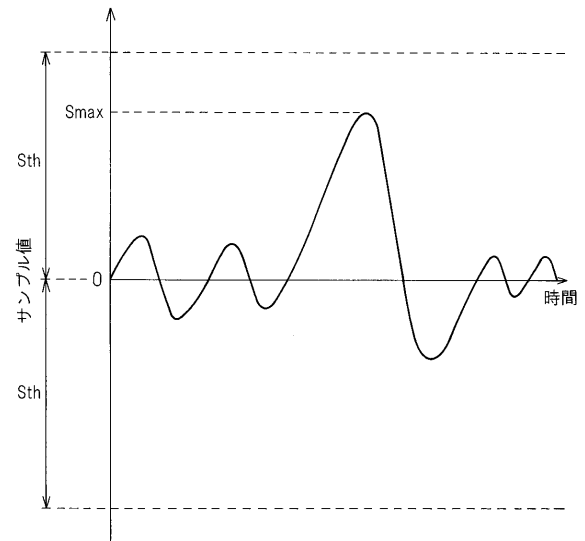
【図 5】

図 5

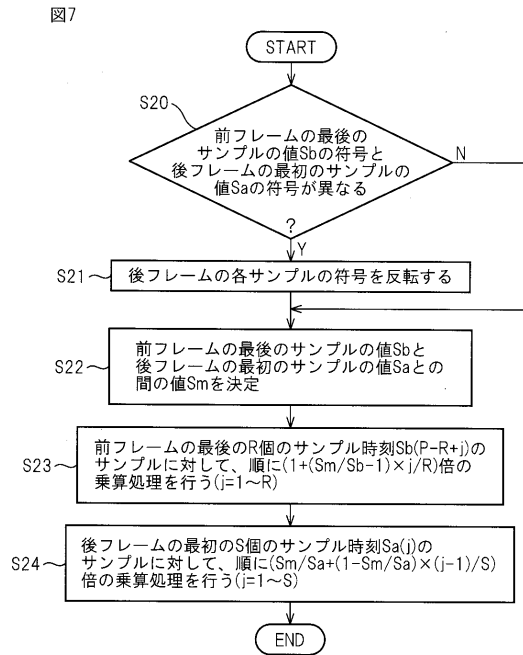


【図 6】

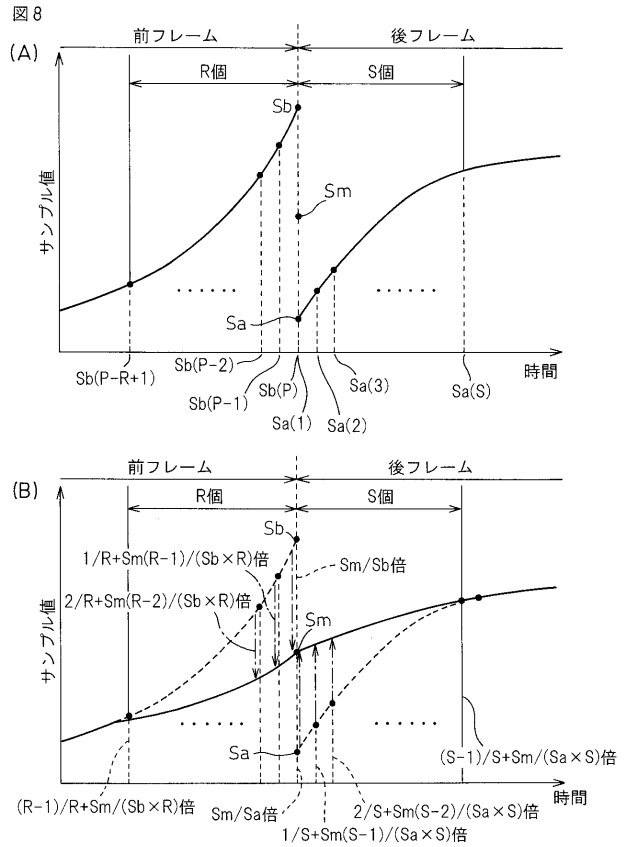
図 6



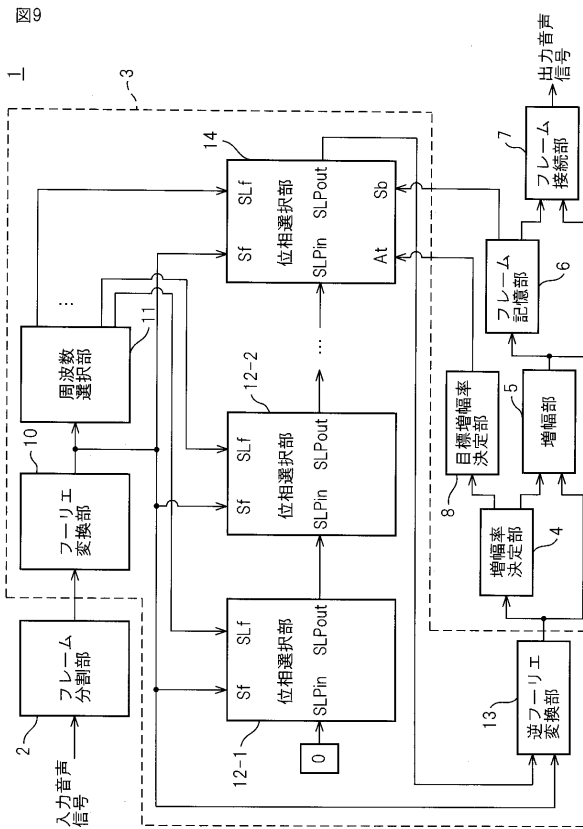
【図 7】



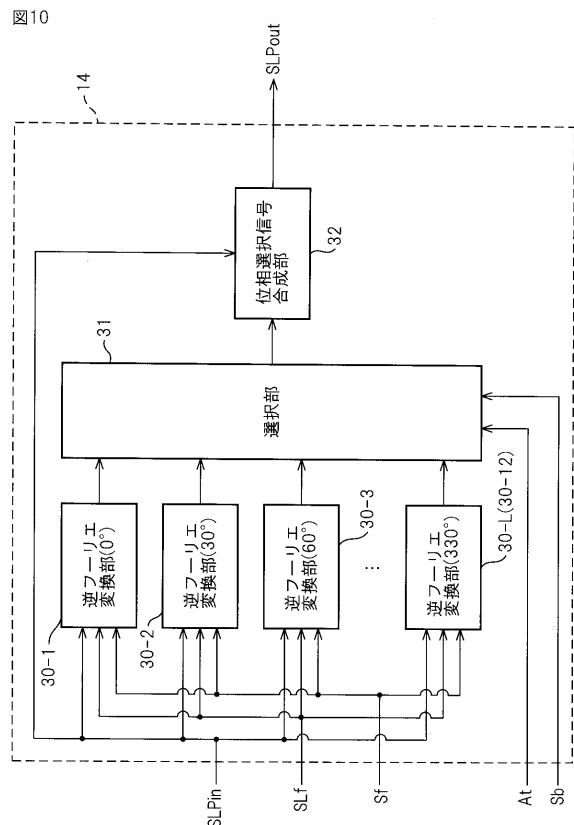
【図 8】



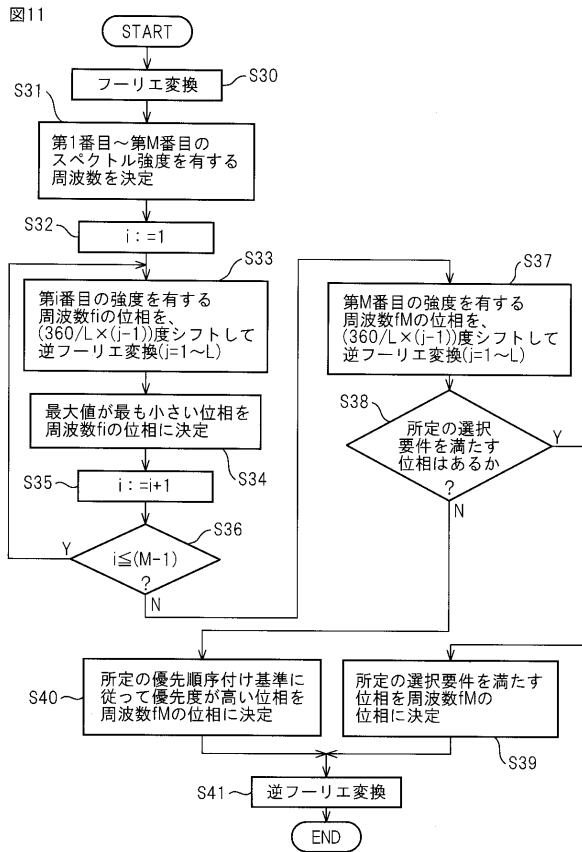
【図 9】



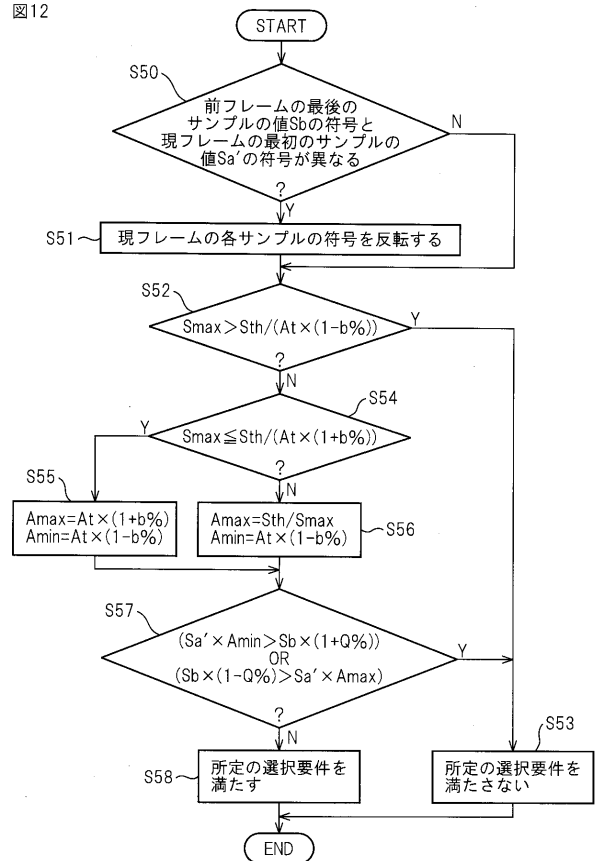
【図 10】



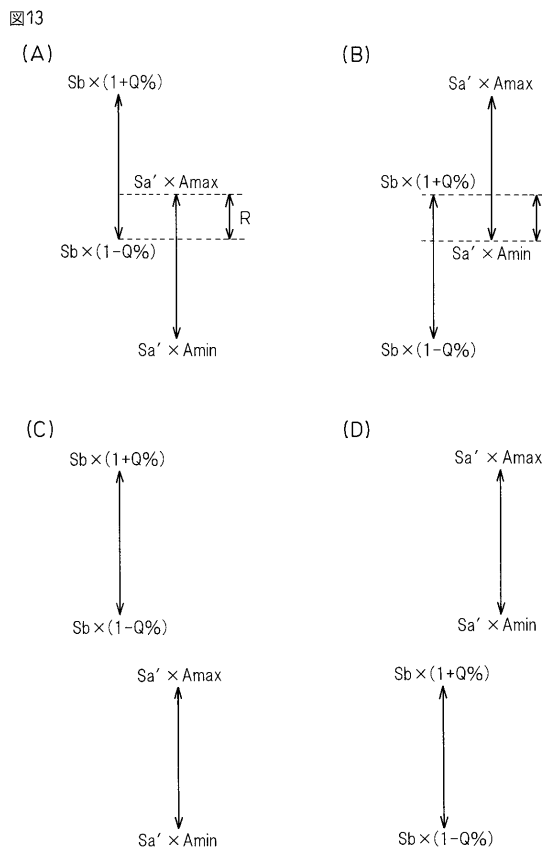
【図 1 1】



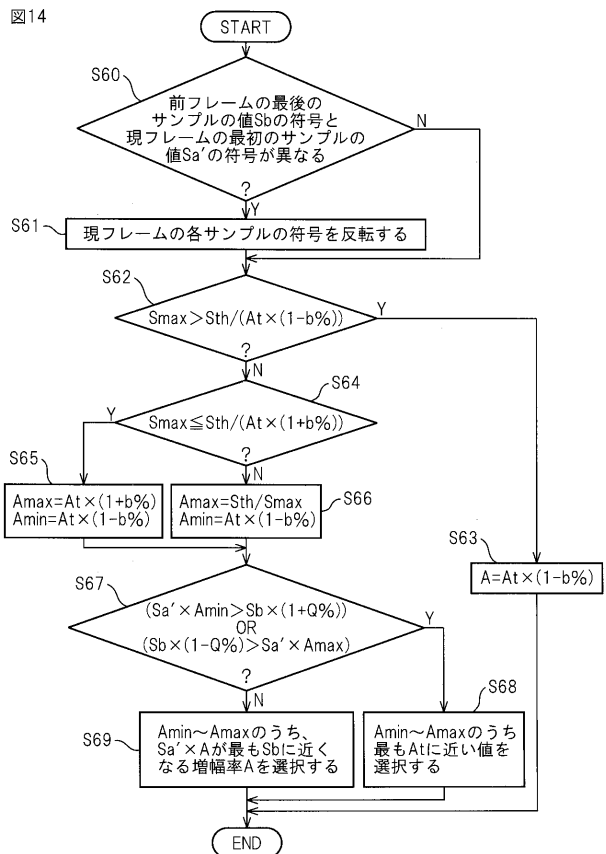
【図 1 2】



【図 1 3】

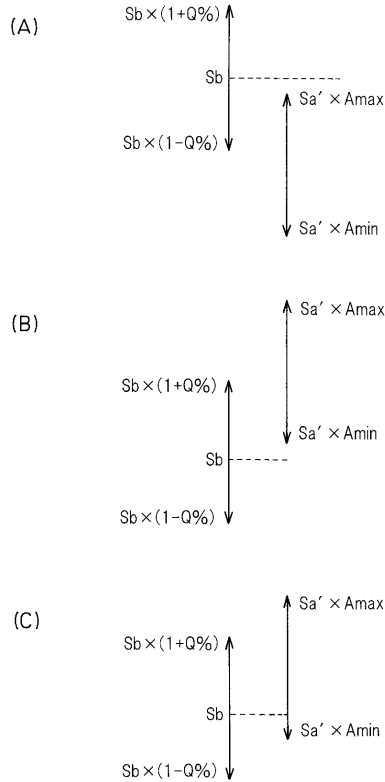


【図 1 4】



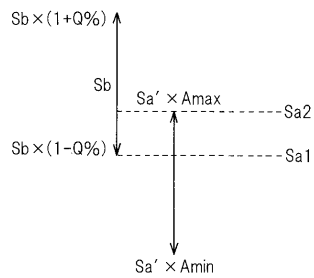
【図 15】

図15



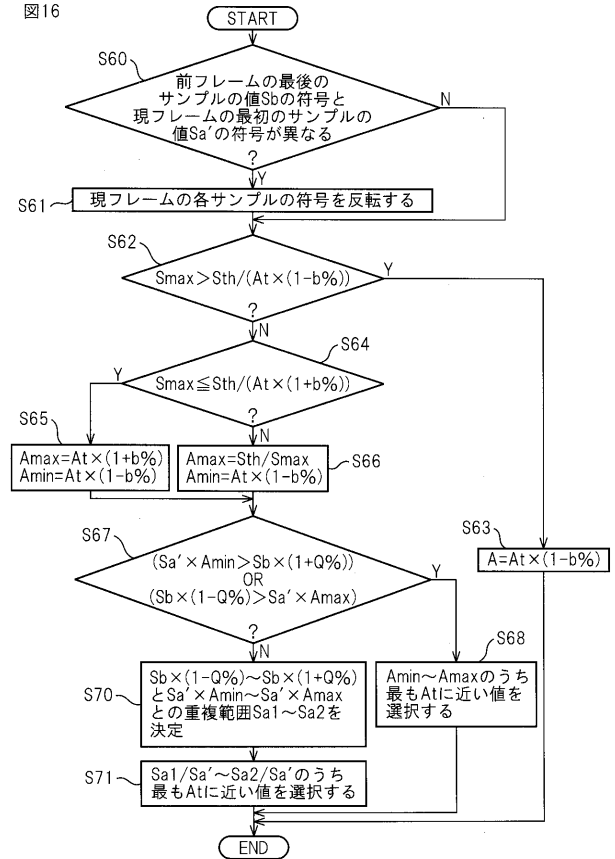
【図 17】

図17



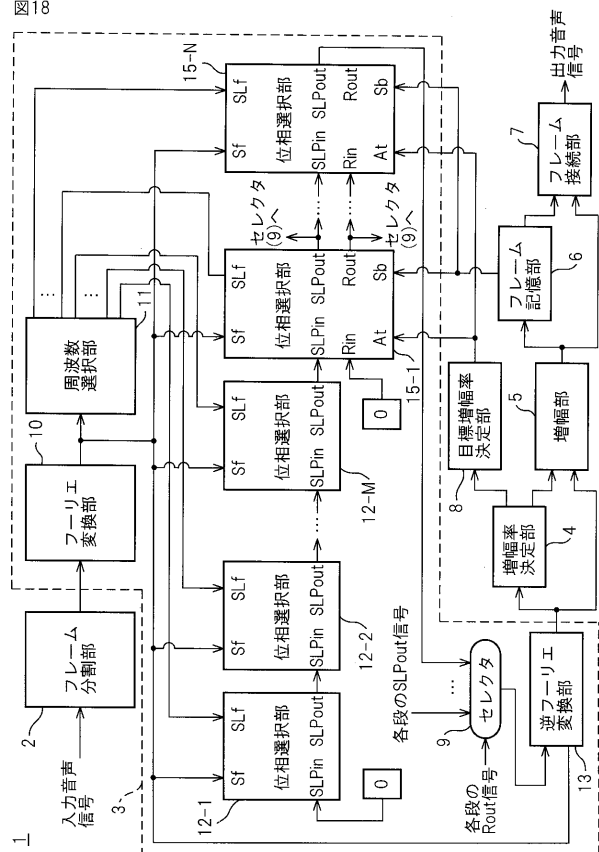
【図 16】

図16

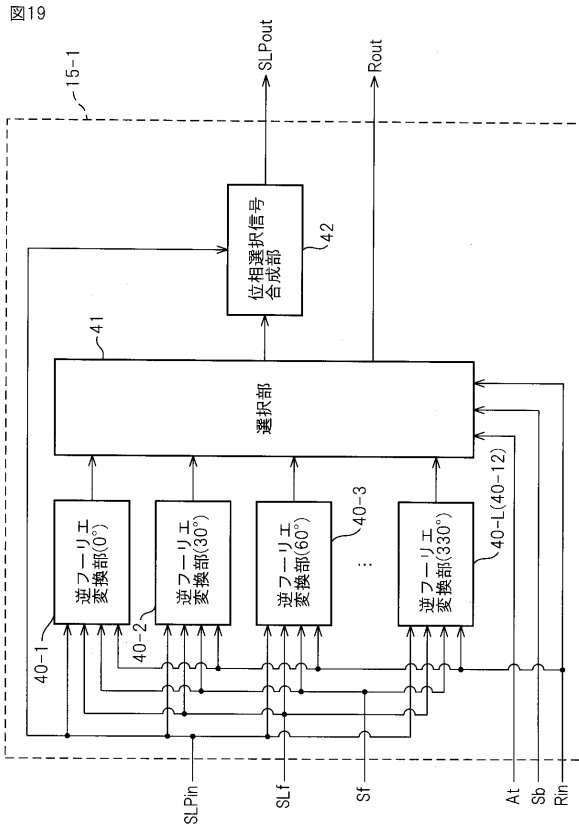


【図 18】

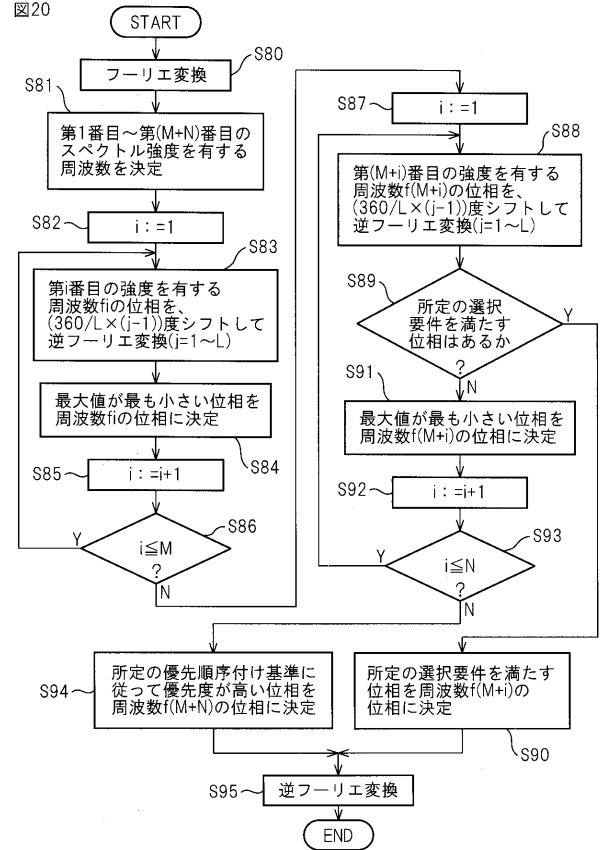
図18



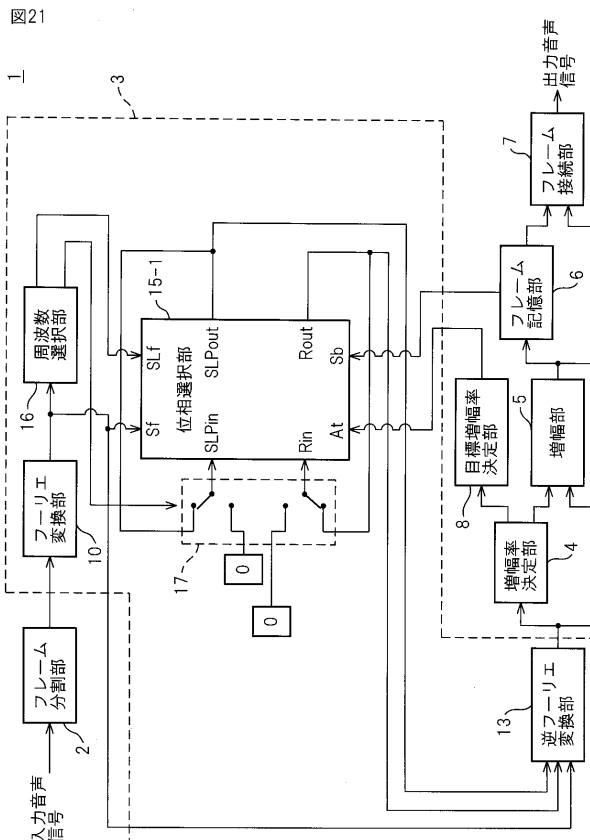
【図 19】



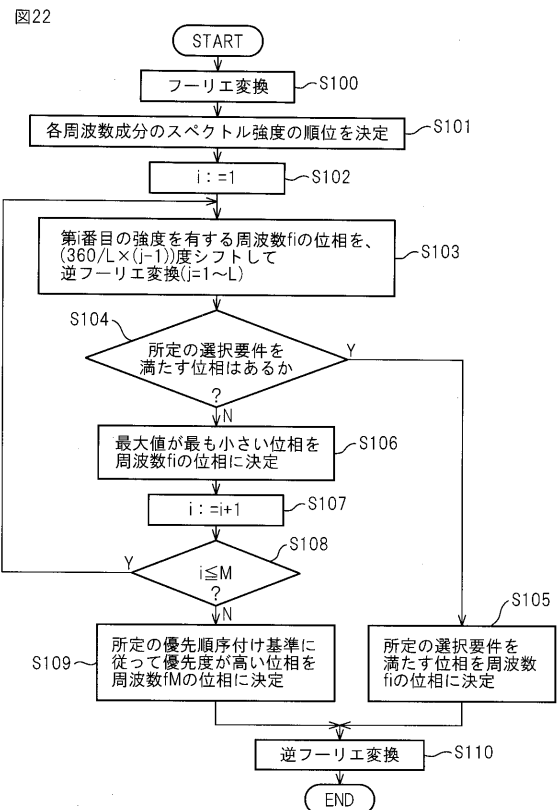
【図 20】



【図 21】

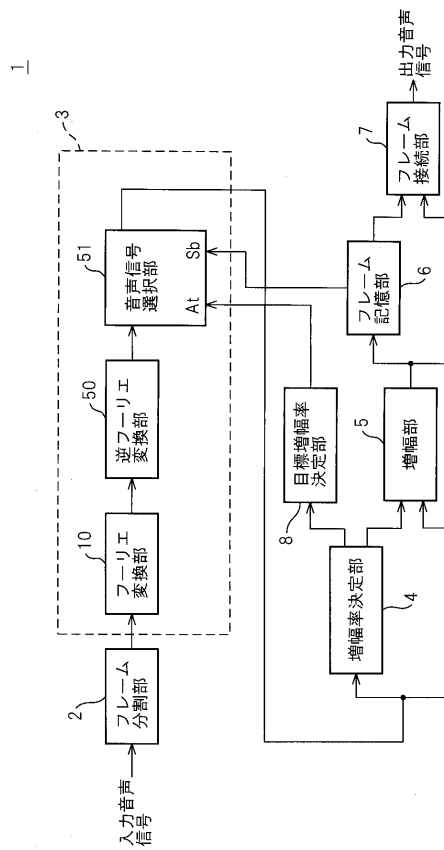


【図 22】



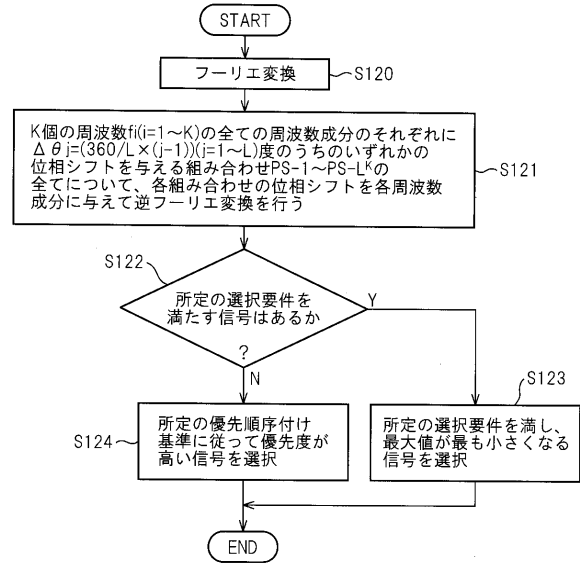
【図 2 3】

図23



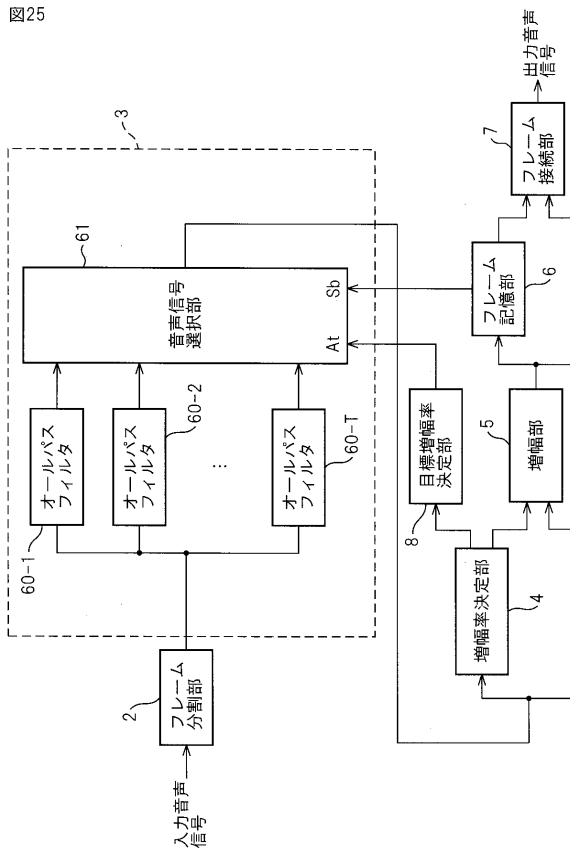
【図 2 4】

図24



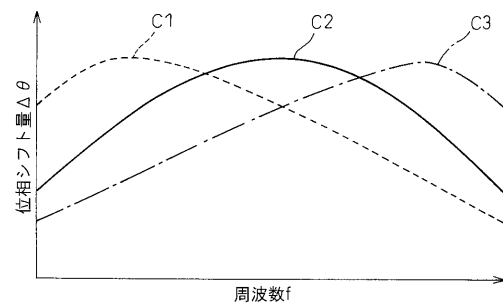
【図 2 5】

図25



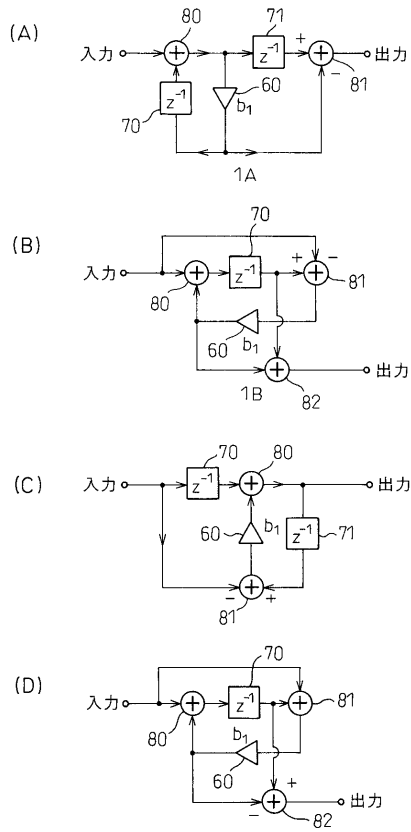
【図 2 6】

図26



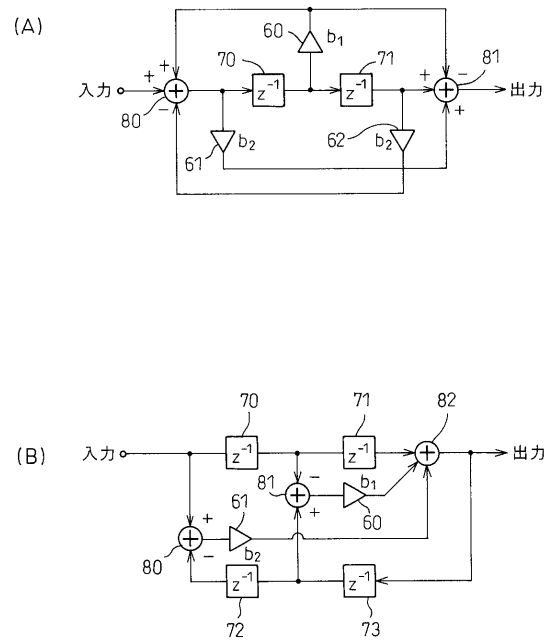
【図 27】

図27



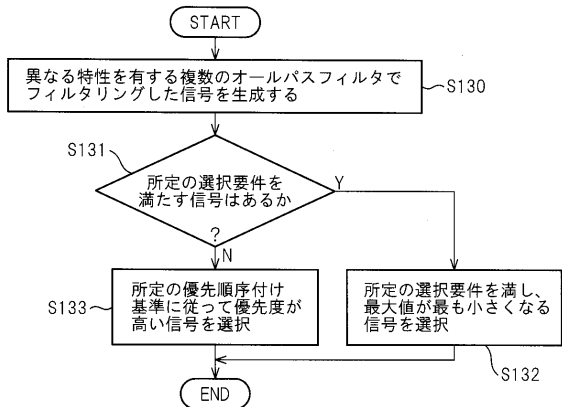
【図 28】

図28



【図 29】

図29



フロントページの続き

(74)代理人 100114177

弁理士 小林 龍

(72)発明者 天野 文雄

神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内