

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2009年3月26日 (26.03.2009)

PCT

(10) 国際公開番号
WO 2009/038136 A1

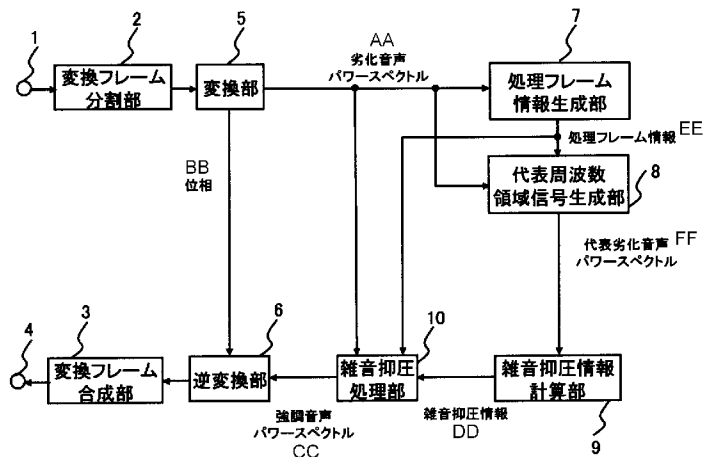
- (51) 国際特許分類:
G10L 21/02 (2006.01) G10L 15/20 (2006.01)
- (21) 国際出願番号: PCT/JP2008/066871
- (22) 国際出願日: 2008年9月18日 (18.09.2008)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願2007-243001 2007年9月19日 (19.09.2007) JP
- (71) 出願人 (米国を除く全ての指定国について): 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目7番1号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 嶋田 修 (SHIMADA, Osamu) [JP/JP]; 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP).
- (74) 代理人: 宇高 克己 (UDAKA, Katsuki); 〒1010025 東京都千代田区神田佐久間町1-1-4 第二東ビル5階 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ユーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,

[続葉有]

(54) Title: NOISE SUPPRESSION DEVICE, ITS METHOD, AND PROGRAM

(54) 発明の名称: 雑音抑圧装置、その方法及びプログラム

[図1]



(57) Abstract: A noise suppression device includes: conversion means which converts an input signal into a frequency region signal for each predetermined first frame; frame generation means which generates a second frame which is different from the first frame; representative frequency region signal generation means which generates a representative frequency region signal from the frequency region signal of the first frame contained in the second frame; and noise suppression degree calculation means which obtains a noise suppression degree of the second frame according to the representative frequency region signal.

(57) 要約: 本発明は、入力信号を定められた第一のフレーム毎に周波数領域信号に変換する変換手段と、第一のフレームと異なるように第二のフレームを生成するフレーム生成手段と、第二のフレームに含まれる第一のフレームの周波数領域信号から代表周波数領域信号を生成する代表周波数領域信号生成手段と、代表周波数領域信号に基づいて、第二のフレームの雑音抑圧の程度を求める雑音抑制程度算出手段とを有することを特徴とする雑音抑圧装置である。

- AA DEGRADED AUDIO POWER SPECTRUM
- BB PHASE
- CC EMPHASIZED AUDIO POWER SPECTRUM
- DD NOISE SUPPRESSION INFORMATION
- EE PROCESSING FRAME INFORMATION
- FF REPRESENTATIVE DEGRADED AUDIO POWER SPECTRUM
- 2 CONVERSION FRAME DIVISION UNIT
- 5 CONVERSION UNIT
- 7 PROCESSING FRAME INFORMATION GENERATION UNIT
- 8 REPRESENTATIVE FREQUENCY REGION SIGNAL GENERATION UNIT
- 3 CONVERSION FRAME SYNTHESIS UNIT
- 6 INVERSE CONVERSION UNIT
- 10 NOISE SUPPRESSION PROCESSING UNIT
- 9 NOISE SUPPRESSION INFORMATION CALCULATION UNIT



WO 2009/038136 A1



IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE,
SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

添付公開書類：
— 国際調査報告書

明 細 書

雑音抑圧装置、その方法及びプログラム

技術分野

[0001] 本発明は、所望の音声信号に重畳されている雑音を抑圧する雑音抑圧装置、その方法及びプログラムに関する。

背景技術

[0002] 所望音声と背景雑音とから構成される入力信号の背景雑音を抑圧する装置として、雑音抑圧装置(以下、ノイズサプレッサという。)が知られている。ノイズサプレッサは、所望の音声信号に重畳されている雑音(ノイズ)を抑圧する装置である。一般的に、ノイズサプレッサは、周波数領域に変換した入力信号を用いて雑音成分のパワースペクトルを推定し、この推定したパワースペクトルを入力信号から差し引くことにより、所望の音声信号に混在する雑音を抑圧するように動作する。さらに、ノイズサプレッサは、雑音成分のパワースペクトルを継続的に推定することにより、非定常な雑音の抑圧にも適用することができる。ノイズサプレッサとしては、例えば、特許文献1に記載されている方式がある。

[0003] 図35を参照して、特許文献1に開示されたノイズサプレッサの構成を説明する。図35の入力端子901にサンプル値系列として供給された所望音声信号と雑音の混在する信号(以下、劣化音声信号という)は、変換フレーム分割部902において、定められたサンプル毎の変換フレームに分割される。変換フレームに分割された劣化音声信号は、変換部905において、フーリエ変換などの変換を施して複数の周波数成分に分割される。そして、変換部905は、周波数成分に分割された信号の振幅値を用いて求めた劣化音声信号のパワースペクトルを、雑音抑圧情報計算部907、雑音抑圧処理部908へ供給する。変換部905は、劣化音声信号位相を、逆変換部906に伝達する。雑音抑圧情報計算部907は、劣化音声パワースペクトルを用いて、抑圧係数を各周波数で算出し、雑音抑圧情報として生成し、雑音抑圧処理部908に出力する。抑圧係数は、劣化音声信号に乗算することによって雑音が抑圧された強調音声を求めるための係数である。雑音抑圧処理部908は、劣化音声パワースペクトルに雑音抑

圧情報である各周波数の抑圧係数を乗じ、強調音声パワースペクトルを求め、逆変換部906へ出力する。逆変換部906は、雑音抑圧処理部908から供給された強調音声パワースペクトルと変換部905から供給された劣化音声信号の位相を合わせて変換フレーム毎に逆変換を行い、変換フレームに分割されている強調音声信号を変換フレーム合成部903に出力する。変換フレーム合成部903は、変換フレームに分割されている強調音声信号を合成し、強調音声信号サンプルとして、出力端子4に出力する。これまでの処理ではパワースペクトルを用いた例を説明したが、代わりにその平方根に相当する振幅値を用いることができることは、広く知られている。

特許文献1:特開2002-204175号公報

発明の開示

発明が解決しようとする課題

[0004] しかしながら、図35を用いて説明した従来の構成では、変換フレーム毎に雑音抑圧情報を算出している。すなわち、従来の構成では、雑音抑圧情報を算出する処理フレーム長は、変換フレーム長と同じ長さのものを使用している。このため、変換フレーム長が長いと、入力信号が変換フレーム内の途中で変化した場合、入力信号の変化に追従できない。このとき、従来の構成では、精度の悪い雑音抑圧情報を算出し、出力信号の音質が劣化するという問題がある。一方で、変換フレーム長が短いと、入力信号の変化に追従することはできるが、雑音抑圧情報を算出する回数が増大し、演算量が増加するという問題がある。ノイズサプレッサに係る演算量が増加すると、ノイズサプレッサ以外の重要な機能があるときには、ノイズサプレッサ機能を組み込むことができない、あるいはノイズサプレッサ機能を組み込むことにより他の機能を組み込むことができないという問題がある。すなわち、従来の方法では、少ない演算量で高品質な雑音抑圧を実現できないという問題がある。

[0005] 本発明は上記課題に鑑みて発明されたものであって、その目的は、少ない演算量で高品質な雑音抑圧を実現することのできる雑音抑圧装置、その方法及びプログラムを提供することである。

課題を解決するための手段

[0006] 上記課題を解決する本発明は、入力信号を定められた第一のフレーム毎に周波数

領域信号に変換する変換手段と、前記第一のフレームと異なるように第二のフレームを生成するフレーム生成手段と、前記第二のフレームに含まれる第一のフレームの前記周波数領域信号から代表周波数領域信号を生成する代表周波数領域信号生成手段と、前記代表周波数領域信号に基づいて、前記第二のフレームの雑音抑圧の程度を求める雑音抑制程度算出手段とを有することを特徴とする雑音抑圧装置である。

[0007] 上記課題を解決する本発明は、入力信号を定められた第一のフレーム毎に周波数領域信号に変換する変換ステップと、前記第一のフレームと異なるように第二のフレームを生成するフレーム生成ステップと、前記第二のフレームに含まれる第一のフレームの前記周波数領域信号から代表周波数領域信号を生成する代表周波数領域信号生成ステップと、前記代表周波数領域信号に基づいて、前記第二のフレームの雑音抑圧の程度を求める雑音抑制程度算出ステップとを有することを特徴とする雑音抑圧方法である。

[0008] 上記課題を解決する本発明は、コンピュータに、入力信号を定められた第一のフレーム毎に周波数領域信号に変換する変換処理と、前記第一のフレームと異なるように第二のフレームを生成するフレーム生成処理と、

前記第二のフレームに含まれる第一のフレームの前記周波数領域信号から代表周波数領域信号を生成する代表周波数領域信号生成処理と、前記代表周波数領域信号に基づいて、前記第二のフレームの雑音抑圧の程度を求める雑音抑制程度算出処理とを実行させる雑音抑圧プログラムである。

発明の効果

[0009] 本発明の構成では、2つ以上の変換フレームを統合した処理フレーム毎に雑音抑圧情報を算出する。このため、本発明の構成によると、少ない演算量で高音質な雑音抑圧を実現できる。

図面の簡単な説明

[0010] [図1]本発明の最良の実施の形態を示すブロック図である。

[図2]図1に含まれる処理フレーム情報生成部の構成を示すブロック図である。

[図3]図2に含まれる時間グループ生成部における処理フレームの一例を示す図で

ある。

[図4]図2に含まれる周波数グループ生成部における統合周波数帯域の一例を示す図である。

[図5]図1に含まれる処理フレーム情報生成部の第2の構成を示すブロック図である。

[図6]図5に含まれる周波数グループ生成部における統合周波数帯域の一例を示す図である。

[図7]図1に含まれる雑音抑圧情報計算部の構成を示すブロック図である。

[図8]図7に含まれる雑音推定部の構成を示すブロック図である。

[図9]図8に含まれる推定雑音計算部の構成を示すブロック図である。

[図10]図9に含まれる更新判定部の構成を示すブロック図である。

[図11]図8に含まれる重みつき劣化音声計算部の構成を示すブロック図である。

[図12]図11に含まれる非線形処理部における非線形関数の一例を示す図である。

[図13]図7に含まれる雑音抑圧係数生成部の構成を示すブロック図である。

[図14]図13に含まれる推定先天的SNR計算部の構成を示すブロック図である。

[図15]図13に含まれる雑音抑圧係数計算部の構成を示すブロック図である。

[図16]図7に含まれる抑圧係数補正部の構成を示すブロック図である。

[図17]図1に含まれる雑音抑圧情報計算部の第2の構成を示すブロック図である。

[図18]図17に含まれる抑圧係数補正部の構成を示すブロック図である。

[図19]本発明の第2の実施の形態を示すブロック図である。

[図20]図19に含まれる雑音抑圧情報計算部の構成を示すブロック図である。

[図21]図20に含まれる雑音推定部の構成を示すブロック図である。

[図22]図19に含まれる雑音抑圧情報計算部の第2の構成を示すブロック図である。

[図23]本発明の第3の実施の形態を示すブロック図である。

[図24]図23に含まれる処理フレーム情報生成部の構成を示すブロック図である。

[図25]図23に含まれる処理フレーム情報生成部の第2の構成を示すブロック図である。

[図26]本発明の第4の実施の形態を示すブロック図である。

[図27]図26に含まれる処理フレーム情報生成部の構成を示すブロック図である。

[図28]本発明の第5の実施の形態を示すブロック図である。

[図29]図28に含まれる処理フレーム情報生成部の構成を示すブロック図である。

[図30]本発明の第6の実施の形態を示すブロック図である。

[図31]図30に含まれる処理フレーム情報生成部の構成を示すブロック図である。

[図32]本発明の第7の実施の形態を示すブロック図である。

[図33]本発明の第8の実施の形態を示すブロック図である。

[図34]本発明の第9の実施の形態を示すブロック図である。

[図35]従来構成を示すブロック図である。

[図36]時間グループ生成部の処理動作の一例を表す流れ図である。

符号の説明

- [0011] 1, 901 入力端子
2, 902 変換フレーム分割部
3, 903 変換フレーム合成部
4, 904 出力端子
5, 905 変換部
6, 906 逆変換部
7, 12, 13, 14, 15 処理フレーム情報生成部
8 代表周波数領域信号生成部
9, 11, 907 雑音抑圧情報計算部
10, 16, 908 雑音抑圧処理部
30 記録部
31 再生部
32 多重化部
33 分離部
50, 57 変換フレームエネルギー計算部
51, 55, 58, 59, 60 時間グループ生成部
52, 54, 56 周波数グループ生成部
53 周波数エネルギー計算部

300, 301 雑音推定部
310 推定雑音計算部
320 重み付き劣化音声計算部
330, 331, 480 カウンタ
400 更新判定部
410 レジスタ長記憶部
420, 3201 推定雑音記憶部
430, 1595 スイッチ
440 シフトレジスタ
450, 6208 加算器
460 最小値選択部
470 除算部
601, 602 雑音抑圧係数生成部
610 後天的SNR計算部
620 推定先天的SNR計算部
630 雑音抑圧係数計算部
640 音声非存在確率記憶部
660, 1597, 3203, 6204, 6205 乗算器
670 音声存在確率計算部
680 仮出力SNR計算部
1000 コンピュータ
1501, 1502 抑圧係数補正部
1591, 6511 最大値選択部
1592 抑圧係数下限値記憶部
1593 閾値記憶部
1594, 4002, 4004 比較部
1596 修正値記憶部
3202 SNR計算部

3204 非線形処理部
4001 論理和計算部
4003, 4005 閾値記憶部
4006 閾値計算部
6201 値域限定処理部
6202 後天的SNR記憶部
6203 抑圧係数記憶部
6206 重み記憶部
6207 重み付き加算部
6301 MMSE STSA ゲイン関数値計算部
6302 一般化尤度比計算部
6303 抑圧係数計算部
6512 抑圧係数下限値計算部

発明を実施するための最良の形態

- [0012] 本発明の雑音抑圧装置の実施の形態について図面を参照して詳細に説明する。
- [0013] 図1を参照して、本発明の最良の実施の形態の構成を説明する。本発明の雑音抑圧装置は、入力端子1、変換フレーム分割部2、変換フレーム合成部3、出力端子4、変換部5、逆変換部6、処理フレーム情報生成部7、代表周波数領域信号生成部8、雑音抑圧情報計算部9、及び雑音抑圧処理部10から構成される。
- [0014] 劣化音声信号である入力信号が、サンプル値系列として入力端子1に供給される。入力信号サンプルは、変換フレーム分割部2に供給され、定められた変換フレーム長に分割される。変換フレーム分割部2は、第 n 変換フレームの入力信号サンプルを、変換部5に出力する。変換部5は、第 n 変換フレームの入力信号サンプルを周波数領域の信号である劣化音声スペクトル $Y_n(k)$ に変換する。ここで、 n は変換フレームの時間方向のインデックスを表す。 k は周波数方向のインデックスを表し、第 n 変換フレームの入力信号サンプルを K 個の周波数帯域に分割するものとする($0 \leq k < K$)。変換部5は、劣化音声スペクトル $Y_n(k)$ を位相と振幅に分離し、位相である $\arg Y_n(k)$ は逆変換部6に、劣化音声パワースペクトル $|Y_n(k)|^2$ は、処理フレーム情報生成部7、代表周

波数領域信号生成部8、雑音抑圧処理部10に出力する。

- [0015] 変換部5は、第n変換フレームの入力信号サンプルを劣化音声スペクトル $Y_n(k)$ へ変換する方法として、変換フレームに分割された入力信号サンプルに対して周波数変換を適用している。周波数変換の例としては、フーリエ変換、コサイン変換、KL(カーネンレーベ)変換などが知られている。これらの変換の具体的な演算に関連する技術及びその性質は、非特許文献1(1990年、「デジタル・コーディング・オブ・ウェーブフォームズ」、プレントイス・ホール(DIGITAL CODING OF WAVEFORMS, PRINCIPLES AND APPLICATIONS TO SPEECH AND VIDEO, PRENTICE-HALL, 1990.))に開示されている。また、アダマール変換、ハール変換、ウェーブレット変換などの他の変換を用いることができることは広く知られている。
- [0016] 変換部5は、当該変換フレームの入力信号サンプルを窓関数Wで重み付けした結果に対して、前述の変換を適用することができる。このような窓関数としては、ハミング、ハニング(ハン)、ケイザー、ブラックマンなどの窓関数が知られている。また、さらに複雑な窓関数を用いることもできる。これらの窓関数に関連する技術は、非特許文献2(1975年、「デジタル・シグナル・プロセッシング」、プレントイス・ホール(DIGITAL SIGNAL PROCESSING, PRENTICE-HALL, 1975.))及び非特許文献3(1993年、「マルチレートシステムズ・アンド・フィルタバンクズ」、プレントイス・ホール(MULTIRATE SYSTEMS AND FILTER BANKS, PRENTICE-HALL, 1993.))に開示されている。さらに、連続する2変換フレーム以上の一部を重ね合わせ(オーバーラップ)て窓がけすることも広く行なわれている。この場合、重ね合わせて窓がけした信号に対して、前述の周波数変換を用いる。オーバーラップを有するブロック化と変換に関連する技術は、非特許文献2に開示されている。
- [0017] さらに変換部5は、帯域分割フィルタバンクで構成して劣化音声スペクトル $Y_n(k)$ を算出してもよい。帯域分割フィルタバンクは、複数の帯域通過フィルタから構成される。帯域分割フィルタバンクの各周波数帯域は等間隔であってもよいし、不等間隔であってもよい。不等間隔に帯域分割することによって、低域では狭帯域に分割して時間分解能を低く、高域では広い帯域に分割して時間分解能を高くすることができる。不等間隔分割の代表例には、低域に向かって帯域が逐次半分になるオクターブ分割

や人間の聴覚特性に対応した臨界帯域分割などがある。等間隔の周波数帯域に分割した後、低域の周波数帯域の周波数分解能を向上させるために、低域のみをさらに帯域分割するハイブリッドフィルタバンクを用いてもよい。帯域分割フィルタバンクとその設計法に関連する技術は、非特許文献3に開示されている。

[0018] 処理フレーム情報生成部7は、劣化音声パワースペクトルから、後述する代表劣化音声パワースペクトルを生成するための処理フレーム情報を計算する。処理フレーム情報には、複数の劣化音声パワースペクトルを時間方向および周波数方向に統合するための情報を含む。図2を参照して、図1に含まれる処理フレーム情報生成部7について詳細に説明する。処理フレーム情報生成部7は、変換フレームエネルギー計算部50、時間グループ生成部51及び周波数グループ生成部52から構成される。

[0019] 変換フレームエネルギー計算部50は、劣化音声パワースペクトル $|Y_n(k)|^2$ から当該変換フレームの変換フレームエネルギー $E(n)$ を求め、時間グループ生成部51に出力する。変換フレームエネルギー $E(n)$ は、

[0020] [数1]

$$E(n) = \sum_{k=0}^{K-1} |Y_n(k)|^2$$

となる。

[0021] ここでは、全周波数帯域の劣化音声パワースペクトルのエネルギー和を変換フレームエネルギーとしている。しかし、一部の周波数帯域のみの劣化音声パワースペクトルから変換フレームエネルギーを算出してもよい。たとえば、音声信号のパワーが集中する帯域のみの劣化音声パワースペクトルから変換フレームエネルギーを算出してもよい。これにより、後述する処理フレーム生成を高品位に行うことができる。また、変換フレームエネルギーの計算に低周波数帯域の信号を使わずに算出した場合、低域に集中しやすい雑音成分の影響を除去することができる。

[0022] さらに、劣化音声パワースペクトルを周波数方向に重み付けして、重み付けした値

の和を変換フレームエネルギーとしてを用いてもよい。他にも、算出した変換フレームエネルギーを時間方向に平滑化してもよい。

[0023] ここで、算出した変換フレームエネルギーを聴覚上の特性に合わせて修正することもできる。例えば、人間の聴覚特性として、音の強さの知覚は対数に比例していることが知られている。この特性を用いて、エネルギーを対数化したものを変換フレームエネルギーとすることもできる。単なる対数ばかりでなく、より複雑な関数や多項式などを用いて、変換フレームエネルギーを修正することも可能である。対数を近似する多項式は、これらの例の一つであり、演算量削減に貢献する。

[0024] 時間グループ生成部51は、変換フレームエネルギーに基づいて、後述する代表劣化音声パワースペクトルを生成するための処理フレームの区切り位置を定める。定められた処理フレーム区切り位置に基づいて生成した処理フレームを周波数グループ生成部52へ出力する。処理フレームの区切り位置を定める方法として、変換フレームエネルギーの変化に基づいて処理フレームの区切り位置を定める方法がある。

[0025] 図3を参照して、変換フレームエネルギーの変化例を説明する。図3において、 n_{L-1} , n_L , n_{L+1} で変換フレームエネルギーが大きく変化している。これらの箇所では処理フレームが分割されるように処理フレームの区切り位置を定めると、第L-1処理フレームの区切り位置は、 n_{L-1} および n_L となり、第L処理フレームの区切り位置は、 n_L および n_{L+1} となる。この結果、第L-1処理フレームは、第 n_{L-1} 変換フレームから第 $n_L - 1$ 変換フレームを統合して処理フレームが生成されている。第L-1処理フレームの処理フレーム長は $n_L - n_{L-1}$ である。一方、第L処理フレームは、第 n_L 変換フレームから第 $n_{L+1} - 1$ 変換フレームを統合して処理フレームが生成されている。その第L処理フレーム長は $n_{L+1} - n_L$ となる。

[0026] 変換フレームエネルギーが大きく変化する箇所の検出には、たとえば、あらかじめ定めた閾値 TH_A を用いて

[0027] [数2]

$$E(n_L) - E(n_L - 1) > TH_A$$

を満たす場合に、変換フレームエネルギーが大きく変化したと判定する方法がある。この方法の場合、 $n=n_L$ で処理フレームが分割されるように処理フレームの区切り位置を定める。このとき、閾値 TH_A を変えることも可能である。たとえば、数2を満たす割合が、ある一定区間内で均一になるように、変換フレームエネルギーの平均や分散の値に基づいて閾値 TH_A を適応的に変更する。このようにすることにより、後述する雑音抑圧情報の演算回数のばらつきを少なくすることができる。

[0028] 処理フレームの区切り位置を生成する他の方法として、隣接する2つの変換フレームエネルギーのみから変化量を算出するのではなく、複数の変換フレームエネルギーを用いて変化量を算出し、処理フレームの区切り位置を生成する方法がある。例えば、3つの変換フレームエネルギーを用いて、次の条件式を満たすとき、 $n=n_L$ で処理フレームが分割されるように処理フレームの区切り位置を定めることができる。

[0029] [数3]

$$(E(n_L) - E(n_L - 1)) \cdot (E(n_L) - E(n_L - 2)) > TH_B$$

ここで、 TH_B は閾値である。このとき、閾値 TH_B を変えることも可能である。たとえば、[数3]を満たす割合が、ある一定区間内で均一になるように、変換フレームエネルギーの平均や分散の値に基づいて閾値 TH_B を適応的に変更する。このようにすることにより、後述する雑音抑圧情報の演算回数のばらつきを少なくすることができる。

[0030] 処理フレームの区切り位置を定める他の方法として、当該処理フレーム内に含まれる変換フレームエネルギーの最小値と最大値が定められた閾値以下になるように処理フレームの区切り位置を定める方法がある。この場合、当該処理フレーム内に含まれる信号が同程度のエネルギーを有していることになり、後述する雑音抑圧情報を高品位に算出することができる。また、変換フレームエネルギーが大きく変化した箇所から、固定の処理フレーム長になるように処理フレームの区切り位置を生成してもよい。この場合、エネルギー変化の判定回数を削減することができるため、演算量を削減できる。

[0031] 上記説明では、変換フレーム毎に変換フレームエネルギーを算出し、処理フレームの

区切り位置を生成する方法について説明した。上記方法に対し、複数の変換フレームを統合した単位で、変換フレームエネルギーを算出し、算出した変換フレームエネルギーに基づいて処理フレームの区切り位置を生成することも可能である。この場合、変換フレーム毎に変換フレームエネルギーを算出する必要がないので、時間グループ生成部51の演算量を削減することができる。また、周波数帯域毎に信号の変化を分析し、処理フレームの区切り位置を定めることもできる。この結果、周波数帯域ごとに定めた重要度を反映させることができる。たとえば、音声信号が含まれる帯域の重要度を大きくすることで、当該帯域の信号の変化を反映しやすくすることができる。

[0032] 処理フレームの区切り位置を定める指標として、変換フレームエネルギー以外の劣化音声スペクトルの特性を用いても良い。たとえば、心理聴覚エントロピーなどの指標に基づいて区切り位置を定めることができる。すなわち、人の聴覚特性である大きい音の近傍の小さい音が聞こえないという心理聴覚マスキングなどを積極的に用いる方法である。心理聴覚マスキングを用いて、人に聞こえる音の成分が変化する箇所処理フレームが分割されるように処理フレームの区切り位置を定める。この方法により、人の聴覚特性に基づいた処理フレームを生成することができ、後述する雑音抑圧情報を高品位に算出することができる。

[0033] 処理フレームの区切り位置を定める際には、上記いずれかの方法のみを用いるのではなく、組み合わせて用いることができるのは明らかである。

[0034] ここで、図36の流れ図を参照して、時間グループ生成部51の処理動作の一例を説明する。

[0035] 時間グループ生成部51は、ある定められた一定区間内のN個の変換フレームにおいて、変換フレームエネルギーの分散を計算する(S001)。その後、上記一定区間内のN個の変換フレームにおいて、時間グループ生成部51は、上述した数2または数3を満たしているか判定する(S002)。1つでも満たしている変換フレームが存在する場合、処理は、S007へ進む。逆に、上述した数2または数3を満たしている変換フレームが1つもない場合、処理は、S003へ進む。

[0036] S003では、時間グループ生成部51は、算出した分散値が閾値Thr1より大きいかどうか判定し、分散値が閾値Thr1より大きい場合は処理をS007へ進める。一方で、分散

値が閾値Thr1より小さい場合、処理は、S004へ進む。S004では、時間グループ生成部51は、算出した分散値が閾値Thr2より大きいかどうかを判定し、分散値が閾値Thr2より小さい場合、処理をS005へ進める。

[0037] S005では、当該N個の変換フレームを1つの処理フレームとする。ここで、 n_0 および n_1 は、処理フレームの区切り位置を表し、Kosuは当該N個の変換フレームからいくつの処理フレームを生成したかを表す。一方で、S004において分散値が閾値Thr2より大きい場合、処理はS006へ進む。S006では、当該N個の変換フレームを2つの処理フレームとする。このとき、2つの処理フレームの処理フレーム長が同じになるように区切り位置を設定する。すなわち $n_1 = N/2$ となる。

[0038] 続いて、S007以降の動作について説明する。S007において、時間グループ生成部51は、必要な変数を初期化後、当該N個の変換フレームを $n=0$ から $N-1$ まで順番に調査し、処理フレームの区切り位置となるかどうかを判定していく。次に、S008において、時間グループ生成部51は、当該処理フレームに含まれる変換エネルギーの最小値と最大値の差の絶対値が定められた閾値より大きいかどうかを判定する。定められた閾値より大きい場合、処理はS010へ進み、定められた閾値より小さい場合、処理はS009へ進む。続いて、S009では、時間グループ生成部51は、変換フレーム n において、上述した数2または数3を満たしているかを判定する。S009において、上述した数2または数3を満たしている場合、処理はS010へ進む。一方で、満たしていない場合、処理はS011へ進む。S010では、時間グループ生成部51は、変換フレーム n において処理フレームが分割されるように処理フレームの区切り位置を定め、処理フレームの個数を1つ増やし、処理をS011へ進める。S011では、時間グループ生成部51は、変換フレーム $N-1$ まで調査したかどうかを判定し、調査すべき変換フレームが残っている場合は、 $n=n+1$ (S012)とし、S008へ戻る。当該N変換フレーム全てを調査した場合、処理フレームの生成を終了する。

[0039] 以上で、図36を参照した時間グループ生成部51の処理動作の一例の説明を終える。

[0040] 周波数グループ生成部52は、時間グループ生成部51より供給された各処理フレームに対して、周波数帯域を統合し、後述する代表劣化音声パワースペクトルを計算

するための統合周波数帯域の区切り位置を定める。その後、処理フレームの区切り位置と統合周波数帯域の区切り位置とを処理フレーム情報として代表周波数領域信号生成部8に出力する。

[0041] 図4を参照して、周波数帯域を統合する様子を説明する。破線で囲まれた一つ一つの格子が1つの劣化音声パワースペクトルを表す。横軸は時間方向を表し、横軸の1マスが1変換フレームを表す。縦軸は周波数方向を表し、縦軸の1マスが変換部5において変換された1つの周波数帯域を表す。前述の時間グループ生成部51の処理は、図4の横軸である時間方向のマスで統合する区切りを決定することに相当する。図4には、時間グループ生成部51により生成した第L-1処理フレームおよび第L処理フレームが表されている。第L-1処理フレームおよび第L処理フレームは、 $n=n_{L-1}, n_L, n_{L+1}$ で処理フレームを区切ることで生成されたものである。また、周波数グループ生成部52での処理は、図4の縦軸である周波数方向のマスで統合することに相当する。図4では、K個の周波数帯域をM個の周波数帯域に統合した場合を表している。第L処理フレームの周波数方向の区切り位置を、 $k_{L,p}$ ($p=0,1,\dots,M$)、 $k_{L,0}=0, k_{L,M}=K$ とする。第L処理フレームの処理フレーム情報は、時間方向の処理フレームの区切り位置($n=n_L, n_{L+1}$)と、周波数方向の統合周波数帯域の区切り位置($k=k_{L,0}, \dots, k_{L,M}$)とから構成される。

[0042] このとき、高周波領域では低周波領域に比べてより多くの帯域を一つに統合してもよい。すなわち、高周波領域成分ほどたくさんの周波数成分を一つに統合することになり、不等分割されることになる。このような不等分割の例としては、高周波領域側に向かって2のべき乗で帯域が広がるオクターブ分割や人間の聴覚特性に基づいて帯域分割された臨界帯域に従った分割などが知られている。特に、臨界帯域に従った帯域分割は、人間の聴覚特性と整合性が高いために、広く用いられている。周波数帯域を統合するときに、臨界帯域よりも細かいグループに統合することによって、雑音抑圧特性の劣化を防ぐこともできる。

[0043] 次に、図5を参照して、処理フレーム情報生成部7の第2の構成例を詳細に説明する。図2の処理フレーム情報生成部7と比較すると、新たに周波数エネルギー計算部53を備え、周波数グループ生成部52が周波数グループ生成部54に置換されていること

を特徴とする。以下、本構成例の特徴である周波数エネルギー計算部53と周波数グループ生成部54について詳細に説明する。

[0044] 周波数エネルギー計算部53は、劣化音声パワースペクトルと処理フレームとから、当該処理フレームにおける同一周波数帯域の劣化音声パワースペクトルのエネルギー和である周波数エネルギー $Ef_L(k)$ を求める。周波数エネルギー計算部53は、周波数エネルギー $Ef_L(k)$ を周波数グループ生成部54に出力する。すなわち、処理フレームLの周波数エネルギー $Ef_L(k)$ は、

[0045] [数4]

$$Ef_L(k) = \sum_{n=n_L}^{n_L-1} |Y_n(k)|^2$$

となる。

[0046] 周波数グループ生成部54は、時間グループ生成部51から供給された処理フレームと周波数エネルギー計算部53から供給された周波数エネルギー $Ef_L(k)$ に基づいて、処理フレーム単位で劣化音声パワースペクトルの特性が似ている周波数帯域を統合する。これにより、周波数グループ生成部54は、統合周波数帯域の区切り位置を定める。

[0047] 図6を参照して、各処理フレームにおける周波数帯域を統合する様子を説明する。横軸、縦軸は、図4と同一である。第L-1処理フレームにおいてK個の周波数帯域を M_{L-1} 個に、第L処理フレームにおいてK個の周波数帯域を M_L 個の周波数帯域に統合した場合を表している。処理フレームLの周波数方向の区切り位置を、 $k_{L,p}$ ($p=0,1,\dots,M_L$), $k_{L,0}=0, k_{L,M_L}=K$ とする。処理フレーム情報は、時間方向の区切り位置である処理フレームの区切り位置と周波数方向の区切り位置である統合周波数帯域の区切り位置とから構成される。

[0048] 周波数帯域の統合は、周波数エネルギーの変化が大きい箇所で統合周波数帯域が分割されるように、統合周波数帯域の区切り位置を定める。たとえば、時間グループ生成部51で説明したエネルギー変化に基づく方法を周波数方向に適用することにより

周波数帯域を統合してもよい。このように構成することにより、各処理フレームで最適な周波数帯域の統合を実現できる。このため、信号の変化が少ない場合には、不必要に多くの帯域に統合することを抑制でき、演算量を削減できる。

[0049] 以上で、処理フレーム情報生成部7の第2の構成例の説明を終える。

[0050] 処理フレーム情報生成部7を上記のように構成することにより、複数の変換フレームを統合した処理フレームを生成することができる。このとき、処理フレームに含まれる各変換フレームについて劣化音声パワースペクトルの特性が似ているため、当該変換フレーム毎に算出したそれぞれの雑音抑圧情報が似た値となる。雑音抑圧情報については後述する。このため、変換フレーム毎に算出した雑音抑圧情報による雑音抑圧と、処理フレーム毎に算出した雑音抑圧情報による雑音抑圧とで効果の差はほとんど生じない。これにより、処理フレーム毎に算出された雑音抑圧情報を用いても、雑音抑圧の効果を低下させることはない。従って、処理フレーム毎に雑音抑圧情報を算出することにより演算量を削減しても、最終的な雑音抑圧に影響を与えることはない。

以上で、処理フレーム情報生成部7の説明を終える。

[0051] 代表周波数領域信号生成部8は、処理フレーム情報と劣化音声パワースペクトルを用いて、代表劣化音声パワースペクトルを生成する。そして、代表周波数領域信号生成部8は、代表劣化音声パワースペクトルを雑音抑圧情報計算部9へ出力する。代表劣化音声パワースペクトルを生成する方法として、当該処理フレームおよび当該統合周波数領域に含まれる劣化音声パワースペクトルの平均を用いる方法がある。この場合、第L処理フレームの代表劣化音声パワースペクトル $|Z_L(m)|^2 (m=0, \dots, M_L-1)$ は、

[0052] [数5]

$$|Z_L(m)|^2 = \frac{\sum_{k=k_{L,m}}^{k_{L,m+1}-1} \sum_{n=n_L}^{n_{L+1}-1} |Y_n(k)|^2}{(k_{L,m+1} - k_{L,m}) \cdot (n_{L+1} - n_L)}$$

となる。すなわち、図4および図6において、灰色で囲まれている格子1つにつき、1つの値を計算していることに相当する。

[0053] また、全ての劣化音声パワースペクトルの平均値を用いる方法以外にも、大きな劣化音声パワースペクトルや小さな劣化音声パワースペクトルを除いて平均値を求める方法がある。このようにすると、突発的な劣化音声パワースペクトルを除くことができるため、代表劣化音声パワースペクトルが安定し、後述する雑音抑圧の程度を高品位に算出できる。

[0054] 他にも、平均値を用いるのではなく、特定の劣化音声パワースペクトルを代表劣化音声パワースペクトルとする方法もある。たとえば、当該処理フレームおよび当該統合周波数領域に含まれる劣化音声パワースペクトルの最大値を代表劣化音声パワースペクトルとすると、後述する雑音抑圧情報の計算において雑音成分を大きく推定することになる。この場合、雑音抑圧された強調音声に含まれる残留雑音を小さくすることができる。一方、当該処理フレームおよび当該統合周波数領域に含まれる劣化音声パワースペクトルの最小値を代表劣化音声パワースペクトルとすると、後述する雑音抑圧情報の計算において雑音成分を小さく推定することになる。この場合、雑音抑圧された強調音声の歪を小さくすることができる。

[0055] 雑音抑圧情報計算部9は、代表劣化音声パワースペクトル毎に1つの雑音抑圧の程度を表す雑音抑圧情報を求める。そして、雑音抑圧情報計算部9は、雑音抑圧情報を雑音抑圧処理部10に出力する。すなわち、雑音抑圧情報計算部9は、複数の劣化音声パワースペクトルに対して共通の雑音抑圧情報を算出する。これは、図4および図6において、灰色で囲まれている格子1つにつき、1つの雑音抑圧情報 $C_L(m)$

$m=0, \dots, M_L-1$)を計算していることに相当する。

- [0056] 図7を参照して、雑音抑圧情報計算部9の第1の構成例について詳細に説明する。雑音抑圧情報計算部9は、雑音推定部300、雑音抑圧係数生成部601及び抑圧係数補正部1501から構成される。
- [0057] 雑音推定部300は、代表劣化音声パワースペクトルに基づいて劣化音声に含まれる雑音成分のエネルギーを推定する。雑音推定部300は、推定した雑音成分のエネルギーを推定雑音パワースペクトルとして、雑音抑圧係数生成部601へ出力する。雑音抑圧係数生成部601は、代表劣化音声パワースペクトルと推定雑音パワースペクトルと後述する補正抑圧係数に基づいて、抑圧係数を求め、入力信号に含まれる音声と雑音の比を表す先天的SNRを推定する。推定先天的SNRについては、後述する。雑音抑圧係数生成部601は、抑圧係数と推定先天的SNRとを抑圧係数補正部1501へ出力する。抑圧係数補正部1501は、推定された先天的SNRに基づいて入力された抑圧係数を補正し、補正抑圧係数を求める。抑圧係数補正部1501は、補正抑圧係数を雑音抑圧情報として出力すると同時に、雑音抑圧係数生成部601へ出力する。
- [0058] 図8を参照して、図7に含まれる雑音抑圧推定部300の構成例を説明する。雑音推定部300は、推定雑音計算部310、重みつき劣化音声計算部320及びカウンタ330から構成される。雑音推定部300に入力された代表劣化音声パワースペクトルは、推定雑音計算部310、及び重み付き劣化音声計算部320に入力される。重み付き劣化音声計算部320は、入力された代表劣化音声パワースペクトルと推定雑音パワースペクトルを用いて重み付き劣化音声パワースペクトルを計算する。重み付き劣化音声計算部320は、重み付き劣化音声パワースペクトルを推定雑音計算部310に出力する。推定雑音計算部310は、代表劣化音声パワースペクトル、重み付き劣化音声パワースペクトル、及びカウンタ330から入力されるカウント値を用いて雑音のパワースペクトルを推定する。推定雑音計算部310は、雑音推定部300の出力として推定雑音パワースペクトルを出力する。さらに、推定雑音計算部310は、推定雑音パワースペクトルを重み付き劣化音声計算部320に出力する。カウンタ330は、カウント値を出力する。カウント値の初期値は0に設定されている。カウンタ330は、処理フレーム毎にカウント値を1ずつ増加していく。

- [0059] 図9を参照して、図8に含まれる推定雑音計算部310の構成を詳細に説明する。推定雑音計算部310は、更新判定部400、レジスタ長記憶部410、推定雑音記憶部420、スイッチ430、シフトレジスタ440、加算器450、最小値選択部460、除算部470及びカウンタ480から構成される。スイッチ430には、重み付き劣化音声パワースペクトルが入力されている。スイッチ430が回路を閉じたときに、重み付き劣化音声パワースペクトルは、シフトレジスタ440に入力される。シフトレジスタ440は、更新判定部400から入力される制御信号に応じて、内部レジスタの記憶値を隣接レジスタにシフトする。シフトレジスタ長は、後述するレジスタ長記憶部410に記憶されている値に等しい。シフトレジスタ440の全レジスタ出力は、加算器450に出力される。加算器450は、入力された全レジスタ出力を加算する。加算器450は、加算結果を除算部470に出力する。
- [0060] 一方、更新判定部400には、カウント値、代表劣化音声パワースペクトル及び推定雑音パワースペクトルが入力されている。更新判定部400は、1又は0の信号をカウンタ480、スイッチ430、及びシフトレジスタ440に出力する。更新判定部400は、入力されるカウント値が予め設定された値に到達するまでは常に1を出力する。また、更新判定部400は、カウント値が予め設定された値に到達した後は入力された劣化音声信号が雑音であると判定されたときに1を、それ以外るときに0を出力する。スイッチ430は、更新判定部400から入力された信号が1のときに回路を閉じ、0のときに開く。カウンタ480は、更新判定部400から入力された信号が1のときにカウント値を増加し、0のときには変更しない。シフトレジスタ440は、更新判定部400から入力された信号が1のときにスイッチ430から入力される信号サンプルを1サンプル取り込む。さらに、シフトレジスタ440は、1サンプルを取り込むと同時に、内部レジスタの記憶値を隣接レジスタにシフトする。最小値選択部460には、カウンタ480の出力とレジスタ長記憶部410の出力が入力されている。
- [0061] 最小値選択部460は、入力されたカウント値とレジスタ長のうち、小さい方を選択して、除算部470に出力する。除算部470は、加算器450から入力された代表劣化音声パワースペクトルの加算値をカウント値又はレジスタ長の小さい方の値で除算する。除算部470は、除算された商を推定雑音パワースペクトル $\lambda_L(m)$ として出力する。 $B_l(m)$ ($l=0, 1, \dots, P-1$)をシフトレジスタ440に保存されている重みつき劣化音声パワース

ペクトルのサンプル値とすると、 $\lambda_L(m)$ は、

[0062] [数6]

$$\lambda_L(m) = \frac{1}{P} \sum_{l=0}^{P-1} B_l(m)$$

で与えられる。ただし、Pはカウント値とレジスタ長のうち、小さい方の値である。カウント値はゼロから始まって単調に増加するので、最初はカウント値で除算が行なわれる。カウント値がレジスタ長より大きくなった後には、レジスタ長で除算が行なわれる。レジスタ長で除算が行なわれることは、シフトレジスタに格納された値の平均値を求めることになる。最初は、シフトレジスタ440に十分多くの値が記憶されていないために、実際に値が記憶されているレジスタの数で除算する。実際に値が記憶されているレジスタの数は、カウント値がレジスタ長より小さいときはカウント値に等しく、カウント値がレジスタ長より大きくなると、レジスタ長と等しくなる。

[0063] 図10を参照して、図9に含まれる更新判定部400の構成を詳細に説明する。更新判定部400は、論理和計算部4001、比較部4004、4002、閾値記憶部4005、4003、閾値計算部4006から構成される。図8のカウンタ330から入力されるカウント値は、比較部4002に入力される。閾値記憶部4003の出力である閾値は、比較部4002に入力される。比較部4002は、入力されたカウント値と閾値を比較し、カウント値が閾値より小さいときに1を、カウント値が閾値より大きいときに0を、論理和計算部4001に出力する。一方、閾値計算部4006は、図9の推定雑音記憶部420から供給される推定雑音パワースペクトルに応じた値を計算し、閾値として閾値記憶部4005に出力する。最も簡単な閾値の計算方法は、推定雑音パワースペクトルを定数倍する方法である。その他に、高次多項式や非線形関数を用いて閾値を計算する方法もある。閾値記憶部4005は、閾値計算部4006から出力された閾値を記憶する。そして、閾値計算部4006は、1処理フレーム前に記憶された閾値を比較部4004へ出力する。比較部4004は、閾値記憶部4005から入力される閾値と図1の代表周波数領域信号生成部8から入力

される代表劣化音声パワースペクトルを比較する。このとき、比較部4004は、代表劣化音声パワースペクトルが閾値よりも小さければ1を、大きければ0を論理和計算部4001に出力する。すなわち、推定雑音パワースペクトルの大きさをもとに、劣化音声信号が雑音であるか否かを判別している。論理和計算部4001は、比較部4202の出力値と比較部4204の出力値との論理和を計算する。そして、論理和計算部4001は、計算結果を図9のスイッチ430、シフトレジスタ440及びカウンタ480に出力する。このように、初期状態や無音区間だけでなく、有音区間でも劣化音声パワーが小さい場合には、更新判定部400は1を出力する。すなわち、有音区間でも劣化音声パワーが小さい場合には、推定雑音の更新が行われる。閾値の計算は各周波数で行われるため、各周波数で推定雑音の更新を行うことができる。

[0064] 図11を参照して、雑音推定部300に含まれる重み付き劣化音声計算部320の構成を詳細に説明する。重み付き劣化音声計算部320は、推定雑音記憶部3201、SNR計算部3202、非線形処理部3204、及び乗算器3203から構成される。推定雑音記憶部3201は、図8の推定雑音計算部310から入力される推定雑音パワースペクトルを記憶する。さらに、推定雑音記憶部3201は、1処理フレーム前に記憶された推定雑音パワースペクトルをSNR計算部3202へ出力する。SNR計算部3202は、推定雑音記憶部3201から入力される推定雑音パワースペクトルと図1の代表周波数領域信号生成部8から入力される代表劣化音声パワースペクトルとを用いてSNRを統合周波数帯域毎に求め、非線形処理部3204に出力する。具体的には、次式に従って、供給された代表劣化音声パワースペクトルを推定雑音パワースペクトルで除算して第L処理フレームのSNR $\hat{\gamma}_L(m)$ を求める。

[0065] [数7]

$$\hat{\gamma}_L(m) = \frac{|Z_L(m)|^2}{\lambda_{L-1}(m)}$$

ここで、 $\lambda_{L-1}(m)$ は1処理フレーム前に記憶された推定雑音パワースペクトルである。

[0066] 非線形処理部3204は、SNR計算部 3202 から入力されるSNRを用いて重み係数ベクトルを計算する。そして、非線形処理部3204は、重み係数ベクトルを乗算器3203に出力する。乗算器3203は、図1の代表周波数領域信号生成部8からから入力される代表劣化音声パワースペクトルと、非線形処理部3204から入力される重み係数ベクトルとの積を周波数帯域毎に計算する。そして、乗算器3203は、重み付き劣化音声パワースペクトルを図8の推定雑音計算部310に出力する。

[0067] 非線形処理部3204は、多重化された入力値それぞれに応じた実数値を出力することが出来る非線形関数を有する。図12に、非線形関数の例を示す。 f_1 を入力値としたとき、図12に示される非線形関数の出力値 f_2 は、

[0068] [数8]

$$f_2 = \begin{cases} 1, & f_1 \leq a \\ \frac{f_1 - b}{a - b}, & a < f_1 \leq b \\ 0, & b < f_1 \end{cases}$$

で与えられる。但し、 a と b は任意の実数である。

[0069] 非線形処理部3204は、SNR計算部3202から入力されるSNRを、非線形関数によって処理して重み係数を求め、乗算器3203に出力する。すなわち、非線形処理部3204はSNRに応じた1から0までの重み係数を出力する。SNRが小さい時は1を、大きい時は0を出力する。

[0070] 図11の乗算器3203は、代表劣化音声パワースペクトルと重み係数とを乗算する。重み係数は、SNRに応じた値になっている。つまり、SNRが大きい程すなわち劣化音声に含まれる音声成分が大きい程、重み係数の値は小さくなる。推定雑音の更新に

は一般に代表劣化音声パワースペクトルが用いられる。しかし、本発明では、推定雑音の更新に用いる代表劣化音声パワースペクトルに対して、SNRに応じた重みづけを行う。このことにより、代表劣化音声パワースペクトルに含まれる音声成分の影響を小さくすることができ、より精度の高い雑音推定を行うことができる。なお、重み係数の計算に非線形関数を用いた例を示したが、非線形関数以外にも線形関数や高次多項式など、他の形で表されるSNRの関数を用いてもよい。

[0071] 以上で、雑音推定部300の説明を終える。

[0072] 続いて、図13を参照して、図7の雑音抑圧係数生成部601を詳細に説明する。

[0073] 雑音抑圧係数生成部601は、後天的SNR計算部610、推定先天的SNR計算部620、雑音抑圧係数計算部630、音声非存在確率記憶部640から構成される。後天的SNR計算部610は、入力された代表劣化音声パワースペクトルと推定雑音パワースペクトルとを用いて統合周波数帯域毎にSNRを計算する。そして、後天的SNR計算部610は、計算結果を後天的SNRとして、推定先天的SNR計算部620と雑音抑圧係数計算部630とに出力する。推定先天的SNR計算部620は、入力された後天的SNR及び抑圧係数補正部650から入力された補正抑圧係数を用いて先天的SNRを推定する。推定先天的SNR計算部620は、推定先天的SNRを雑音抑圧係数補正部1501に出力する。さらに、推定先天的SNR計算部620は、推定先天的SNRを雑音抑圧係数計算部630に出力する。

[0074] 雑音抑圧係数計算部630は、入力された後天的SNR、推定先天的SNR及び音声非存在確率記憶部640から入力される音声非存在確率を用いて抑圧係数を生成する。音声非存在確率は、入力信号に音声が含まれていないあらかじめ定めた確率を表す。そして、雑音抑圧係数計算部630は、抑圧係数を出力する。

[0075] 図14を参照して、図13に含まれる推定先天的SNR計算部620の構成を詳細に説明する。推定先天的SNR計算部620は、値域限定処理部6201、後天的SNR記憶部6202、抑圧係数記憶部6203、乗算器6204、6205、重み記憶部6206、重み付き加算部6207、加算器6208から構成される。

[0076] 図13の後天的SNR計算部610から入力される後天的SNR $\gamma_L(m)$ ($m=0, 1, \dots, M_L-1$) は、後天的SNR記憶部6202と加算器6208に入力される。後天的SNR記憶部6202は、

第L処理フレームにおける後天的SNR $\gamma_L(m)$ を記憶する。これと同時に、後天的SNR記憶部6202は、1つ前の処理フレームである第L-1処理フレームにおける後天的SNR $\gamma_{L-1}(m)$ を乗算器6205に出力する。図7の抑圧係数補正部1501から入力される第L処理フレームの補正抑圧係数 $C_L(m)$ ($m=0, 1, \dots, M_L-1$)は、抑圧係数記憶部6203に入力される。抑圧係数記憶部6203は、第L処理フレームにおける補正抑圧係数 $C_L(m)$ を記憶する。これと同時に抑圧係数記憶部6203は、1つ前の処理フレームである第L-1処理フレームにおける補正抑圧係数 $C_{L-1}(m)$ を乗算器6204に出力する。乗算器6204は、供給された $C_L(m)$ を2乗して $C_{L-1}^2(m)$ を求め、乗算器6205に出力する。乗算器6205は、 $C_{L-1}^2(m)$ と $\gamma_{L-1}(m)$ を $m=0, 1, \dots, M_L-1$ に対して乗算して $C_{L-1}^2(m)\gamma_{L-1}(m)$ を求める。そして、乗算器6205は、計算結果を過去の推定SNRとして重み付き加算部6207に出力する。

[0077] 加算器6208の他方の端子には-1が供給されており、加算結果 $\gamma_L(m)-1$ が値域限定処理部6201に出力される。値域限定処理部6201は、加算器6208から入力された加算結果 $\gamma_L(m)-1$ に値域限定演算子 $P[\cdot]$ による演算を行う。そして、値域限定処理部6201は、演算の結果である $P[\gamma_L(m)-1]$ を瞬時推定SNRとして重み付き加算部6207に伝達する。ただし、 $P[x]$ は次式で定められる。

[0078] [数9]

$$P[x] = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

重み付き加算部6207には、重み記憶部6206から重みが入力されている。重み付き加算部6207は、これらの入力された瞬時推定SNR、過去の推定SNR、重みを用いて推定先天的SNRを求める。重みを α とし、 $\xi_L(m)$ ハットを推定先天的SNRとすると、 $\xi_L(m)$ ハットは、次式によって計算される。

[0079] [数10]

$$\hat{\xi}_L(m) = \alpha \gamma_{L-1}(m) C_{L-1}^2(m) + (1 - \alpha) P[\gamma_L(m) - 1]$$

ここで、 $\gamma_{-1}(m) C_{-1}^2(m) = 1$ とする。

[0080] 図15を参照して、図13に含まれる雑音抑圧係数生成部630の構成を詳細に説明する。雑音抑圧係数生成部630は、MMSE STSA ゲイン関数値計算部 6301、一般化尤度比計算部 6302、及び抑圧係数計算部 6303から構成される。以下、非特許文献4(1984年12月、アイ・イー・イー・イー・トランザクションズ・オン・アコースティクス・スピーチ・アンド・シグナル・プロセッシング、第32巻、第6号(IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, VOL.32, NO.6, PP.1109-1121, DEC, 1984)、1109~1121 ページ)に記載されている計算式をもとに、抑圧係数の計算方法を説明する。

[0081] 処理フレーム番号をL、周波数番号をmとし、 $\gamma_L(m)$ を、図13の後天的SNR 計算部610から入力される周波数別後天的SNR、 $\hat{\xi}_L(m)$ ハットを図13の推定先天的SNR計算部620から入力され推定先天的SNR、q を図13の音声非存在確率記憶部640から入力される音声非存在確率とする。また、 $\eta_L(m) = \hat{\xi}_L(m)$ ハット / (1-q)、 $v_L(m) = (\eta_L(m) \gamma_L(m)) / (1 + \eta_L(m))$ とする。

[0082] MMSE STSA ゲイン関数値計算部6301は、図13の後天的SNR計算部610から入力される後天的SNR $\gamma_L(m)$ 、図13の推定先天的SNR計算部620から入力される推定先天的SNR $\hat{\xi}_L(m)$ ハット及び図13の音声非存在確率記憶部640から入力される音声非存在確率qをもとに、周波数帯域毎にMMSE STSAゲイン関数値を計算し、抑圧係数計算部6303に出力する。第L処理フレームの統合周波数帯域毎のMMSE STSAゲイン関数値 $G_L(m)$ は、

[0083] [数11]

$$G_L(m) = \frac{\sqrt{\pi}}{2} \frac{\sqrt{v_L(m)}}{\gamma_L(m)} \exp\left(-\frac{v_L(m)}{2}\right) \left[\left(1 + v_L(m)\right) I_0\left(\frac{v_L(m)}{2}\right) + v_L(m) I_1\left(\frac{v_L(m)}{2}\right) \right]$$

で与えられる。ここで、 $I_0(z)$ は0次変形ベッセル関数、 $I_1(z)$ は1次変形ベッセル関数

である。

[0084] 一般化尤度比計算部6302は、図13の後天的SNR計算部610から入力される後天的SNR $\gamma_L(m)$ 、図13の推定先天的SNR計算部620から入力される推定先天的SNR $\xi_L(m)$ ハット及び図13の音声非存在確率記憶部 640から入力される音声非存在確率 q をもとに、周波数帯域毎に一般化尤度比を計算する。そして、一般化尤度比計算部 6302 は、一般化尤度比を抑圧係数計算部 6303 へ出力する。第L処理フレームの周波数帯域毎の一般化尤度比 $\Lambda_L(m)$ は、

[0085] [数12]

$$\Lambda_L(m) = \frac{1 - q \exp(v_L(m))}{q (1 + \eta_L(m))}$$

で与えられる。

[0086] 抑圧係数計算部6303は、MMSE STSA ゲイン関数値計算部6301から入力される MMSE STSA ゲイン関数値 $C_L(m)$ バーと一般化尤度比計算部6302から入力される一般化尤度比 $\Lambda_L(m)$ から周波数帯域毎に抑圧係数を計算する。そして、抑圧係数計算部 6303は、抑圧係数を図7の抑圧係数補正部1501へ出力する。第L処理フレームの周波数帯域毎の抑圧係数 $C_L(m)$ バーは、

[0087] [数13]

$$\bar{C}_L(m) = \frac{\Lambda_L(m)}{\Lambda_L(m) + 1} G_L(m)$$

で与えられる。周波数帯域別にSNRを計算する代わりに、複数の周波数帯域から構成される広い帯域に共通なSNRを求めて、求められた共通SNRを用いることも可能で

ある。

- [0088] 図16を参照して、抑圧係数補正部1501の構成を詳細に説明する。抑圧係数補正部1501は、最大値選択部1591、抑圧係数下限値記憶部1592、閾値記憶部1593、比較部1594、スイッチ1595、修正値記憶部1596及び乗算器1597から構成される。比較部1594は、閾値記憶部1593から入力される閾値と、雑音抑圧係数生成部601からの入力として図13の推定先天的SBR計算部620から入力される推定先天的SNRとを比較する。そして、比較部1594は、推定先天的SNRが閾値よりも大きければ0を、小さければ1をスイッチ1595に inputs。スイッチ1595は、図13の雑音抑圧係数計算部630から入力される抑圧係数を、比較部1594の出力値が1のときに乗算器1597に出力し、0のときに最大値選択部1591に出力する。すなわち、推定先天的SNRが閾値よりも小さいときに、抑圧係数の補正が行われる。乗算器1597は、スイッチ1595の出力値と修正値記憶部1596の出力値との積を計算し、最大値選択部1591に出力する。
- [0089] 一方、抑圧係数下限値記憶部1592は、記憶している抑圧係数の下限値を、最大値選択部1591に出力する。最大値選択部1591は、図13の雑音抑圧係数計算部630から入力される統合周波数帯域別抑圧係数又は乗算器1597で計算された積と、抑圧係数下限値記憶部1592から入力される抑圧係数下限値とを比較し、大きい方の値を補正抑圧係数 $C_L(m)$ として出力する。すなわち、抑圧係数は、必ず抑圧係数下限値記憶部1592が記憶する下限値以上の値になる。このとき、最大値選択部1591の出力である補正抑圧係数が雑音抑圧情報となる。抑圧係数を補正しない場合は、 $C_L(m) = \bar{C}(m)$ となる。
- [0090] これまで、雑音抑圧情報の計算は、過去の処理フレームの状態を表す値を出力するシフトレジスタ440、推定雑音記憶部3201、後天的SNR記憶部6202などにおいて、現処理フレームの統合周波数帯域のインデックス番号と同一のインデックス番号で表される過去の処理フレームの値を出力する場合について説明してきた。しかしながら、処理フレーム毎に統合周波数帯域が異なる場合、現処理フレームにおける統合周波数帯域のインデックス番号と過去の処理フレームにおける統合周波数帯域のインデックス番号が同一でも、実際の周波数帯域が異なる場合がある。この場合、現処理フレームにおいて、記憶されている過去の処理フレームの値のうち、当該帯域に最も

近い帯域のインデックス番号が指し示す値を出力するようにすると、高品質な雑音抑圧を実現できる。また、記憶されている過去の処理フレームの値をそのまま使うのではなく、現処理フレームの当該帯域に相当する値を算出し、これを用いてもよい。

[0091] 以上で、雑音抑圧情報計算部9の第1の構成説明を終える。

[0092] 続いて、図17を参照して、図1の雑音抑圧情報計算部9の第2の構成例を詳細に説明する。図7の雑音抑圧情報計算部9と比較すると、雑音抑圧係数生成部601が雑音抑圧係数生成部602に、抑圧係数補正部1501が抑圧係数補正部1502に置換されていることが異なる。雑音抑圧係数生成部602は、図13に示す雑音抑圧係数生成部601と比較すると、推定先天的SNR計算部620の出力である推定先天的SNRを出力しない点が異なり、残りの部分の動作は同一である。

[0093] 図18を参照して、図17に含まれる抑圧係数補正部1502の構成を詳細に説明する。抑圧係数補正部1502は、乗算器660、音声存在確率計算部670、仮出力SNR計算部680、抑圧係数下限値計算部6512及び最大値選択部6511から構成される。

[0094] 乗算器660は、代表劣化音声パワースペクトルと抑圧係数の積を求め、仮強調音声パワースペクトルとして音声存在確率計算部670と仮出力SNR計算部680に出力する。音声存在確率計算部670は、仮強調音声パワースペクトルと推定雑音パワースペクトルとから第L処理フレームの音声存在確率 V_L を求めて、仮出力SNR計算部680と抑圧係数下限値計算部6512とに出力する。音声存在確率の一例として、仮強調音声パワースペクトルと推定雑音パワースペクトルとの比を用いることができる。この比が大きいときには音声存在確率が高く、小さいときには音声存在確率が低い。仮出力SNR計算部680は、音声存在確率 V_L を用いて、仮出力と推定雑音パワースペクトルとから仮出力SNR $D_L(m)$ を求め、抑圧係数下限値計算部6512に出力する。仮出力SNRの一例として、仮出力の長時間平均と推定雑音パワースペクトルとから求められる長時間出力SNRを用いることができる。仮出力SNR計算部680は、音声存在確率計算部670から入力された音声存在確率 V_L の大きさに応じて仮出力の長時間平均を更新する。

[0095] 抑圧係数下限値計算部6512は、仮出力SNR $D_L(m)$ と音声存在確率 V_n とから、抑圧係数の下限値を計算し、最大値選択部6511に出力する。抑圧係数の下限値 $A(V_L, D_L)$

$A(D_L(m))$ は、次式に基づいて、関数 $A(D_L(m))$ と音声区間とに対応した抑圧係数最小値 f_s を用いて表すことができる。

[0096] [数14]

$$A(V_L, D_L(m)) = f_s \cdot V_L + (1 - V_L) \cdot A(D_L(m))$$

関数 $A(D_L(m))$ は基本的に、大きなSNRに対して小さな値をとるような形状を有する。 $A(D_L(m))$ が仮出力SNR $D_L(m)$ に対応してこのような形状をとる関数であることは、仮出力SNRが高いほど、非音声区間に対応する抑圧係数の下限値が小さくなることを意味する。これは、残留雑音が小さくなることに対応し、音声区間と非音声区間の音質不連続性を低減する効果がある。なお、関数 $A(D_L(m))$ は全ての周波数成分に対して異なってもよいし、複数の周波数成分に対して共有されていてもよい。また、時間と共にその形状が変化することも可能である。

[0097] 最大値計算部6511は、雑音抑圧係数計算部630から入力された抑圧係数 $C_L(m)$ と抑圧係数下限値計算部6512から入力された抑圧係数の下限値とを比較して、大きいほうの値を補正抑圧係数 $C_L(m)$ として出力する。この処理は、次式で表すことができる。

[0098] [数15]

$$C_L(m) = \begin{cases} \bar{C}_L(m) & \bar{C}_L(m) \geq A(V_L, D_L(m)) \\ A(V_L, D_L(m)) & \bar{C}_L(m) < A(V_L, D_L(m)) \end{cases}$$

すなわち、完全に音声区間と思われる場合は f_s が、完全に非音声区間と思われる場合は仮出力 $D_L(m)$ に応じて単調減少関数で定められる値が、抑圧係数最小値となる。両者の中間と思われる状況では、これらの値が適切に混合される。 $A(D_L(m))$ の単調減少性によって、低SNR時の大きな抑圧係数最小値が保証される。これにより、消し残し雑音の多い直前の音声区間からの連続性が保たれる。高SNR時は、抑圧係数

最小値が小さくなり、残留雑音が小さくなるように制御される。これは、音声区間の残留雑音が無視できる程度に小さいので、非音声区間の残留雑音が小さいときも、連続性が保たれるためである。また、 f_s を $A(D_L(m))$ よりも大きく設定することによって、音声区間あるいはその可能性が高い場合に雑音抑圧が軽度になり、音声に生じる歪を低減することができる。これは、符号化・復号によって生じる歪の混入した音声において雑音推定精度が十分に高くできない場合に、特に有効である。

[0099] 以上で、雑音抑圧情報計算部9の第2の構成の説明を終える。

[0100] 図1に戻り、本発明の最良の実施の形態の構成の説明をする。雑音抑圧処理部10は、劣化音声パワースペクトル、処理フレーム情報及び雑音抑圧情報を用いて、強調音声パワースペクトル $|X_n(k)|^2$ バーを計算し、逆変換部6へ出力する。たとえば、第L処理フレームの統合周波数帯域mに含まれる劣化音声パワースペクトルに共通の雑音抑圧情報を適用することで、強調音声パワースペクトルを算出することができる。すなわち、[数5]の代表劣化音声パワースペクトル $Z_L(m)$ の算出の際に使用した劣化音声パワースペクトルに共通の雑音抑圧情報 $C_L(m)$ を乗じる。これは、図4および図6において、灰色で囲まれている格子1つに含まれる全ての劣化音声パワースペクトルに共通の雑音抑圧情報 $C_L(m)$ を適用することに相当する。強調音声パワースペクトル $|X_n(k)|^2$ バーは、

[0101] [数16]

$$|\bar{X}_n(k)|^2 = C_L^2(m) |\bar{Y}_n(k)|^2 \quad (n_L \leq n < n_{L+1}, k_m \leq k < k_{m+1})$$

となる。

[0102] 強調音声パワースペクトルを算出する他の方法として、複数処理フレームの雑音抑圧情報を用いて算出する方法もある。たとえば、1処理フレーム前の雑音抑圧情報 $C_{L-1}(m)$ を用いて補間を行うと、

[0103] [数17]

$$|\bar{X}_n(k)|^2 = \left(C_{L-1}^2(m) + n \cdot \frac{C_L^2(m) - C_{L-1}^2(m)}{n_{L+1} - n_L} \right) \cdot |\bar{Y}_n(k)|^2$$

$$(n_L \leq n < n_{L+1}, k_m \leq k < k_{m+1})$$

となる。このように補間を行った雑音抑圧情報を用いると、処理フレームの境界付近での不連続感を低減することができ、高品質な雑音抑圧を実現できる。また、あらかじめ、複数処理フレームの雑音抑圧情報を平滑化してから上記方法を用いてもよい。この場合、雑音抑圧情報の急激な変化を避けることができ、高品質な雑音抑圧を実現できる。他にも、雑音抑圧情報を周波数方向に補間してから、強調音声パワースペクトルを算出してもよい。また、時間方向および周波数方向の両方に平滑化を行った雑音抑圧情報を劣化音声パワースペクトルに適用してもよい。

[0104] 逆変換部6は、雑音抑圧処理部10から入力される強調音声パワースペクトル $|X_n(k)|^2$ バーを用いて求めた強調音声振幅スペクトル $|X_n(k)|$ バーと変換部5から入力された位相 $\arg Y_n(k)$ を乗算して、強調音声スペクトル $X_n(k)$ バーを求める。すなわち、

[0105] [数18]

$$\bar{X}_n(k) = |\bar{X}_n(k)| \cdot \arg Y_n(k)$$

を実行する。

[0106] 逆変換部6は、得られた強調音声スペクトル $X_n(k)$ バーに逆周波数変換を施し、時間領域信号を生成する。このとき、逆変換部6が適用する逆周波数変換は、変換部5が適用する周波数変換と対応する逆変換が選択されることが望ましい。変換部5において、窓関数 W で重み付けしている場合は、逆周波数変換した信号に、窓関数 W を乗じる。変換部5を帯域分割フィルタバンクで構成するときには、逆変換部6を帯域合成フィルタバンクで構成する。帯域合成フィルタバンクとその設計法に関連する技術は、非特許文献3に開示されている。逆周波数変換された時間領域信号は、変換フレーム合成部3へ出力される。

- [0107] 変換フレーム合成部3は、入力された変換フレーム長に分割されている逆周波数変換された時間領域信号を合成し、強調音声信号サンプルを出力端子4に出力する。
- [0108] このように、入力信号の変化に追従できる短い変換フレーム長を用いながら、変換フレームを統合した処理フレームで雑音抑圧情報の算出を行うため、高品質な雑音抑圧を実現するとともに、雑音抑圧情報の計算回数を削減でき演算量を削減できる。さらに、処理フレームを入力信号に応じて適応的に定めることで、高品質な雑音抑圧を低演算量で実現することができる。
- [0109] 以上で、最良の実施の形態の説明を終える。
- [0110] 続いて、図19を参照して本発明の第2の実施の形態について詳細に説明する。
- [0111] 本発明の第2の実施の形態は、最良の実施の形態を表す図1と比較すると、雑音抑圧情報計算部9が雑音抑圧情報計算部11に置換され、新たに処理フレーム情報が入力されている点が異なる。図1の構成と共通する構成については説明を省略する。以下、雑音抑圧情報計算部11について詳細に説明する。
- [0112] 図20を参照して、図19に含まれる雑音抑圧情報計算部11の第1の構成例を詳細に説明する。図7の雑音抑圧情報部9と比較すると、雑音推定部300が雑音推定部301に置換され、新たに処理フレーム情報が入力されている点が異なる。
- [0113] 図21を参照して、図20に含まれる雑音推定部301の構成を詳細に説明する。図8の雑音推定部300とは、カウンタ330がカウンタ331に置換され、新たに処理フレーム情報が入力されている点が異なる。カウンタ331は、カウント値を出力する。カウンタ値の初期値は、0に設定されている。カウンタ331は、処理フレーム毎に当該処理フレームの処理フレーム長をカウント値に加算していく。すなわち、第L処理フレームのカウント値をCnt(L)とすると、第L+1処理フレームのカウント値Cnt(L+1)は、
- [0114] [数19]

$$Cnt(L + 1) = Cnt(L) + (n_{L+1} - n_L)$$

となる。従って、推定雑音計算部310の更新判定部400において、カウンタ331のカウント値と閾値とを比較する場合、図10の閾値記憶部4003の値は、カウンタ330を用いたときより大きい値を設定することが一般的である。

- [0115] 上述の構成により、処理フレーム長が処理フレーム毎に異なっても、定められた時間を正確に判定することができ、高品位な雑音推定を実現することができる。
- [0116] 図22を参照して、雑音抑圧情報計算部11の第2の構成例を詳細に説明する。図20の雑音抑圧情報計算部11と比較すると、雑音抑圧係数生成部601が雑音抑圧係数生成部602に、抑圧係数補正部1501が抑圧係数補正部1502に置換されていることが異なる。雑音抑圧係数生成部602の構成及び抑圧係数補正部1502の構成は、図17を参照して既に詳細に説明しているため、ここでは説明を省略する。また、雑音推定部301の構成については、図21を参照して既に説明しているため、ここでは説明を省略する。
- [0117] 本実施の形態では、処理フレーム長を用いて制御を行う例としてカウンタ331の動作について説明したが、その他の部分にも適用できる。たとえば、推定雑音パワースペクトルを算出する際に、推定雑音計算部310のシフトレジスタ440に保存されている重みつき劣化音声パワースペクトルのうち、当該処理フレームから定められた過去の時間に含まれる処理フレームの重みつき劣化音声パワースペクトルのみを用いて、これらの平均を推定雑音パワースペクトルとすることもできる。このような構成により、処理フレーム長の長短によらず、一定時間内の信号を用いて推定雑音を算出することができるため、高品位な雑音推定を実現できる。
- [0118] 以上で、本発明の第2の実施の形態の説明を終える。
- [0119] 続いて、図23を参照して、本発明の第3の実施の形態について詳細に説明する。
- [0120] 本発明の第3の実施の形態は、最良の実施の形態を表す図1と比較すると、処理フレーム情報生成部7が処理フレーム情報生成部14に置換されている点で異なる。また、定められた一定時間内における処理フレームの個数の最大値が、処理フレーム情報生成部14に入力されている点が異なる。処理フレーム情報生成部14は、定められた一定時間内における処理フレームの個数が入力された最大値以下になるように、処理フレームを定め、処理フレーム情報を出力する。

- [0121] 図24を参照して、図23の処理フレーム情報生成部14の第1の構成例を詳細に説明する。図2の処理フレーム情報生成部7と比較すると、時間グループ生成部51が時間グループ生成部58に置換されている点が異なる。また、時間グループ生成部58に最大値が入力されている点が異なる。入力された最大値をLNとすると、時間グループ生成部58が生成する処理フレームの個数は、定められた一定時間内における処理フレームの個数が、最大値LN以下になるように、変換フレームを統合し、処理フレームの区切り位置を定める。時間グループ生成部58もにおいて処理フレームの区切り位置を定める方法として、図3を用いて説明した変換フレームエネルギー $E(n)$ の変化量に基づいた方法がある。このとき、時間グループ生成部58は、変化量大きい箇所から順次処理フレームが分割されるように処理フレームの区切り位置を生成する。そして、時間グループ生成部58は、生成した処理フレームの個数がLN個になった時点で区切り位置の生成を終了する。
- [0122] 図25を参照して、図23の処理フレーム情報生成部14の第2の構成例を詳細に説明する。
- [0123] 図24の処理フレーム情報生成部7と比較すると、新たに周波数エネルギー計算部53を備え、周波数グループ生成部52が周波数グループ生成部54に置換されている点で異なる。周波数エネルギー計算部53および周波数グループ生成部52は、図5を参照して既に詳細に説明しているので、ここでは説明を省略する。
- [0124] このように、処理フレーム情報生成部14を構成することにより、一定時間内での処理フレームの個数の最大値を定めることができる。従って、雑音抑圧情報の計算回数を制御することができ、演算量を削減することができる。
- [0125] 以上で、本発明の第3の実施の形態の説明を終える。
- [0126] 続いて、図26を参照して本発明の第4の実施の形態について詳細に説明する。
- [0127] 本発明の第4の実施の形態は、最良の実施の形態を表す図1と比較すると、処理フレーム情報生成部7が処理フレーム情報生成部12に置換されている点が異なる。さらに、新たに、定められた一定時間内における雑音抑圧情報の計算回数の最大値が処理フレーム情報生成部12に入力されている点のみが異なる。処理フレーム情報生成部12は、雑音抑圧情報の計算回数が供給された最大値以下になるように、処理フ

レームと統合周波数帯域とを定め、処理フレーム情報を出力する。

- [0128] 図27を参照して、図26の処理フレーム情報生成部12の構成例を詳細に説明する。図5の処理フレーム情報生成部7と比較すると、時間グループ生成部51が時間グループ生成部55に、周波数グループ生成部54が周波数グループ生成部56に置換されている点で異なる。さらに、時間グループ生成部55と周波数グループ生成部56に最大値が入力されている点が異なる。
- [0129] 処理フレーム情報生成部12に入力された最大値をLMとすると、時間グループ生成部55が生成する処理フレームの個数TNは、関数fを用いて、 $TN=f(LM)$ と表される。ここで関数fの例として最大値をLMの正の平方根を超えない最大の整数としてもよい。他にも、最大値LMを定数で除算した値を超えない最大の整数としてもよい。時間グループ生成部55は、処理フレームの個数がTN個になるように、変換フレームを統合し、処理フレームの区切り位置を定める。処理フレームの区切り位置を定める方法として、図5を参照して既に説明したように、変換フレームエネルギーE(n)の変化量に基づいた方法がある。このとき、時間グループ生成部55は、変化量大きい箇所から順次処理フレームが分割されるように処理フレームを生成していく。そして、時間グループ生成部55は、生成した処理フレームの個数がTN個になった時点で区切り位置の生成を終了する。
- [0130] 周波数グループ生成部56は、各処理フレームにおいて、複数の周波数帯域を統合し、統合周波数帯域の区切り位置を定め、処理フレーム情報を出力する。各処理フレームにおける統合周波数帯域の最大個数FNは、 $FN=\text{int}(LM/TN)$ で定められる。ここで、 $\text{int}(X)$ は、Xを超えない最大の整数である。すなわち、図6を参照して既に説明した第L処理フレームの統合周波数帯域の個数 M_L がFNを越えないように統合周波数帯域を設定する。周波数グループ生成部56は、統合周波数帯域の設定において、周波数エネルギー計算部53から入力された周波数エネルギーの変化が大きい箇所で統合周波数帯域が分割されるように区切り位置を定めていく。
- [0131] このように、処理フレーム情報生成部を構成することにより、一定時間内の雑音抑圧情報の計算回数の最大値を定めることができるため、演算量を削減することができる。

[0132] 以上で、本発明の第4の実施の形態の説明を終える。

[0133] 続いて、図28を参照して、本発明の第5の実施の形態を詳細に説明する。本発明の第5の実施の形態は、最良の実施の形態を表す図1と比較すると、処理フレーム情報生成部7が処理フレーム情報生成部13に置換されている点異なる。さらに、変換フレームに分割された劣化音声信号が処理フレーム情報生成部13に入力されている点異なる。

[0134] 図29を参照して、処理フレーム情報生成部13の構成例を詳細に説明する。図2の処理フレーム情報生成部7と比較すると、変換フレームエネルギー計算部50が変換フレームエネルギー計算部57に置換されている点異なる。変換フレームエネルギー計算部57は、変換フレーム長に分割された入力信号サンプルの2乗和を変換フレームエネルギー $E(n)$ として、時間グループ生成部51に出力する。

[0135] 本実施の形態では、周波数変換された信号を分析して処理フレーム情報を算出するのではなく、時間信号を分析して処理フレーム情報を算出する点を特徴とする。このため、周波数変換と処理フレーム情報の算出を並列で行うことができる。これにより演算量を削減することができる。さらに、並列プロセッサ等を用いることにより、より演算量の削減を実現することができる。

[0136] 以上で、本発明の第5の実施の形態の説明を終える。

[0137] 続いて、図30を参照して、本発明の第6の実施の形態を詳細に説明する。

[0138] 本発明の第6の実施の形態は、最良の実施の形態を表す図1と比較すると、処理フレーム情報生成部7が処理フレーム情報生成部15に置換されている点異なる。処理フレーム情報生成部15は、処理フレーム情報を生成し、代表周波数領域信号生成部8、雑音抑圧処理部10へ出力する。

[0139] 図31を参照して、処理フレーム情報生成部15の構成例を詳細に説明する。処理フレーム情報生成部15は、時間グループ生成部60と周波数グループ生成部52とから構成される。時間グループ生成部60は、代表劣化音声パワースペクトルを計算するための処理フレームの区切り位置を定め、周波数グループ生成部52へ出力する。処理フレームの区切り位置は、あらかじめ定められた処理フレーム長になるように定める。処理フレーム長を定める方法として、入力信号のサンプリング周波数や、演算能

力に応じた方法がある。たとえば、サンプリング周波数が高くなるにつれ、処理フレーム長が長くなるように区切り位置を定める。これにより、1処理フレームの時間を低いサンプリング周波数のときと同等にすることができる。また、演算能力が低い場合、処理フレーム長が長くなるように区切り位置を定めることにより、後段に続く雑音抑圧情報の計算回数を削減できる。また、他の機能とのリソース配分に基づいて、ノイズサプレッサが使用可能なリソースに基づいて処理フレームの区切り位置を定めても良い。この場合、時々刻々とノイズサプレッサが使用可能なリソースが変化するため、ノイズサプレッサが使用可能なリソースに応じて処理フレーム長を定める。周波数グループ生成部52の動作については、図2を参照して詳細は既に説明しているため、ここでは説明を省略する。ここで、演算能力や他の機能とのリソース配分に基づいて、統合周波数帯域の区切り位置を定めることもできる。

[0140] このように、処理フレーム情報生成部15を構成することにより、処理フレーム情報を計算するための演算量を大幅に削減できるため、低演算量で雑音抑圧を行うことができる。

[0141] 以上で、本発明の第6の実施の形態の説明を終える。

[0142] 続いて、図32を参照して、本発明の第7の実施の形態を詳細に説明する。

[0143] 本発明の第7の実施の形態は、最良の実施の形態を表す図1と比較すると、雑音抑圧処理部10が雑音抑圧処理部16に置換されている点が異なる。さらに、雑音抑圧処理部16には、劣化音声パワースペクトルではなく代表劣化音声パワースペクトルが入力されている点が異なる。

[0144] 雑音抑圧処理部16は、雑音抑圧情報 $C_L(m)$ と処理フレーム情報と代表劣化音声パワースペクトルとから強調音声パワースペクトルを算出し、これを逆変換部6へ出力する。強調音声パワースペクトル $|\bar{X}_n(k)|^2$ バーは、

[0145] [数20]

$$|\bar{X}_n(k)|^2 = C_L^2(m) \cdot Z_L(m) \quad (n_L \leq n < n_{L+1}, k_m \leq k < k_{m+1})$$

となる。

[0146] 強調音声パワースペクトルを算出する他の方法として、複数処理フレームの雑音抑圧情報を用いて算出する方法もある。たとえば、1処理フレーム前の雑音抑圧情報 $C_{L-1}(m)$ を用いて補間を行うと、

[0147] [数21]

$$\left| \bar{X}_n(k) \right|^2 = \left(C_{L-1}^2(m) + n \cdot \frac{C_L^2(m) - C_{L-1}^2(m)}{n_{L+1} - n_L} \right) \cdot Z_L(m) \\ (n_L \leq n < n_{L+1}, k_m \leq k < k_{m+1})$$

となる。もちろん、複数処理フレームの雑音抑圧情報から補間を行ってもよい。このように補間を行った雑音抑圧情報を用いると、処理フレームの境界付近での不連続感を低減することができ、高品質な雑音抑圧を実現できる。また、あらかじめ複数処理フレームの雑音抑圧情報を平滑化してから上記方法を用いてもよい。この場合、雑音抑圧情報の急激な変化を避けることができ、高品質な雑音抑圧を実現できる。他にも、雑音抑圧情報を周波数方向に補間してから、強調音声パワースペクトルを算出してもよい。また、時間方向および周波数方向の両方に平滑化を行った雑音抑圧情報を劣化音声パワースペクトルに適用してもよい。

[0148] 以上で、本発明の第7の実施の形態の説明を終える。

[0149] 続いて、図33を参照して、本発明の第8の実施の形態を詳細に説明する。

[0150] 本発明の第8の実施の形態は、記録部30と再生部31とから構成される。記録部30は、入力端子1から入力信号が供給され、入力信号の雑音を抑圧するための情報を算出し、入力信号と算出した情報を多重化し多重化信号を出力する。一方、再生部31は、記録部30が出力した多重化信号を受け取り、多重化信号に含まれる雑音を抑圧するための情報に基づき、多重化信号に含まれる入力信号の雑音を抑圧し、出力端子4へ出力する。

[0151] 記録部30は、変換フレーム分割部2、変換部5、処理フレーム情報生成部7、代表周波数領域信号生成部8、雑音抑圧情報計算部9、多重化部32から構成される。変換フレーム分割部2、変換部5、処理フレーム情報生成部7、代表周波数領域信号生成

部8、雑音抑圧情報計算部9については、図1を参照して既に詳細に説明しているため、ここでは説明を省略する。

[0152] 多重化部32は、入力信号と処理フレーム情報と雑音抑圧情報を多重化し多重化信号を出力する。

[0153] 再生部31は、分離部33、変換フレーム分割部2、変換部5、雑音抑圧処理部10、逆変換部6、変換フレーム合成部3から構成される。変換フレーム分割部2、変換部5、雑音抑圧処理部10、逆変換部6、変換フレーム合成部3については、図1を参照して既に詳細に説明しているため、ここでは説明を省略する。

[0154] 分離部33は、入力された多重化信号を入力信号と処理フレーム情報と雑音抑圧情報に分離し、入力信号を変換フレーム分割部2へ、処理フレーム情報と雑音抑圧情報を雑音抑圧処理部10へ出力する。

[0155] ここで、多重化信号を一度蓄積媒体に保存して、再生時に、蓄積媒体から多重化信号を取出してもよい。また、入力信号をそのまま多重化するのではなく、符号化を行い、情報圧縮したものを多重化してもよい。この場合、再生部31には、記録部30とは逆の入力信号を復号する機能である復号部を備える。同様に、処理フレーム情報および雑音抑圧情報を符号化できることは明らかである。

[0156] ここでは、録音部30と再生部31が同一の端末に存在する場合を仮定して説明したが、録音部30と再生部31が異なる端末に存在していてもよい。この場合、録音部30の出力である多重化信号が伝送路等を通り、別の端末にある再生部31に出力される。また、多重化信号を蓄積媒体に保存し、別の端末にある再生部31に入力してもよい。

[0157] このように構成することにより、録音された信号を再生するさいに、雑音抑圧情報を計算する必要がないため、再生時の演算量を削減することができる。

[0158] 以上で、本発明の第8の実施の形態の説明を終える。

[0159] 続いて、図34を参照して、本発明の第9の実施の形態を詳細に説明する。

[0160] 本発明の第9の実施の形態は、プログラム制御により動作するコンピュータ100を備える。コンピュータ100は、入力端子1から受けた入力信号に対し、上述した最良の実施の形態および本発明の第2から第8の実施の形態のいずれかに係る処理を行い、

出力端子4に強調音声を出力するためのプログラムに基づき動作する。

[0161] 以上で、本発明の第9の実施の形態の説明を終える。

[0162] これまで説明した全ての実施の形態では、雑音抑圧の方式として、最小平均2乗誤差短時間スペクトル振幅法を仮定してきたが、その他の方法にも適用することができる。このような方法の例として、非特許文献5(1979年12月、プロシーディングス・オブ・ザ・アイ・イー・イー・イー、第67巻、第12号(PROCEEDINGS OF THE IEEE, VOL.67, NO.12, PP.1586-1604, DEC, 1979)、1586～1604ページ)に開示されているウィナーフィルタ法や、非特許文献6(1979年4月、アイ・イー・イー・イー・トランザクションズ・オン・アコースティクス・スピーチ・アンド・シグナル・プロセッシング、第27巻、第2号(IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, VOL.27, NO.2, PP.113-120, APR, 1979)、113～120ページ)に開示されているスペクトル減算法などがあるが、これらの詳細な構成例については説明を省略する。

[0163] 以上の如く、実施の形態を説明したが、本発明の実施例は以下の通りである。

[0164] すなわち、本発明の第1の実施例は、入力信号を定められた第一のフレーム毎に周波数領域信号に変換する変換手段と、前記第一のフレームと異なるように第二のフレームを生成するフレーム生成手段と、前記第二のフレームに含まれる第一のフレームの前記周波数領域信号から代表周波数領域信号を生成する代表周波数領域信号生成手段と、前記代表周波数領域信号に基づいて、前記第二のフレームの雑音抑圧の程度を求める雑音抑制程度算出手段とを有することを特徴とする雑音抑圧装置である。

[0165] また、本発明の第2の実施例は、上記実施例において、前記フレーム生成手段が、フレームの長さが前記第一のフレームの長さより長い第二のフレームを生成することを特徴とする。

[0166] また、本発明の第3の実施例は、上記実施例において、前記フレーム生成手段が、前記第二のフレーム同士が互いに独立するように、前記第二のフレームを生成することを特徴とする。

[0167] また、本発明の第4の実施例は、上記実施例において、前記雑音抑制程度算出手

段が、前記第二のフレームに含まれる前記周波数領域信号に、前記雑音抑圧の程度を適用して雑音を抑圧することを特徴とする。

[0168] また、本発明の第5の実施例は、上記実施例において、前記雑音抑制程度算出手段が、前記第二のフレームに含まれる前記周波数領域信号に、他の第二のフレームの前記雑音抑圧の程度を補間して算出した雑音抑圧の程度を適用して雑音を抑圧することを特徴とする。

[0169] また、本発明の第6の実施例は、上記実施例において、前記フレーム生成手段が、前記周波数領域信号の特性に基づいて、第二のフレームを生成することを特徴とする。

[0170] また、本発明の第7の実施例は、上記実施例において、前記周波数領域信号の特性が、前記入力信号のエネルギーの変化であることを特徴とする。

[0171] また、本発明の第8の実施例は、上記実施例において、前記第二のフレーム毎に、周波数方向の区切り位置を生成する周波数区切り位置生成手段を有し、前記代表周波数領域信号生成手段は、前記第二のフレームおよび前記周波数方向の区切り位置に基づいて、前記周波数領域信号から前記代表周波数領域信号を生成することを特徴とする。

[0172] また、本発明の第9の実施例は、上記実施例において、前記フレーム生成手段が、一定区間内の第二のフレームの個数があらかじめ定められた個数の範囲内になるように、前記第二のフレームを生成することを特徴とする。

[0173] また、本発明の第10の実施例は、上記実施例において、前記フレーム生成手段が、一定区間内の前記雑音抑圧の程度を算出する計算回数があらかじめ定められた回数の範囲内になるように、前記第二のフレームと前記周波数方向の区切り位置とを定めることを特徴とする。

[0174] また、本発明の第11の実施例は、上記実施例において、前記雑音抑圧の程度が雑音抑圧係数として表されていることを特徴とする。

[0175] また、本発明の第12の実施例は、上記実施例において、前記雑音抑圧の程度が雑音の推定値として表されていることを特徴とする。

[0176] 本発明の第13の実施例は、入力信号を定められた第一のフレーム毎に周波数領

域信号に変換する変換ステップと、前記第一のフレームと異なるように第二のフレームを生成するフレーム生成ステップと、前記第二のフレームに含まれる第一のフレームの前記周波数領域信号から代表周波数領域信号を生成する代表周波数領域信号生成ステップと、前記代表周波数領域信号に基づいて、前記第二のフレームの雑音抑圧の程度を求める雑音抑制程度算出ステップとを有することを特徴とする雑音抑圧方法である。

[0177] また、本発明の第14の実施例は、上記実施例において、前記フレーム生成ステップは、前記第一のフレームの長さより長い前記第二のフレームを生成することを特徴とする。

[0178] また、本発明の第15の実施例は、上記実施例において、前記フレーム生成ステップが、前記第二のフレーム同士が互いに独立するように、前記第二のフレームを生成することを特徴とする。

[0179] また、本発明の第16の実施例は、上記実施例において、前記雑音抑制程度算出ステップが、前記第二のフレームに含まれる前記周波数領域信号に前記雑音抑圧の程度を適用して雑音を抑圧することを特徴とする。

[0180] また、本発明の第17の実施例は、上記実施例において、前記雑音抑制程度算出ステップが、前記第二のフレームに含まれる前記周波数領域信号に、他の第二のフレームの前記雑音抑圧の程度を補間して算出した雑音抑圧の程度を適用して雑音を抑圧することを特徴とする。

[0181] また、本発明の第18の実施例は、上記実施例において、前記フレーム生成ステップは、前記周波数領域信号の特性に基づいて、前記第二のフレームを生成することを特徴とする。

[0182] また、本発明の第19の実施例は、上記実施例において、前記周波数領域信号の特性が前記入力信号のエネルギーの変化であることを特徴とする。

[0183] また、本発明の第20の実施例は、上記実施例において、前記第二のフレーム毎に、周波数方向の区切り位置を生成する周波数区切り位置生成ステップを有し、前記代表周波数領域信号生成ステップは、前記第二のフレームおよび前記周波数方向の区切り位置に基づいて、前記周波数領域信号から代表周波数領域信号を生成す

ることを特徴とする。

- [0184] また、本発明の第21の実施例は、上記実施例において、前記フレーム生成ステップが、一定区間内の前記第二のフレームの個数があらかじめ定められた個数の範囲内になるように前記第二のフレームを生成することを特徴とする。
- [0185] また、本発明の第22の実施例は、上記実施例において、前記フレーム生成ステップが、一定区間内の前記雑音抑圧の程度を算出する計算回数があらかじめ定められた回数の範囲内になるように、前記第二のフレームと前記周波数方向の区切り位置とを生成することを特徴とする。
- [0186] また、本発明の第23の実施例は、上記実施例において、前記雑音抑制程度算出ステップにおいて、前記雑音抑圧の程度が雑音抑圧係数として表されていることを特徴とする。
- [0187] また、本発明の第24の実施例は、上記実施例において、前記雑音抑制程度算出ステップにおいて、前記雑音抑圧の程度が雑音の推定値として表されていることを特徴とする。
- [0188] 本発明の第25の実施例は、コンピュータに、入力信号を定められた第一のフレーム毎に周波数領域信号に変換する変換処理と、前記第一のフレームと異なるように第二のフレームを生成するフレーム生成処理と、前記第二のフレームに含まれる第一のフレームの前記周波数領域信号から代表周波数領域信号を生成する代表周波数領域信号生成処理と、前記代表周波数領域信号に基づいて、前記第二のフレームの雑音抑圧の程度を求める雑音抑制程度算出処理とを実行させる雑音抑圧プログラムである。
- [0189] また、本発明の第26の実施例は、上記実施例において、前記フレーム生成処理が、前記第一のフレームの長さより長い前記第二のフレームを生成することを特徴とする。
- [0190] また、本発明の第27の実施例は、上記実施例において、前記フレーム生成処理が、前記第二のフレーム同士が互いに独立するように前記第二のフレームを生成することを特徴とする。
- [0191] また、本発明の第28の実施例は、上記実施例において、前記雑音抑制程度算出

処理が、前記第二のフレームに含まれる前記周波数領域信号に前記雑音抑圧の程度を適用して雑音を抑圧することを特徴とする。

[0192] また、本発明の第29の実施例は、上記実施例において、前記雑音抑制程度算出処理が、前記第二のフレームに含まれる前記周波数領域信号に、他の第二のフレームの前記雑音抑圧の程度を補間して算出した雑音抑圧の程度を適用して雑音を抑圧することを特徴とする。

[0193] また、本発明の第30の実施例は、上記実施例において、前記フレーム生成処理が、前記周波数領域信号の特性に基づいて、前記第二のフレームを生成することを特徴とする。

[0194] また、本発明の第31の実施例は、上記実施例において、前記周波数領域信号の特性が前記入力信号のエネルギーの変化であることを特徴とする。

[0195] また、本発明の第32の実施例は、上記実施例において、前記第二のフレーム毎に、周波数方向の区切り位置を生成する周波数区切り位置生成処理を有し、前記代表周波数領域信号生成処理は、前記第二のフレームおよび前記周波数方向の区切り位置に基づいて、前記周波数領域信号から代表周波数領域信号を生成することを特徴とする。

[0196] また、本発明の第33の実施例は、上記実施例において、前記フレーム生成処理は、一定区間内の前記第二のフレームの個数があらかじめ定められた個数の範囲内になるように前記第二のフレームを生成することを特徴とする。

[0197] また、本発明の第34の実施例は、上記実施例において、前記フレーム生成処理は、一定区間内の前記雑音抑圧の程度を算出する計算回数があらかじめ定められた回数の範囲内になるように前記第二のフレームと前記周波数方向の区切り位置とを生成することを特徴とする。

[0198] また、本発明の第35の実施例は、上記実施例において、前記雑音抑制程度算出処理において、前記雑音抑圧の程度が雑音抑圧係数として表されていることを特徴とする。

[0199] また、本発明の第36の実施例は、上記実施例において、前記雑音抑制程度算出処理において、前記雑音抑圧の程度が雑音の推定値として表されていることを特徴

とする。

- [0200] 以上好ましい実施の形態及び実施例をあげて本発明を説明したが、本発明は必ずしも上記実施の形態及び実施例に限定されるものではなく、その技術的思想の範囲内において様々に変形し実施することが出来る。
- [0201] 本出願は、2007年9月19日に出願された日本出願特願2007-243001号を基礎とする優先権を主張し、その開示の全てをここに取り込む。

請求の範囲

- [1] 入力信号を定められた第一のフレーム毎に周波数領域信号に変換する変換手段と、
前記第一のフレームと異なるように第二のフレームを生成するフレーム生成手段と、
前記第二のフレームに含まれる第一のフレームの前記周波数領域信号から代表周波数領域信号を生成する代表周波数領域信号生成手段と、
前記代表周波数領域信号に基づいて、前記第二のフレームの雑音抑圧の程度を求める雑音抑制程度算出手段と
を有することを特徴とする雑音抑圧装置。
- [2] 前記フレーム生成手段は、フレームの長さが前記第一のフレームの長さより長い第二のフレームを生成することを特徴とする請求項1に記載の雑音抑圧装置。
- [3] 前記フレーム生成手段は、前記第二のフレーム同士が互いに独立するように、前記第二のフレームを生成することを特徴とする請求項1又は請求項2に記載の雑音抑圧装置。
- [4] 前記雑音抑制程度算出手段は、前記第二のフレームに含まれる前記周波数領域信号に、前記雑音抑圧の程度を適用して雑音を抑圧することを特徴とする請求項1から請求項3のいずれかに記載の雑音抑圧装置。
- [5] 前記雑音抑制程度算出手段は、前記第二のフレームに含まれる前記周波数領域信号に、他の第二のフレームの前記雑音抑圧の程度を補間して算出した雑音抑圧の程度を適用して雑音を抑圧することを特徴とする請求項1から請求項4のいずれかに記載の雑音抑圧装置。
- [6] 前記フレーム生成手段は、前記周波数領域信号の特性に基づいて、第二のフレームを生成することを特徴とする請求項1から請求項5のいずれかに記載の雑音抑圧装置。
- [7] 前記周波数領域信号の特性が、前記入力信号のエネルギーの変化であることを特徴とする請求項6に記載の雑音抑圧装置。
- [8] 前記第二のフレーム毎に、周波数方向の区切り位置を生成する周波数区切り位置生成手段を有し、

前記代表周波数領域信号生成手段は、前記第二のフレームおよび前記周波数方向の区切り位置に基づいて、前記周波数領域信号から前記代表周波数領域信号を生成すること

を特徴とすることを特徴とする請求項1から請求項7のいずれかに記載の雑音抑圧装置。

[9] 前記フレーム生成手段は、一定区間内の第二のフレームの個数があらかじめ定められた個数の範囲内になるように、前記第二のフレームを生成することを特徴とする請求項1から請求項8のいずれかに記載の雑音抑圧装置。

[10] 前記フレーム生成手段は、一定区間内の前記雑音抑圧の程度を算出する計算回数があらかじめ定められた回数の範囲内になるように、前記第二のフレームと前記周波数方向の区切り位置とを求めることを特徴とする請求項8または9に記載の雑音抑圧装置。

[11] 前記雑音抑圧の程度が雑音抑圧係数として表されていることを特徴とする請求項1から請求項10のいずれかに記載の雑音抑圧装置。

[12] 前記雑音抑圧の程度が雑音の推定値として表されていることを特徴とする請求項1から請求項10のいずれかに記載の雑音抑圧装置。

[13] 入力信号を定められた第一のフレーム毎に周波数領域信号に変換する変換ステップと、

前記第一のフレームと異なるように第二のフレームを生成するフレーム生成ステップと、

前記第二のフレームに含まれる第一のフレームの前記周波数領域信号から代表周波数領域信号を生成する代表周波数領域信号生成ステップと、

前記代表周波数領域信号に基づいて、前記第二のフレームの雑音抑圧の程度を求める雑音抑制程度算出ステップと

を有することを特徴とする雑音抑圧方法。

[14] 前記フレーム生成ステップは、前記第一のフレームの長さより長い前記第二のフレームを生成することを特徴とする請求項13に記載の雑音抑圧方法。

[15] 前記フレーム生成ステップは、前記第二のフレーム同士が互いに独立するように、

前記第二のフレームを生成することを特徴とする請求項13又は請求項14に記載の雑音抑圧方法。

- [16] 前記雑音抑制程度算出ステップは、前記第二のフレームに含まれる前記周波数領域信号に前記雑音抑圧の程度を適用して雑音を抑圧することを特徴とする請求項13から請求項15のいずれかに記載の雑音抑圧方法。
- [17] 前記雑音抑制程度算出ステップは、前記第二のフレームに含まれる前記周波数領域信号に、他の第二のフレームの前記雑音抑圧の程度を補間して算出した雑音抑圧の程度を適用して雑音を抑圧することを特徴とする請求項13から請求項16のいずれかに記載の雑音抑圧方法。
- [18] 前記フレーム生成ステップは、前記周波数領域信号の特性に基づいて、前記第二のフレームを生成することを特徴とする請求項13から請求項17のいずれかに記載の雑音抑圧方法。
- [19] 前記周波数領域信号の特性が前記入力信号のエネルギーの変化であることを特徴とする請求項18に記載の雑音抑圧方法。
- [20] 前記第二のフレーム毎に、周波数方向の区切り位置を生成する周波数区切り位置生成ステップを有し、
前記代表周波数領域信号生成ステップは、前記第二のフレームおよび前記周波数方向の区切り位置に基づいて、前記周波数領域信号から代表周波数領域信号を生成することを特徴とする請求項13から請求項19のいずれかに記載の雑音抑圧方法。
- [21] 前記フレーム生成ステップは、一定区間内の前記第二のフレームの個数があらかじめ定められた個数の範囲内になるように前記第二のフレームを生成することを特徴とする請求項13から請求項20のいずれかに記載の雑音抑圧方法。
- [22] 前記フレーム生成ステップは、一定区間内の前記雑音抑圧の程度を算出する計算回数があらかじめ定められた回数の範囲内になるように、前記第二のフレームと前記周波数方向の区切り位置とを生成することを特徴とする請求項20または請求項21に記載の雑音抑圧方法。
- [23] 前記雑音抑制程度算出ステップにおいて、前記雑音抑圧の程度が雑音抑圧係数

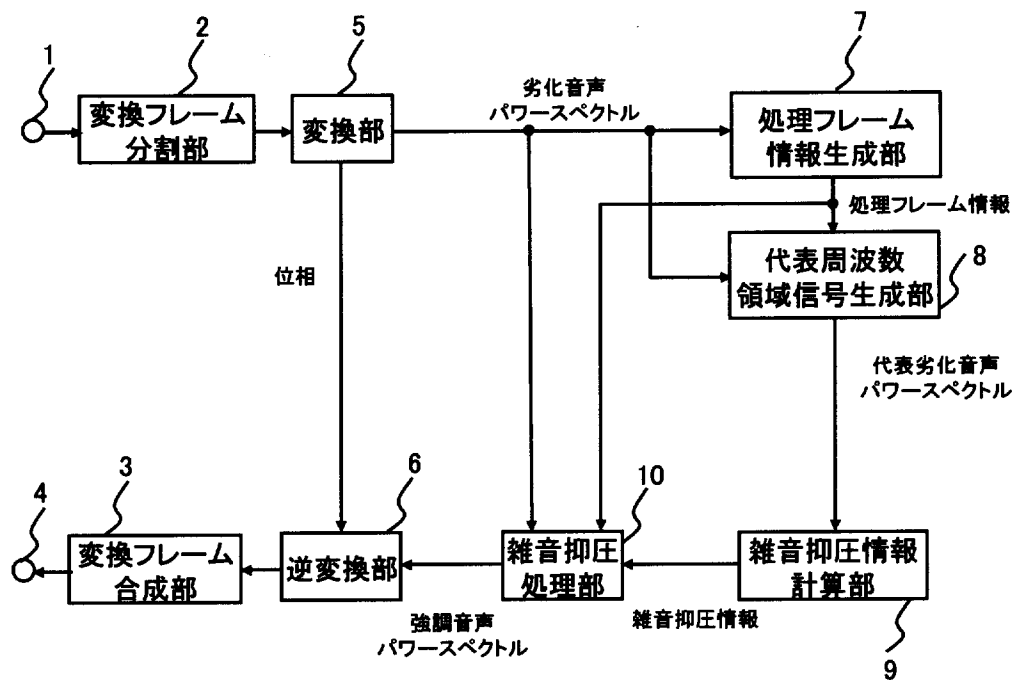
として表されていることを特徴とする請求項13から請求項22のいずれかに記載の雑音抑圧方法。

- [24] 前記雑音抑制程度算出ステップにおいて、前記雑音抑圧の程度が雑音の推定値として表されていることを特徴とする請求項13から請求項22のいずれかに記載の雑音抑圧装置。
- [25] コンピュータに、
入力信号を定められた第一のフレーム毎に周波数領域信号に変換する変換処理と、
前記第一のフレームと異なるように第二のフレームを生成するフレーム生成処理と、
前記第二のフレームに含まれる第一のフレームの前記周波数領域信号から代表周波数領域信号を生成する代表周波数領域信号生成処理と、
前記代表周波数領域信号に基づいて、前記第二のフレームの雑音抑圧の程度を求める雑音抑制程度算出処理と
を実行させる雑音抑圧プログラム。
- [26] 前記フレーム生成処理は、前記第一のフレームの長さより長い前記第二のフレームを生成することを特徴とする請求項25に記載の雑音抑圧プログラム。
- [27] 前記フレーム生成処理は、前記第二のフレーム同士が互いに独立するように前記第二のフレームを生成することを特徴とする請求項25又は請求項26に記載の雑音抑圧プログラム。
- [28] 前記雑音抑制程度算出処理は、前記第二のフレームに含まれる前記周波数領域信号に前記雑音抑圧の程度を適用して雑音を抑圧することを特徴とする請求項25から請求項27のいずれかに記載の雑音抑圧プログラム。
- [29] 前記雑音抑制程度算出処理は、前記第二のフレームに含まれる前記周波数領域信号に、他の第二のフレームの前記雑音抑圧の程度を補間して算出した雑音抑圧の程度を適用して雑音を抑圧することを特徴とする請求項25から請求項28のいずれかに記載の雑音抑圧プログラム。
- [30] 前記フレーム生成処理は、前記周波数領域信号の特性に基づいて、前記第二のフレームを生成することを特徴とする請求項25から請求項29のいずれかに記載の雑

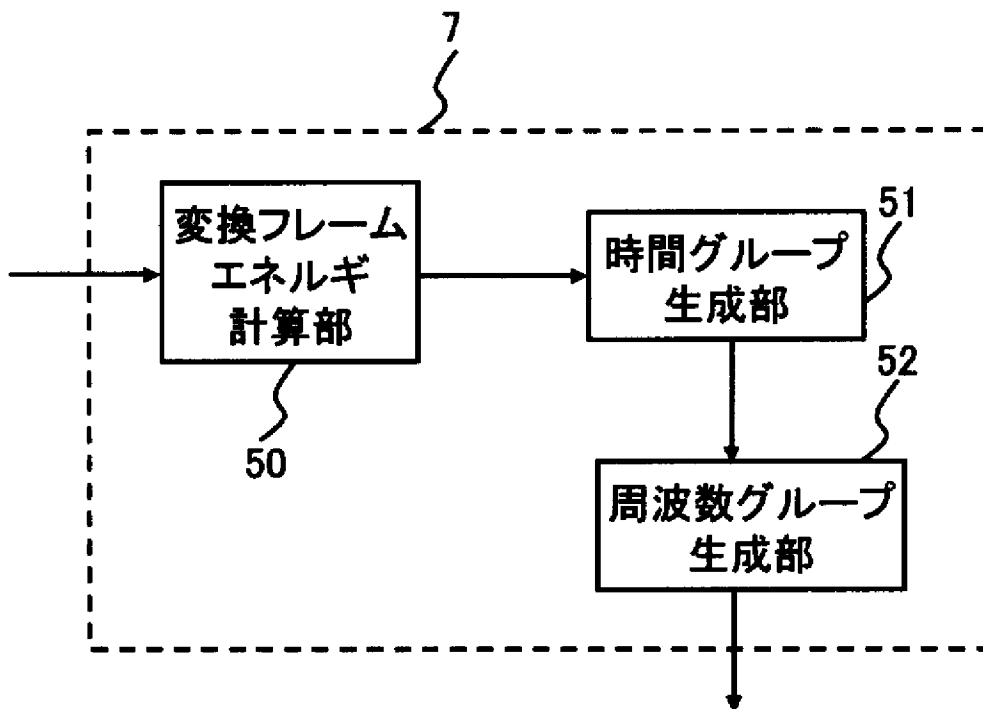
音抑圧プログラム。

- [31] 前記周波数領域信号の特性が前記入力信号のエネルギーの変化であることを特徴とする請求項30に記載の雑音抑圧プログラム。
- [32] 前記第二のフレーム毎に、周波数方向の区切り位置を生成する周波数区切り位置生成処理を有し、
前記代表周波数領域信号生成処理は、前記第二のフレームおよび前記周波数方向の区切り位置に基づいて、前記周波数領域信号から代表周波数領域信号を生成することを特徴とする請求項25から請求項31のいずれかに記載の雑音抑圧プログラム。
- [33] 前記フレーム生成処理は、一定区間内の前記第二のフレームの個数があらかじめ定められた個数の範囲内になるように前記第二のフレームを生成することを特徴とする請求項25から請求項32のいずれかに記載の雑音抑圧プログラム。
- [34] 前記フレーム生成処理は、一定区間内の前記雑音抑圧の程度を算出する計算回数があらかじめ定められた回数の範囲内になるように前記第二のフレームと前記周波数方向の区切り位置とを生成することを特徴とする請求項32または請求項33に記載の雑音抑圧プログラム。
- [35] 前記雑音抑制程度算出処理において、前記雑音抑圧の程度が雑音抑圧係数として表されていることを特徴とする請求項25から請求項34のいずれかに記載の雑音抑圧プログラム。
- [36] 前記雑音抑制程度算出処理において、前記雑音抑圧の程度が雑音の推定値として表されていることを特徴とする請求項25から請求項34のいずれかに記載の雑音抑圧プログラム。

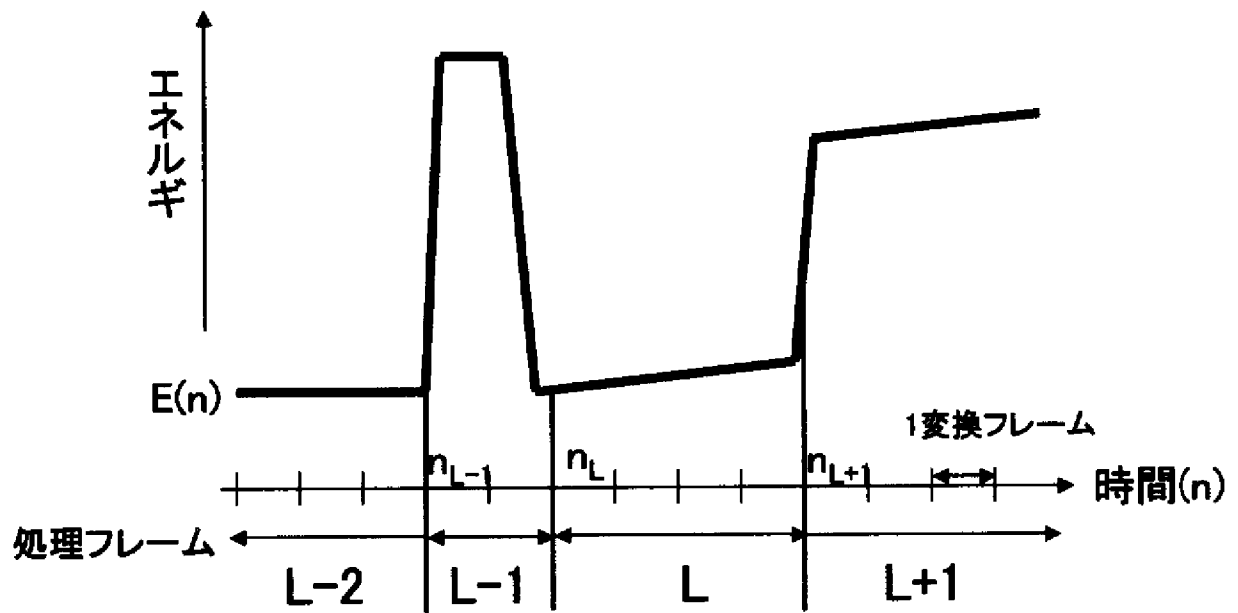
[図1]



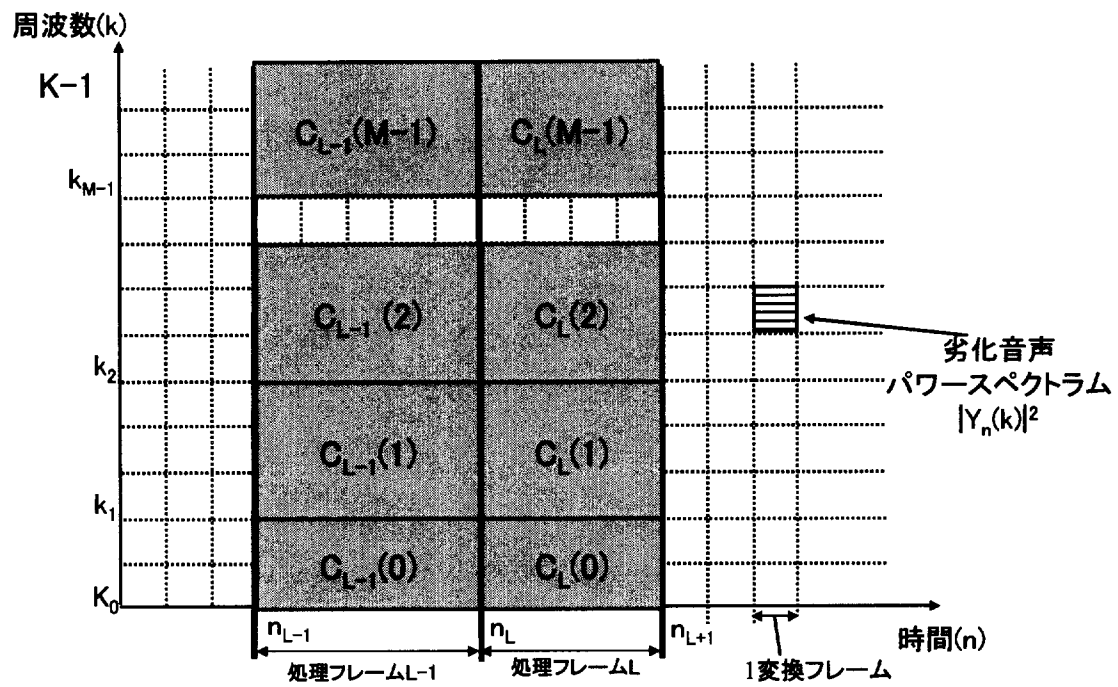
[図2]



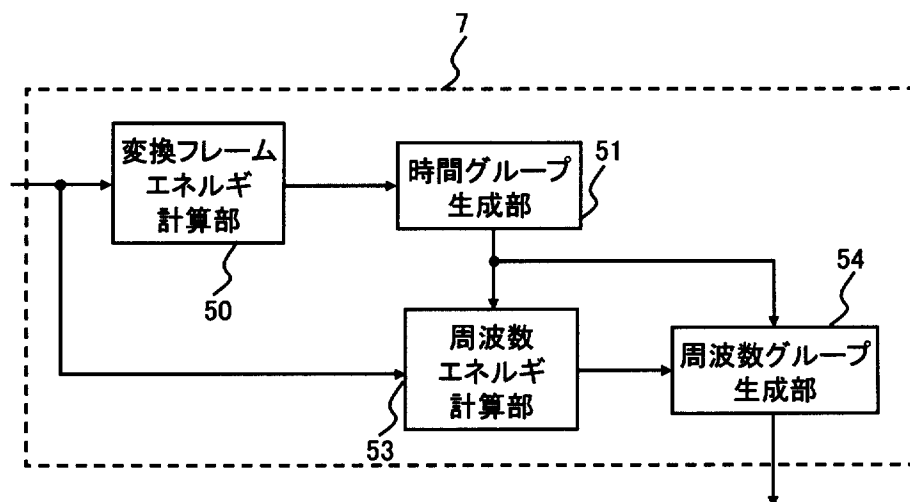
[図3]



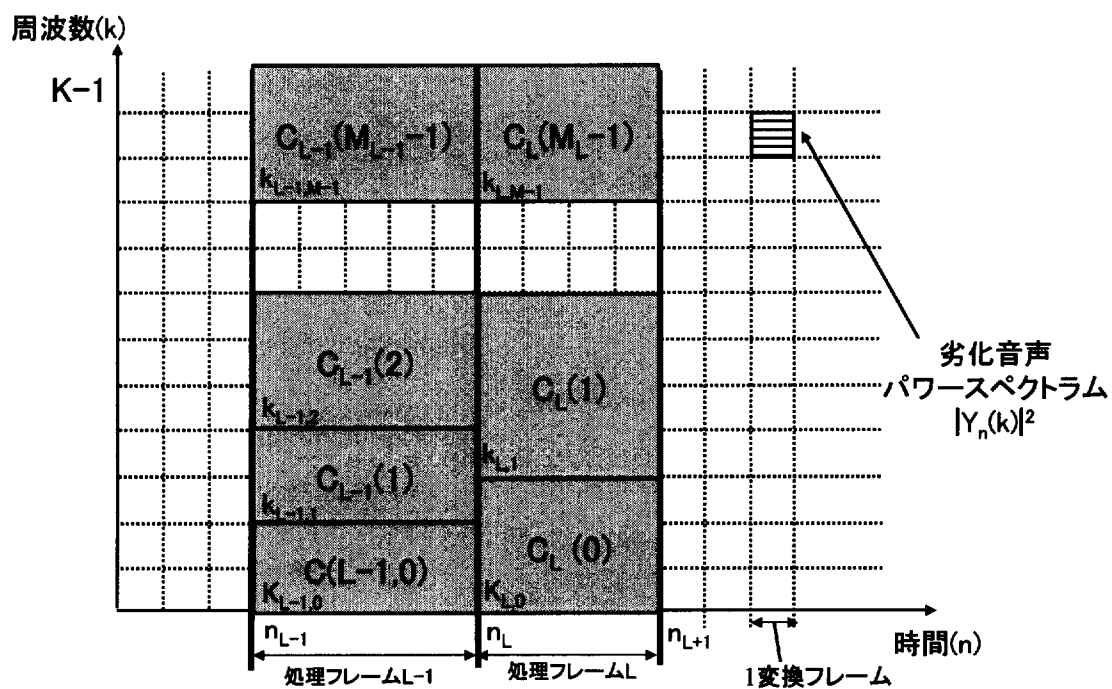
[図4]



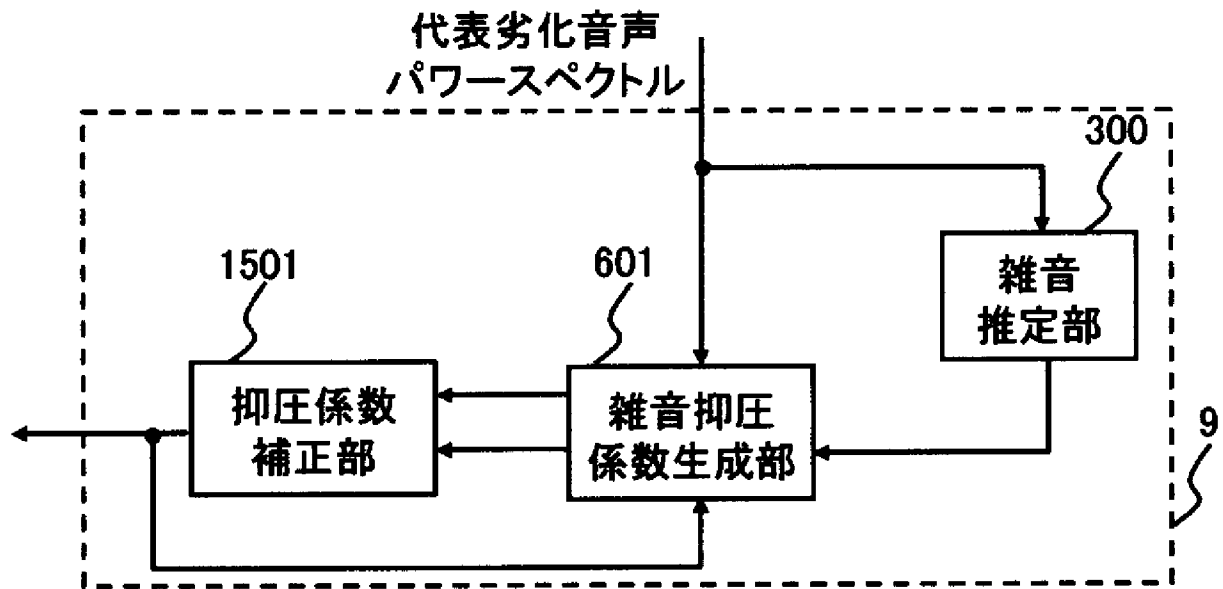
[図5]



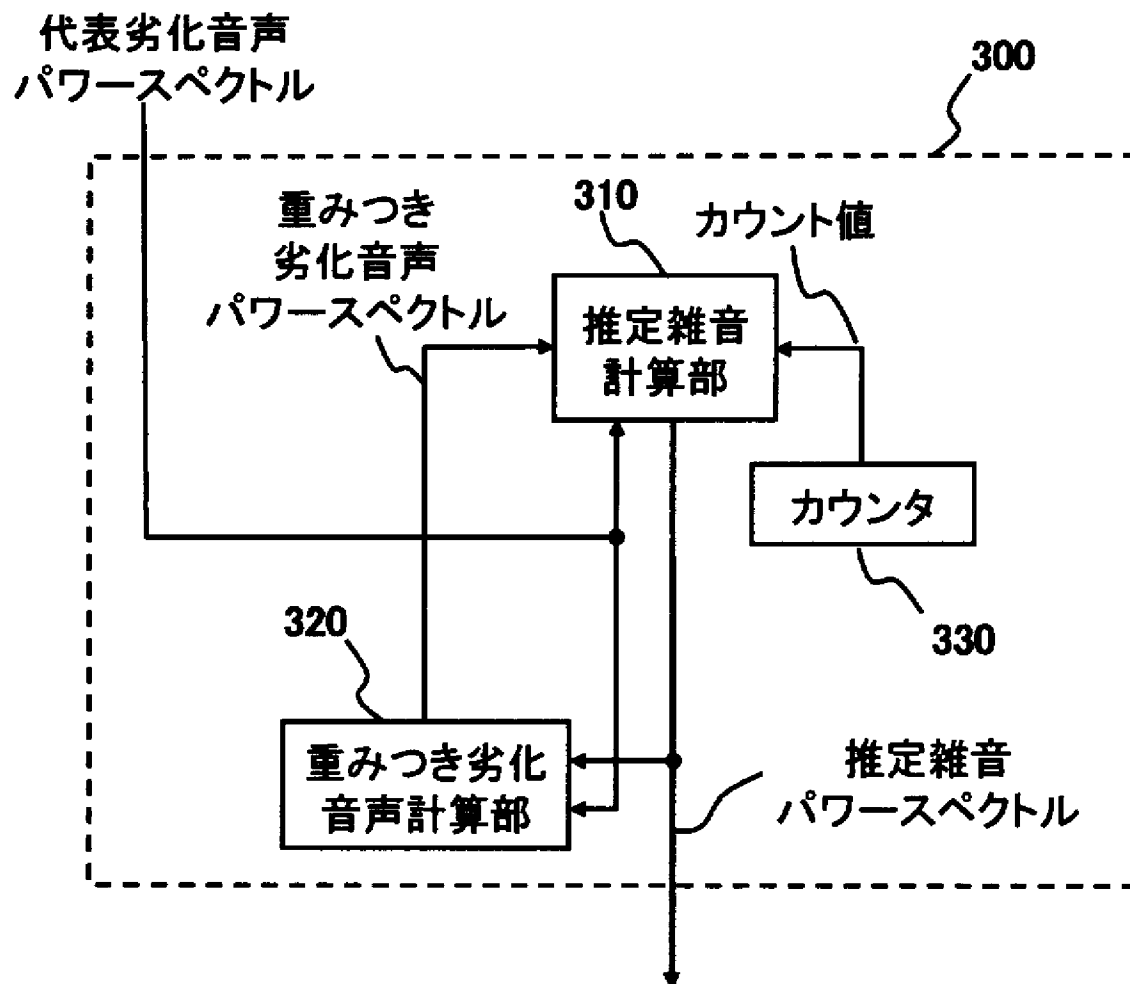
[図6]



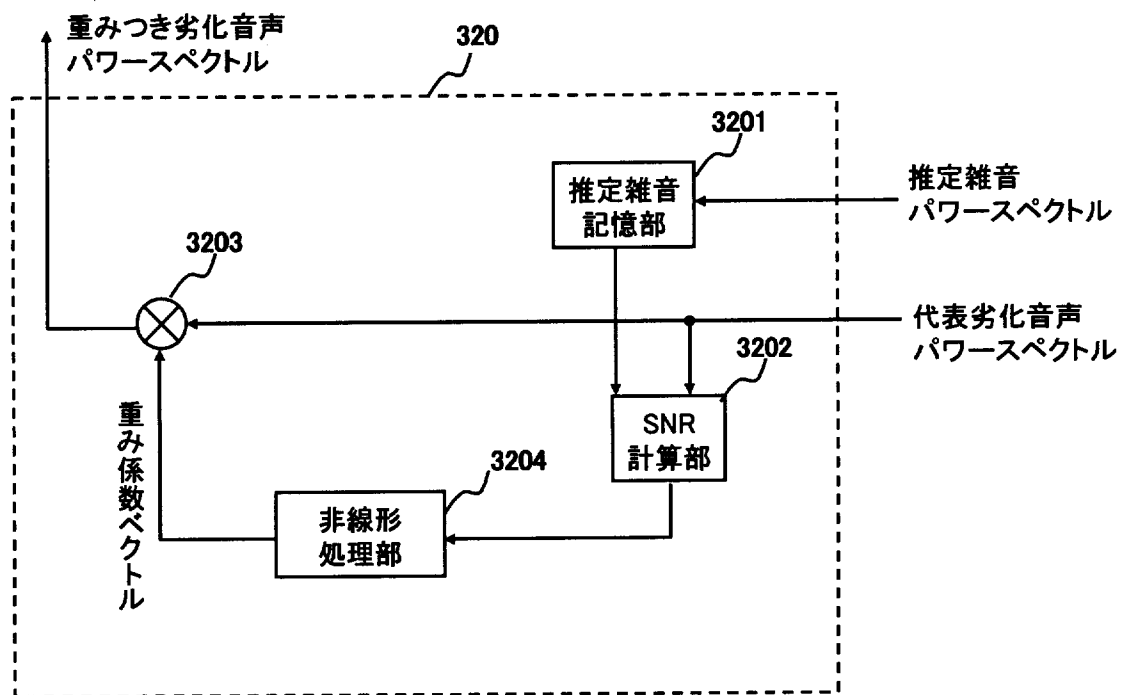
[図7]



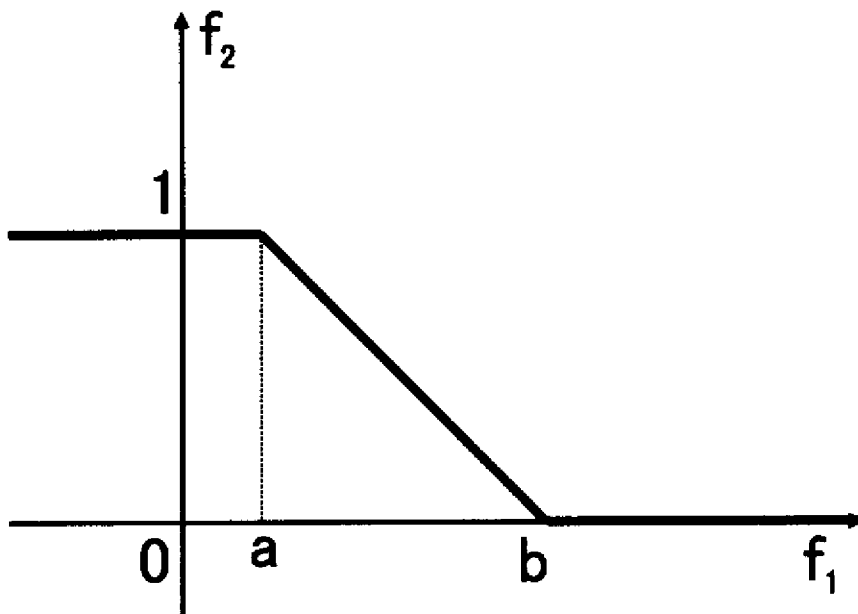
[図8]



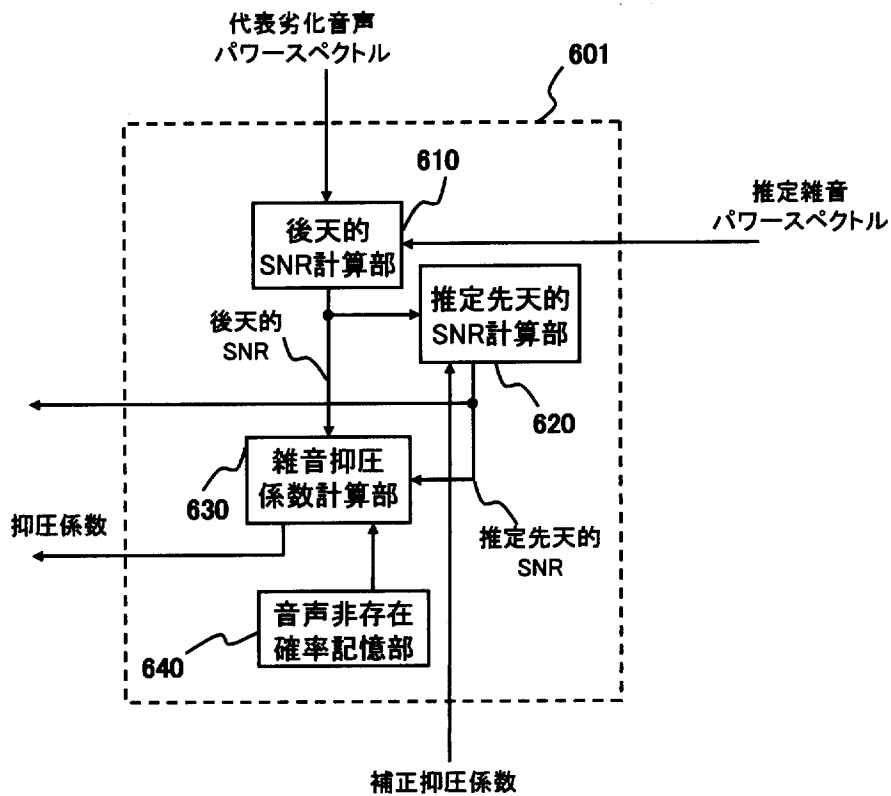
[図11]



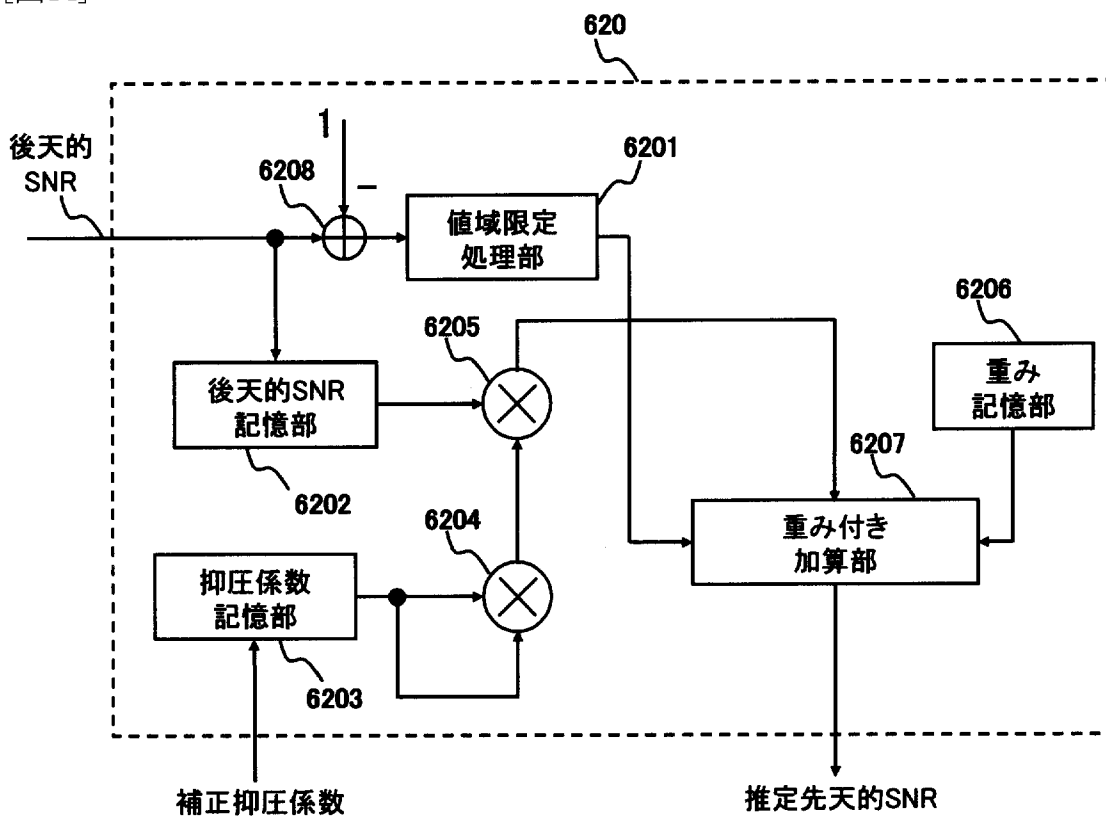
[図12]



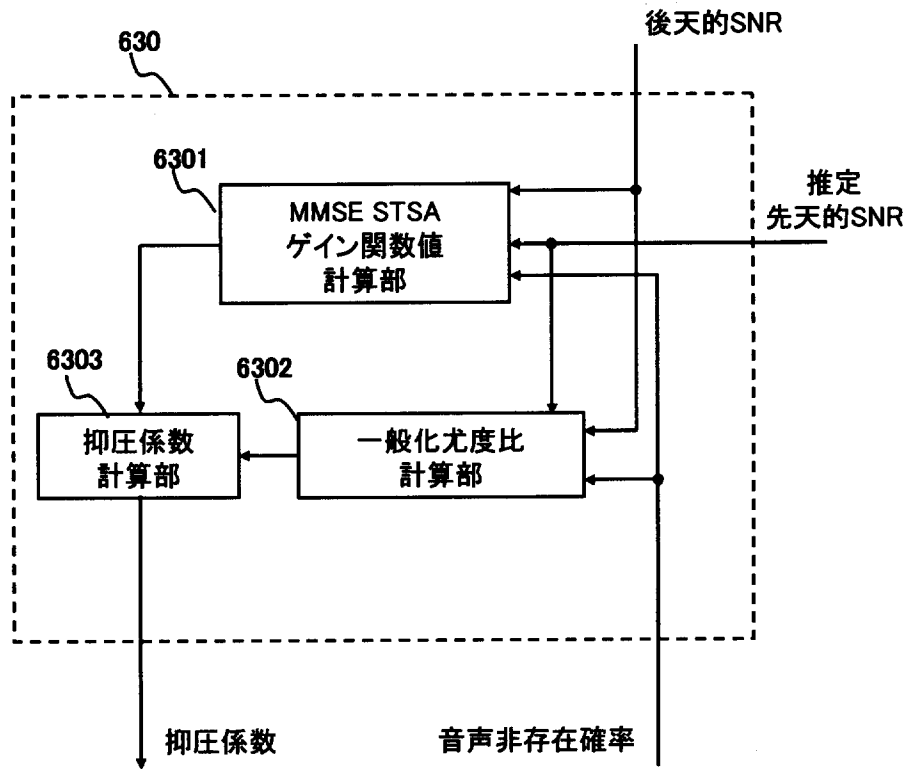
[図13]



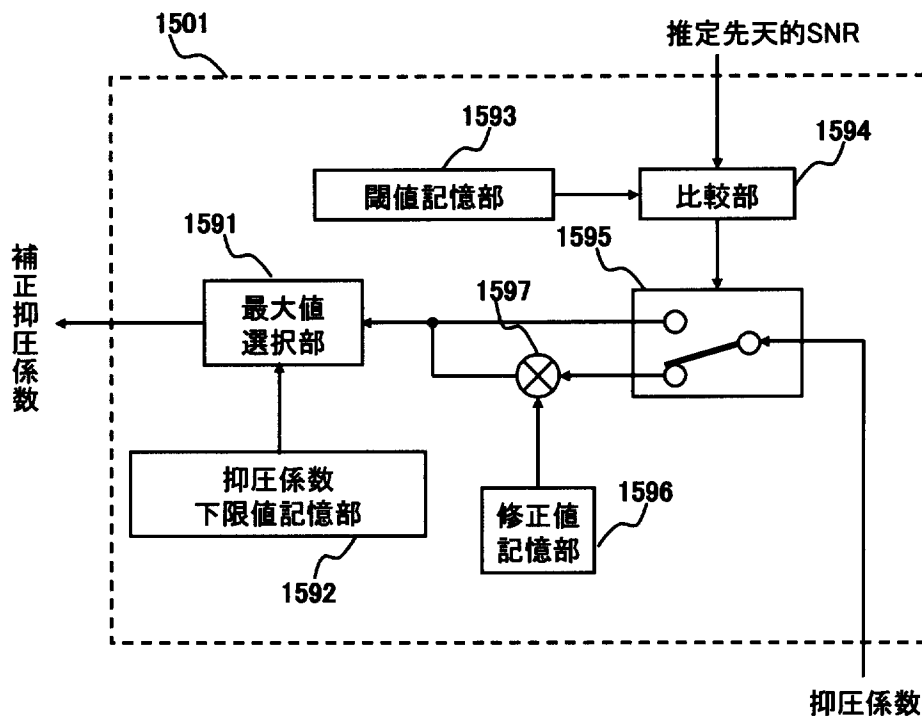
[図14]



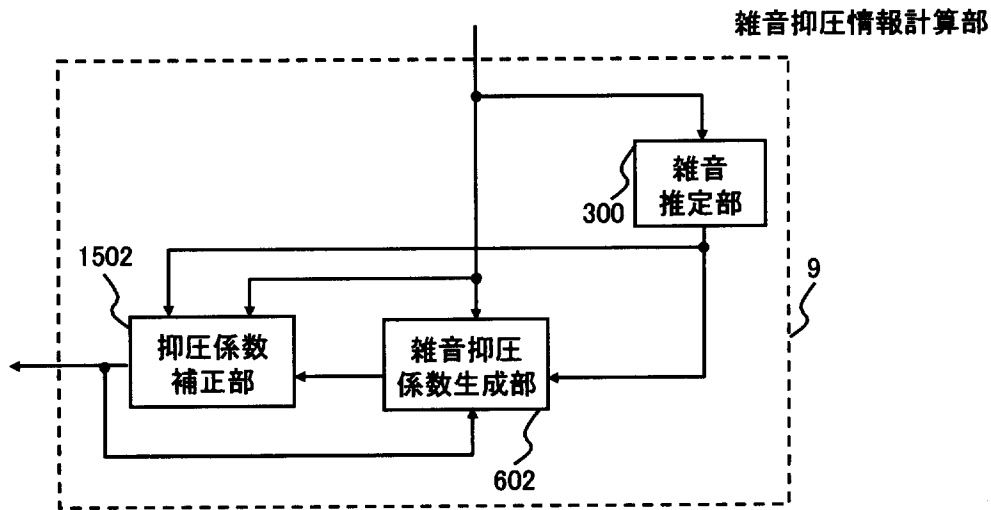
[図15]



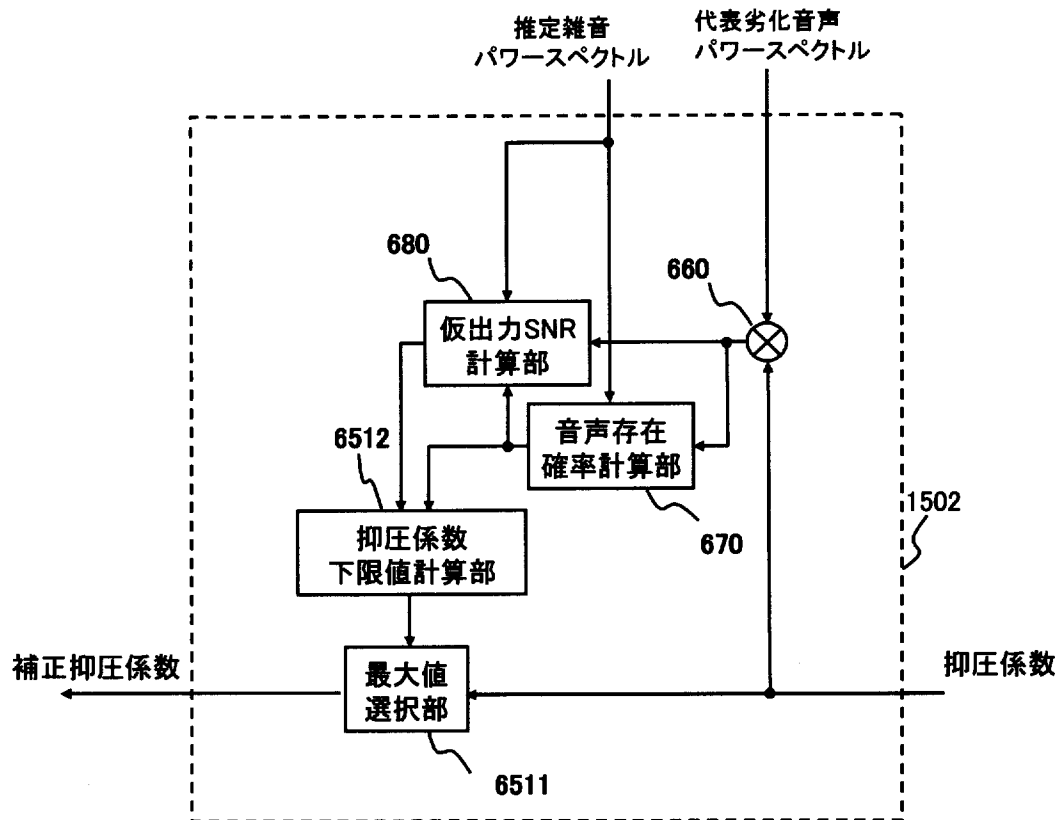
[図16]



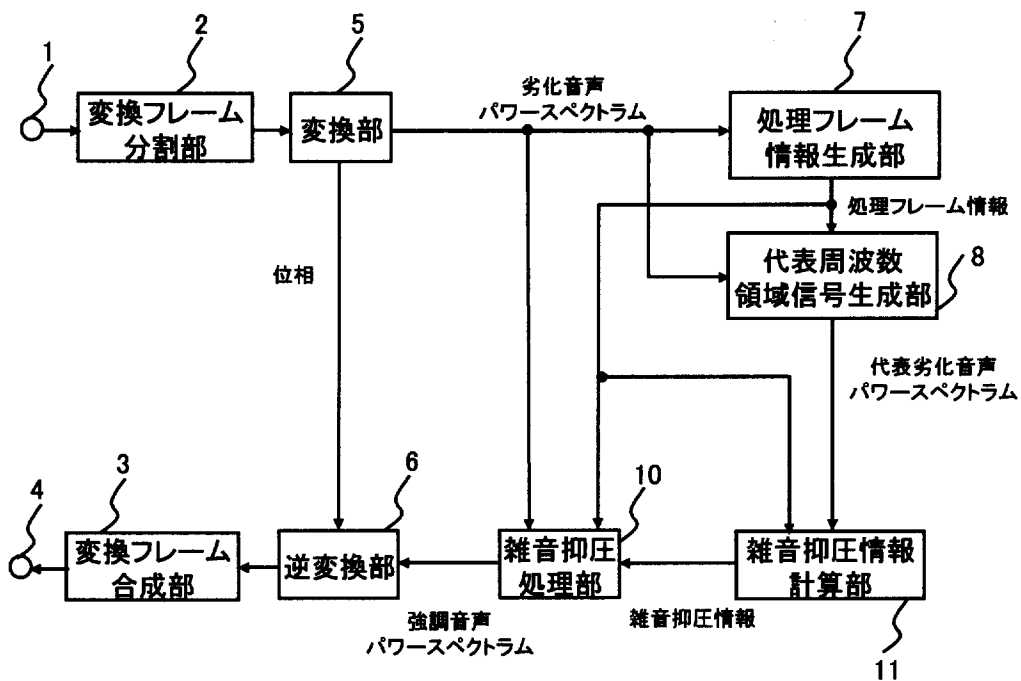
[図17]



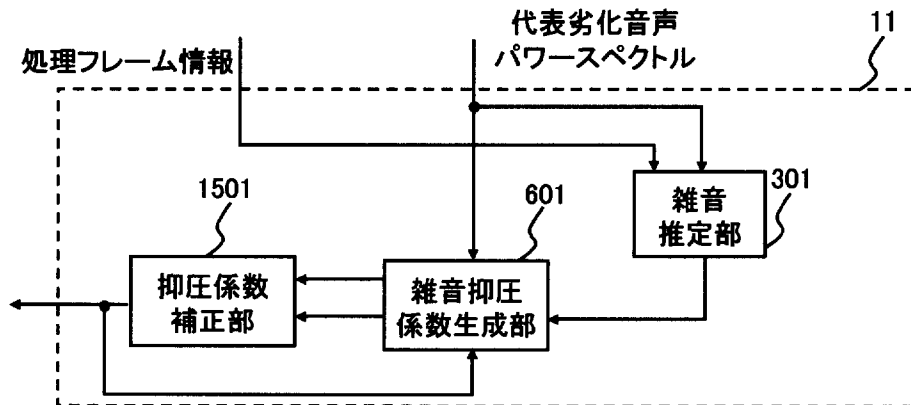
[図18]



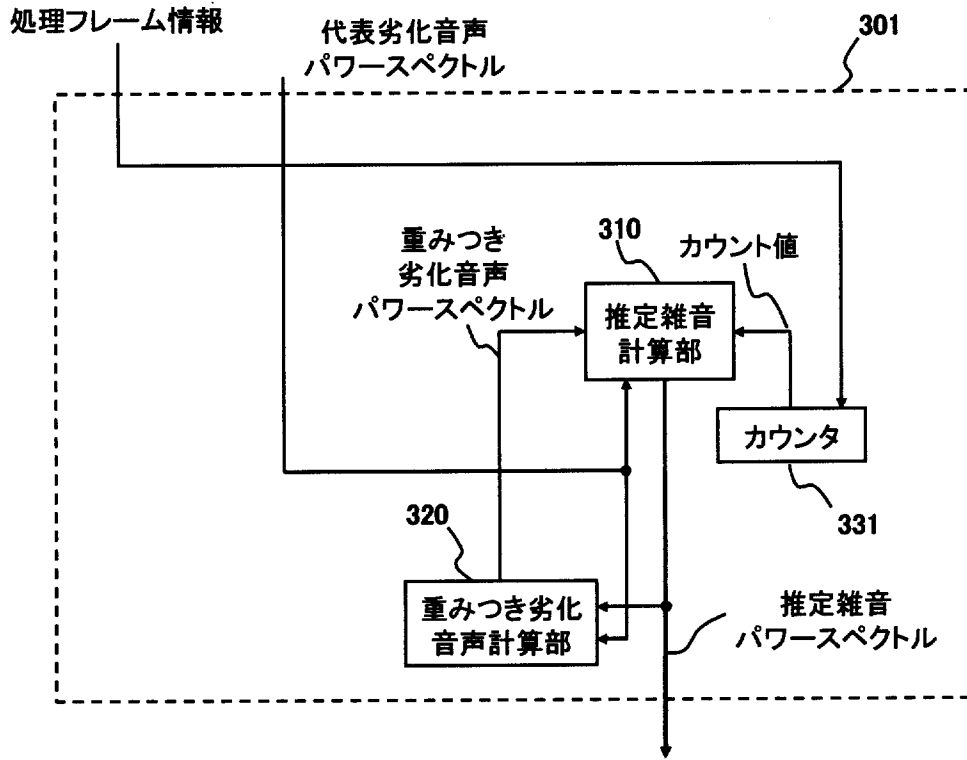
[図19]



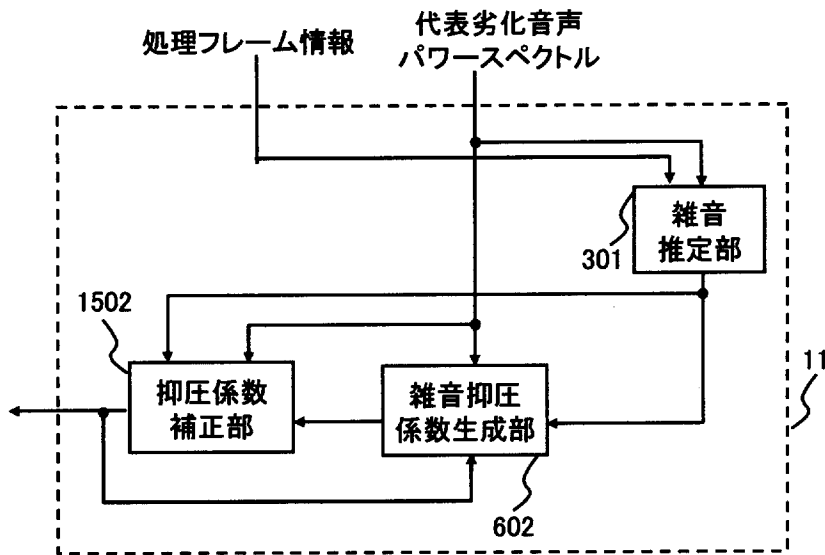
[図20]



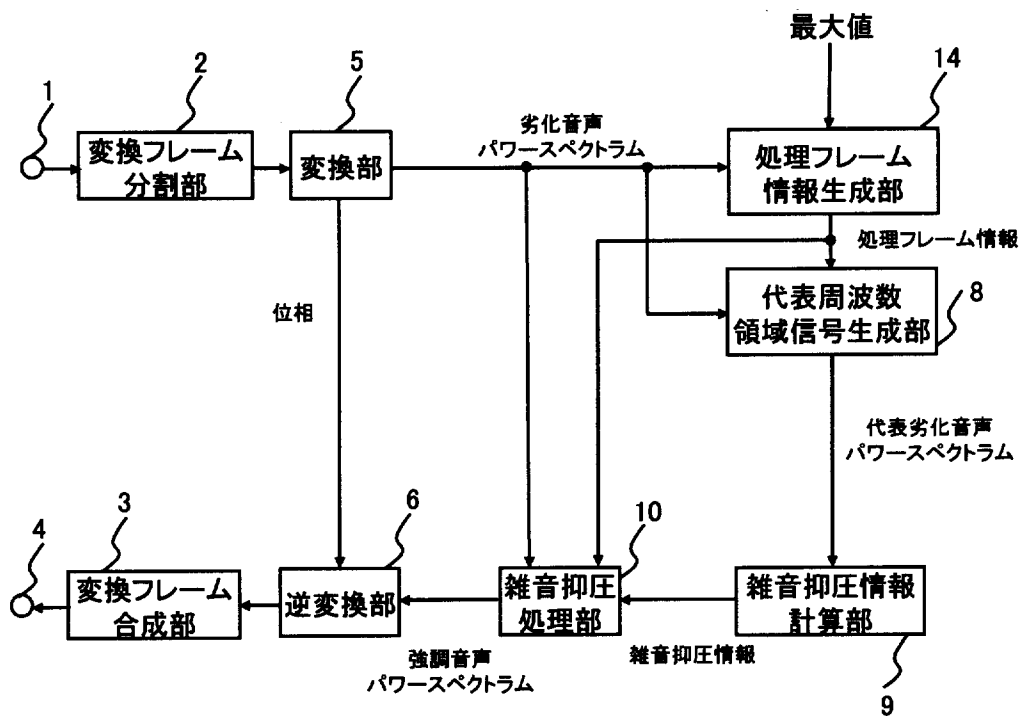
[図21]



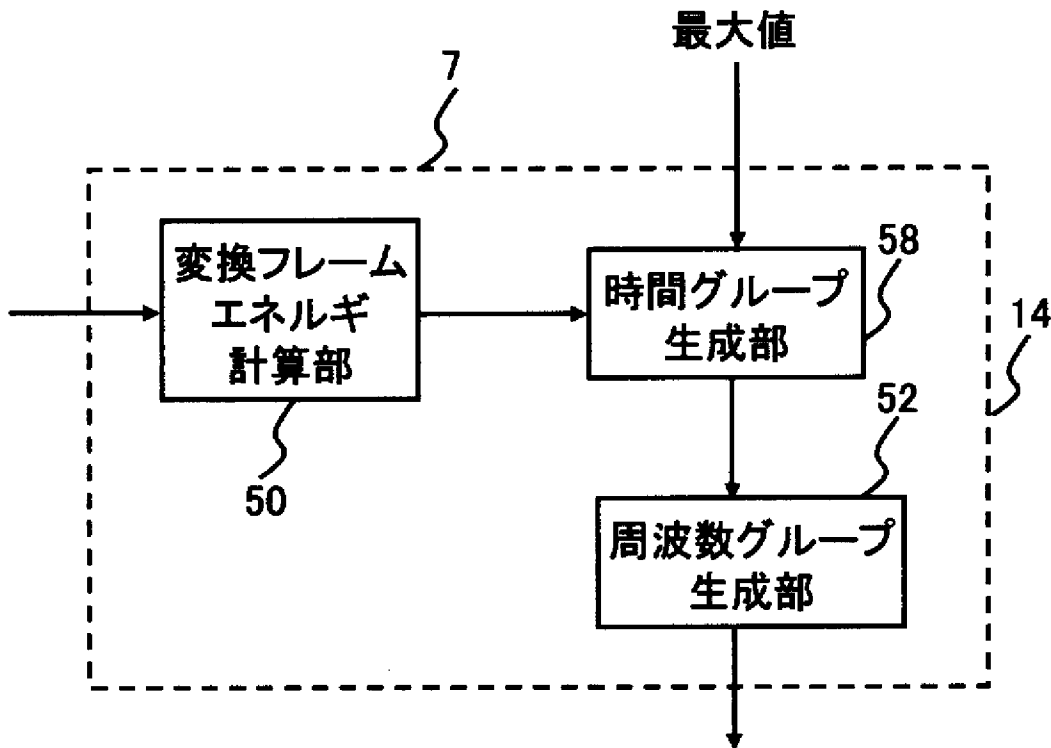
[図22]



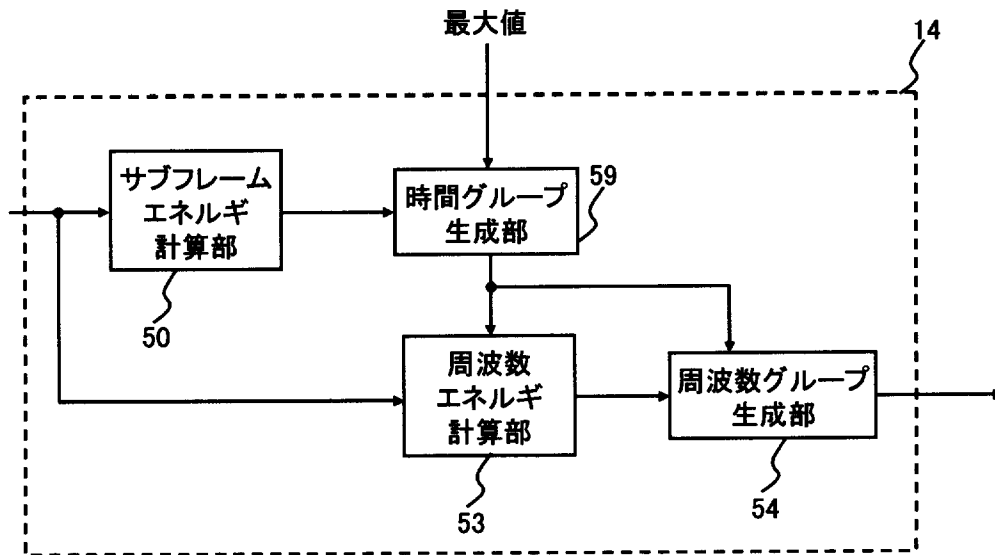
[図23]



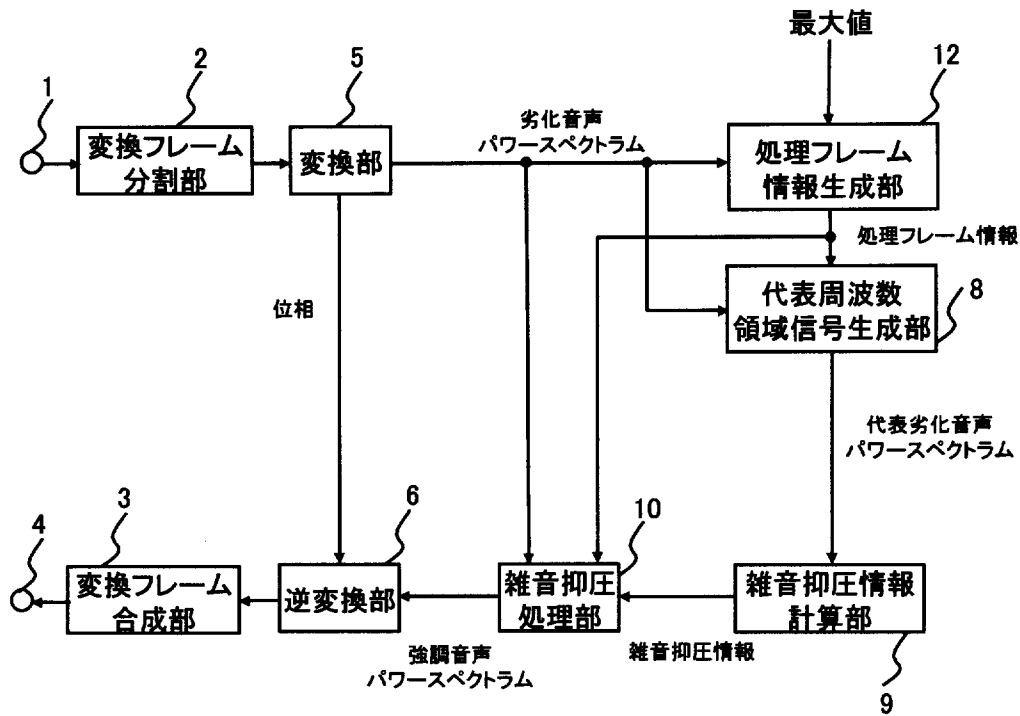
[図24]



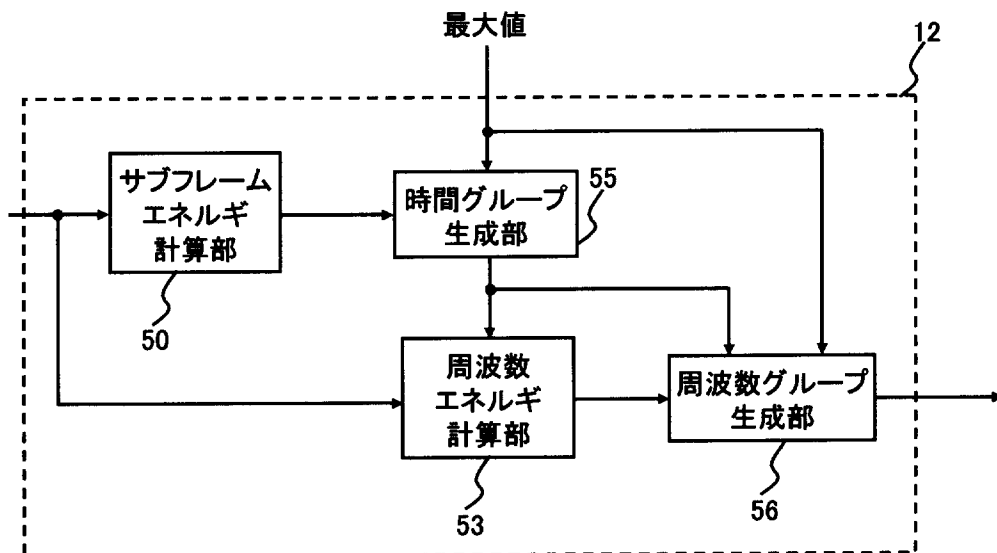
[図25]



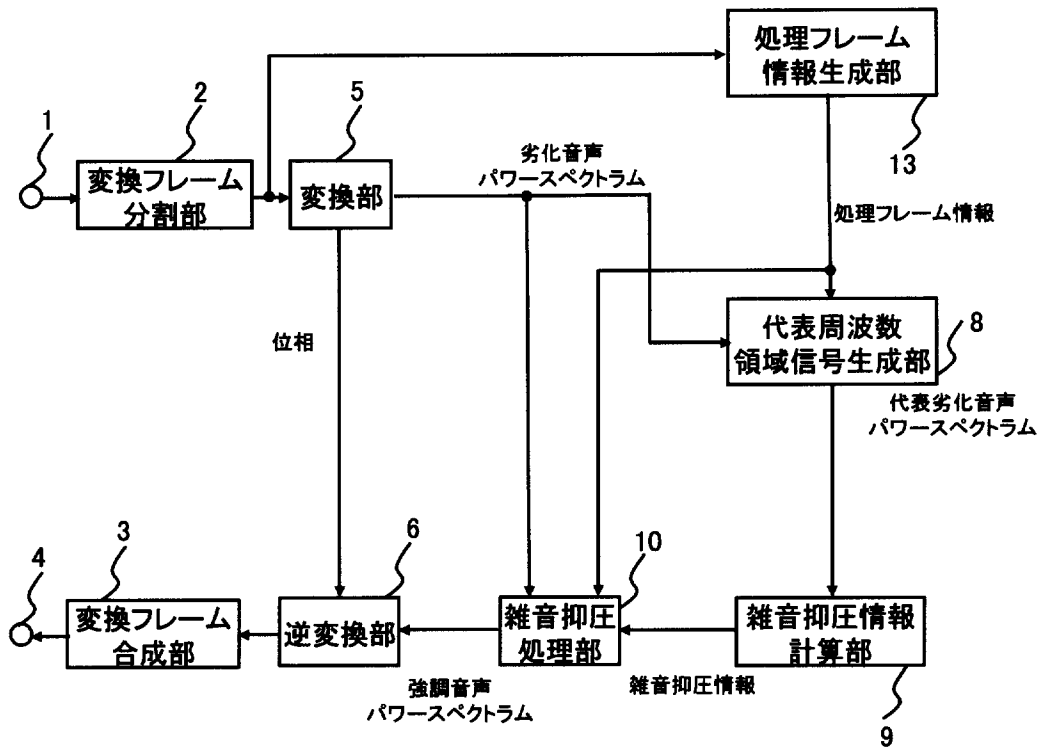
[図26]



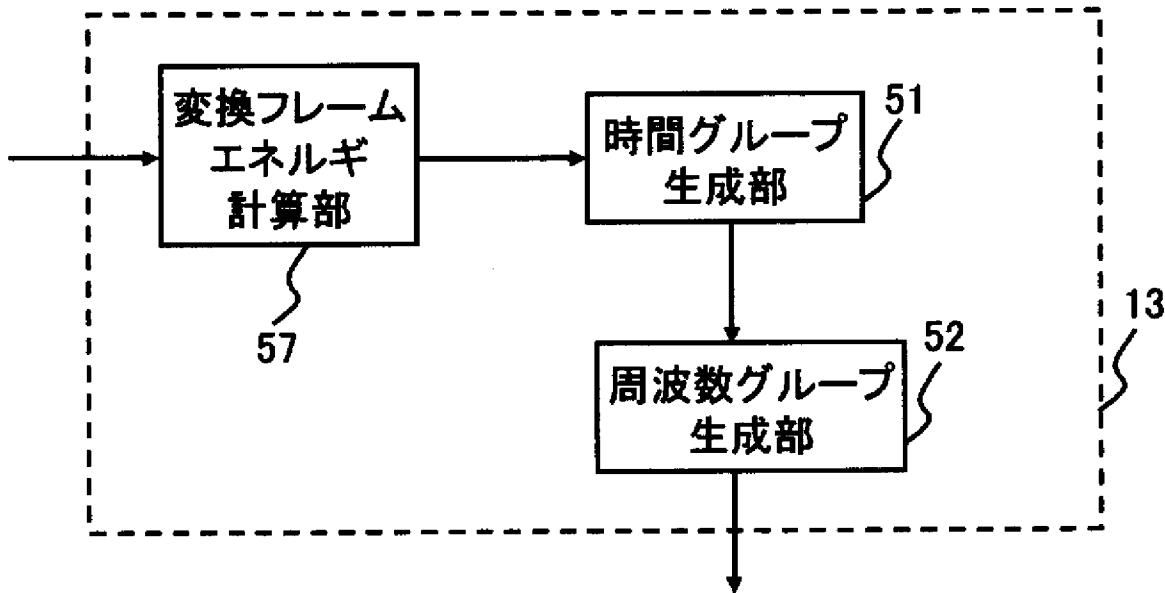
[図27]



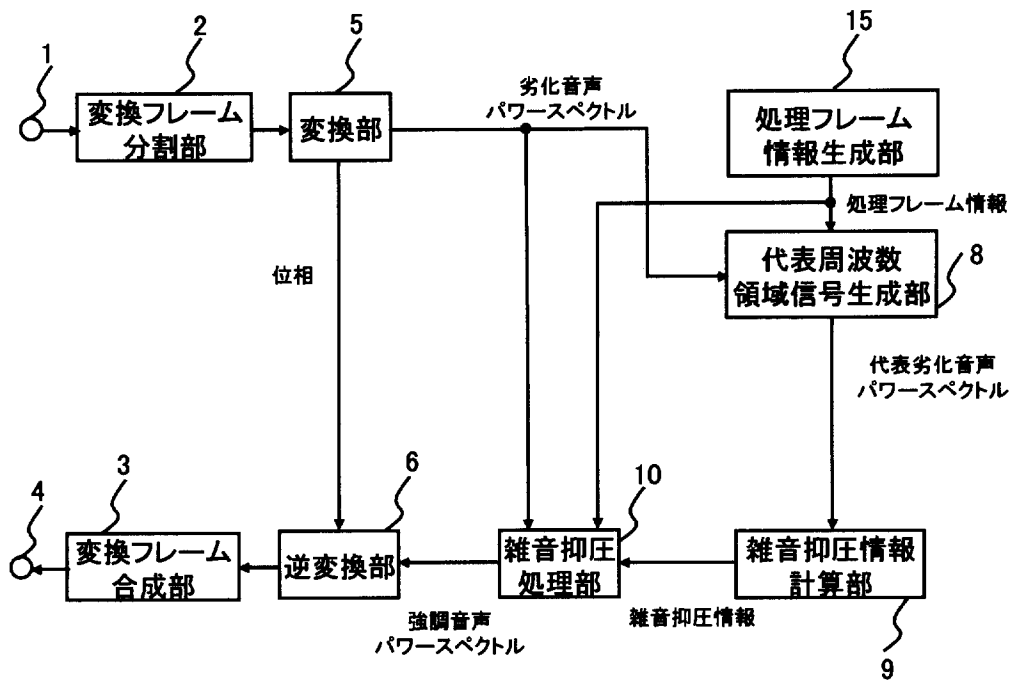
[図28]



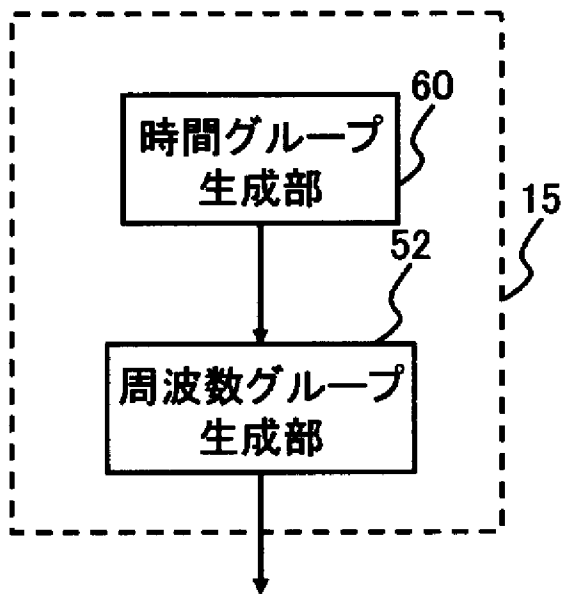
[図29]



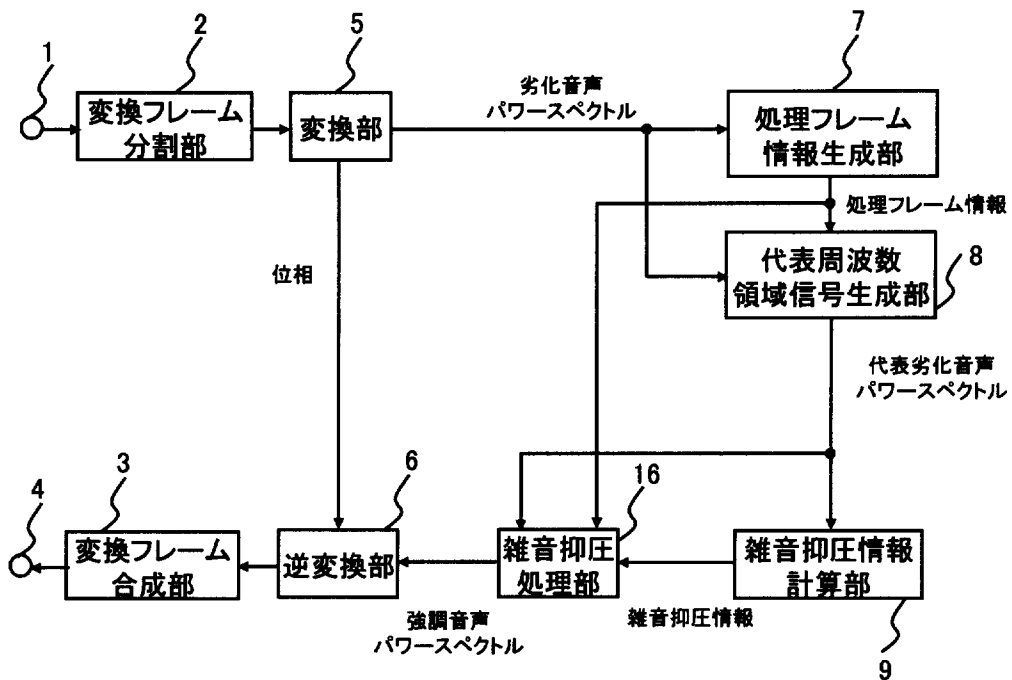
[図30]



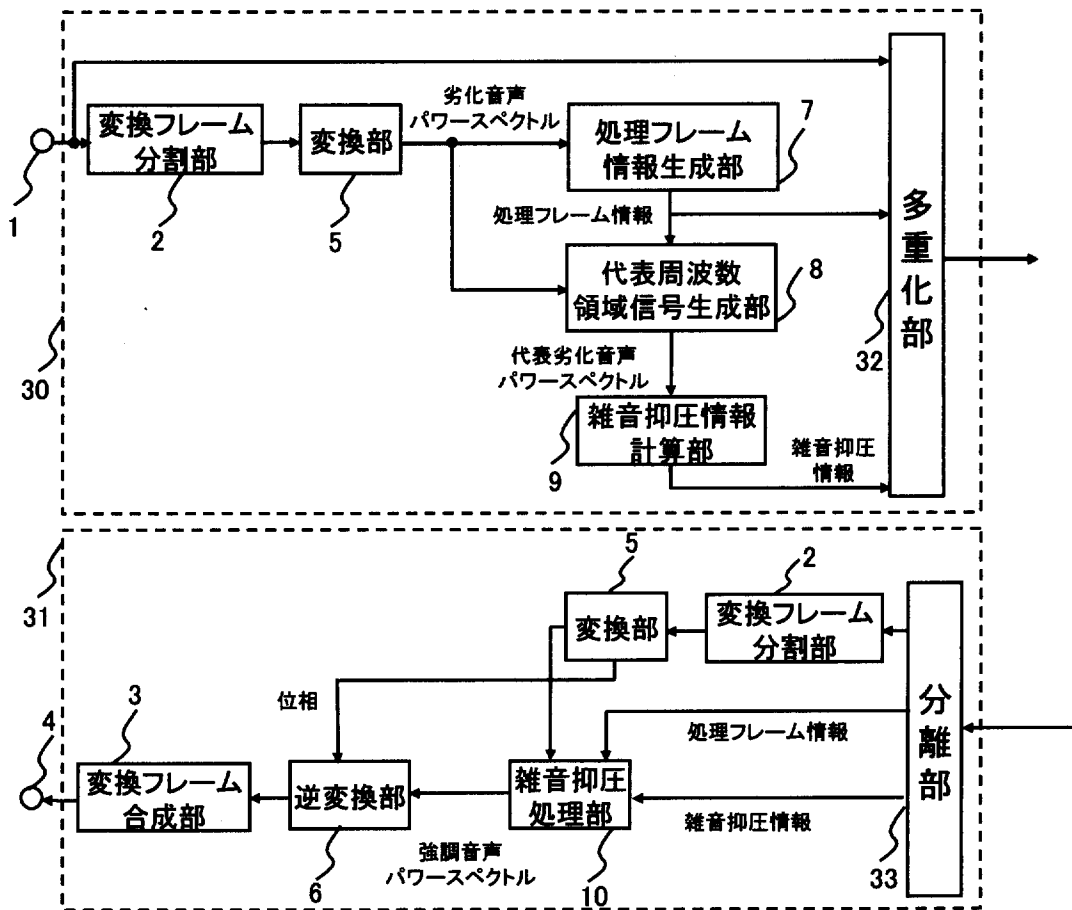
[図31]



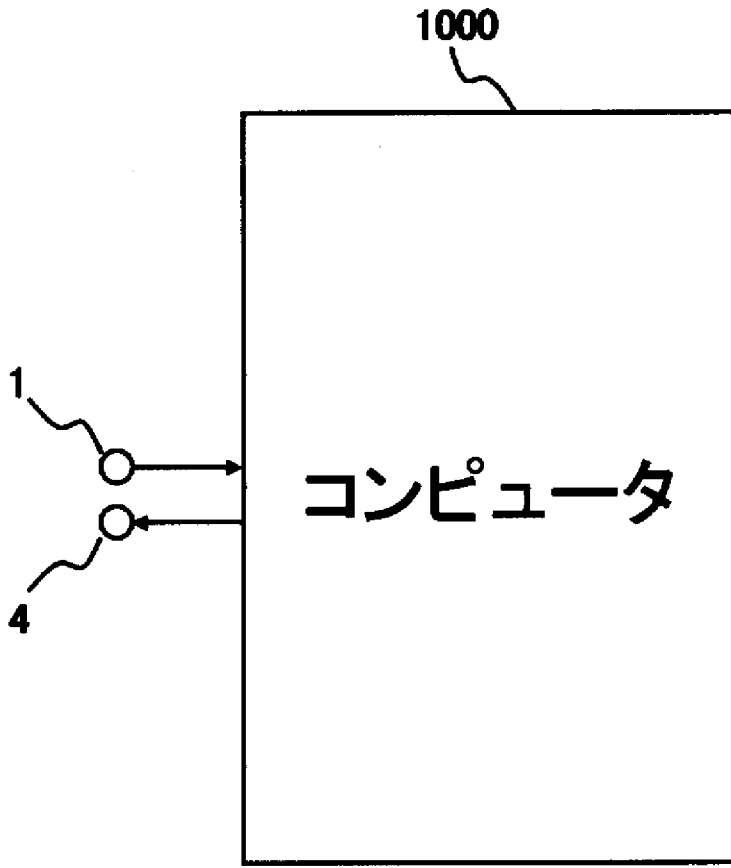
[図32]



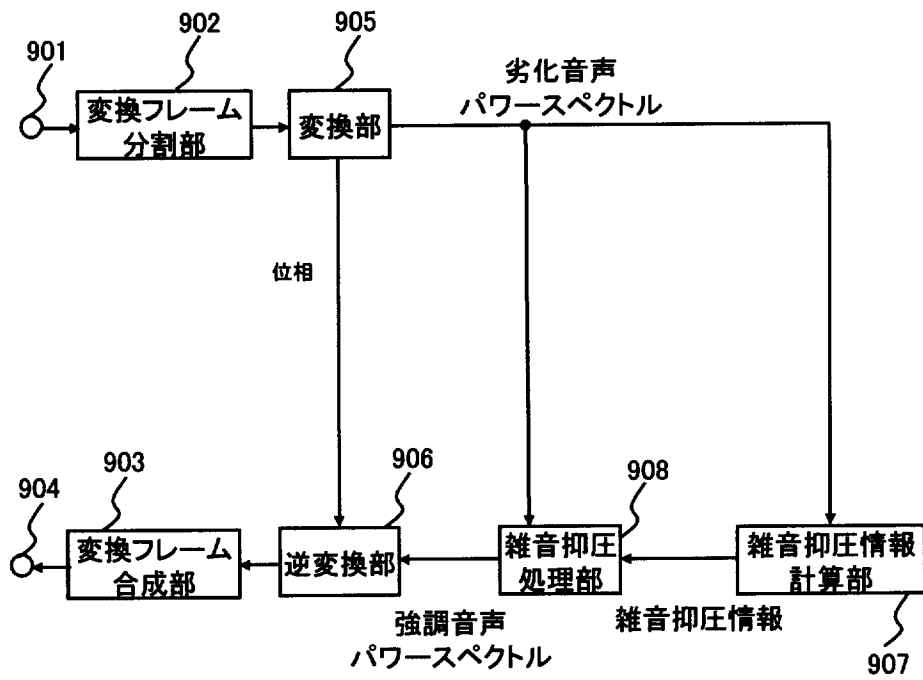
[図33]



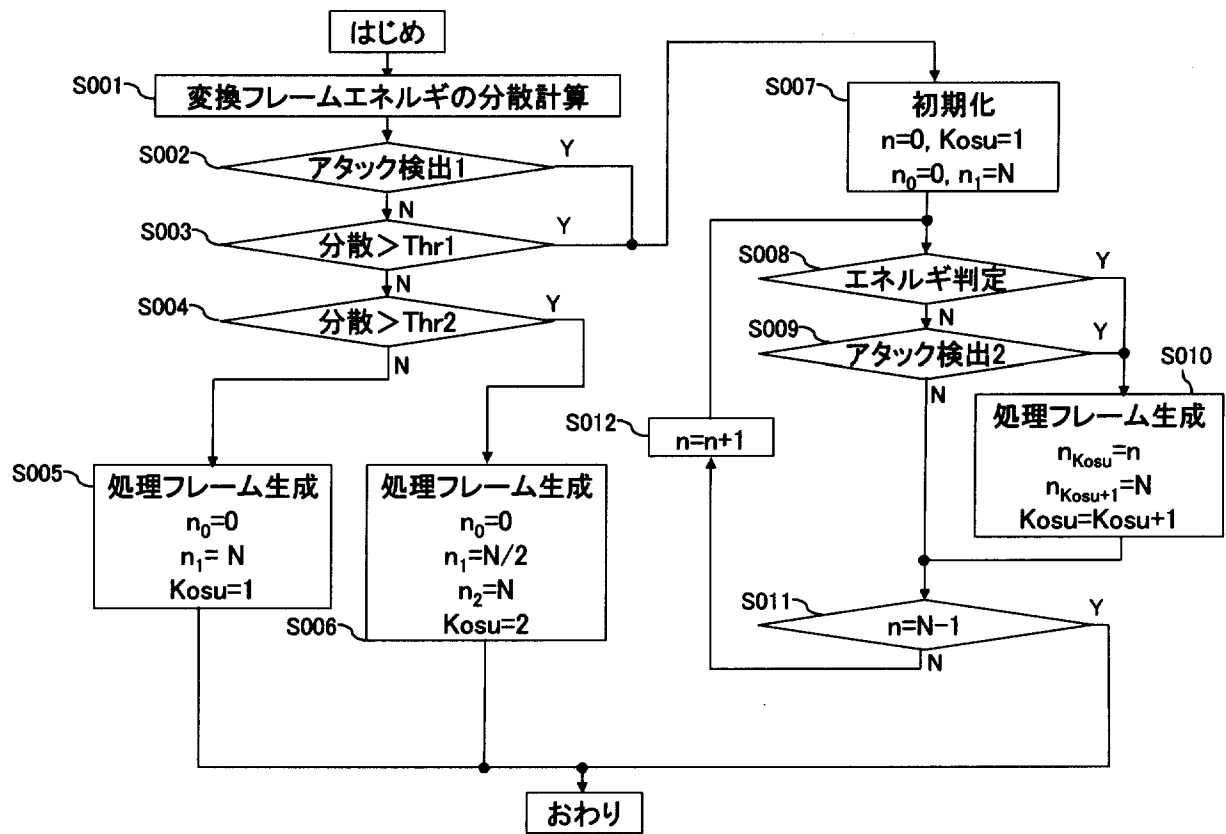
[図34]



[図35]



[図36]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2008/066871

A. CLASSIFICATION OF SUBJECT MATTER

G10L21/02(2006.01)i, G10L15/20(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L15/20, 21/02

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2008

Kokai Jitsuyo Shinan Koho 1971-2008 Toroku Jitsuyo Shinan Koho 1994-2008

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2007-110532 A (Sharp Corp.), 26 April, 2007 (26.04.07), Full text; all drawings (Family: none)	1-36
A	JP 2006-84974 A (Nissan Motor Co., Ltd.), 30 March, 2006 (30.03.06), Full text; all drawings (Family: none)	1-36
A	JP 2005-284016 A (Iwatsu Electric Co., Ltd.), 13 October, 2005 (13.10.05), Full text; all drawings (Family: none)	1-36

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
19 December, 2008 (19.12.08)Date of mailing of the international search report
06 January, 2009 (06.01.09)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2008/066871

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2000-330587 A (NEC Corp.), 30 November, 2000 (30.11.00), Full text; all drawings (Family: none)	1-36
A	JP 2003-337594 A (International Business Machines Corp.), 28 November, 2003 (28.11.03), Full text; all drawings & US 2003/0177006 A1	1-36

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. G10L21/02(2006.01)i, G10L15/20(2006.01)i

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl. G10L15/20, 21/02

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2008年
日本国実用新案登録公報	1996-2008年
日本国登録実用新案公報	1994-2008年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

WPI

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2007-110532 A (シャープ株式会社) 2007.04.26, 全文, 全図 (ファミリーなし)	1-36
A	JP 2006-84974 A (日産自動車株式会社) 2006.03.30, 全文, 全図 (ファミリーなし)	1-36
A	JP 2005-284016 A (岩崎通信機株式会社) 2005.10.13, 全文, 全図 (ファミリーなし)	1-36

C欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的な技術水準を示すもの
 「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」口頭による開示、使用、展示等に言及する文献
 「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献
 「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
 「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
 「&」同一パテントファミリー文献

国際調査を完了した日

19.12.2008

国際調査報告の発送日

06.01.2009

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)
 郵便番号100-8915
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

山下 剛史

電話番号 03-3581-1101 内線 3541

5Z

8946

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2000-330587 A (日本電気株式会社) 2000.11.30, 全文, 全図 (ファミリーなし)	1-36
A	JP 2003-337594 A (インターナショナル・ビジネス・マシーンズ・ コーポレーション) 2003.11.28, 全文, 全図 & US 2003/0177006 A1	1-36