(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3588030号
(P3588030)
(45)発行日　平成16年11月10日(2004.11.10)　(24)登録日　平成16年8月20日(2004.8.20)

(51)Int.Cl.[7]　F I
G10L 11/02　G10L 3/00　513B
G10L 11/00　G10L 7/04　A
G10L 15/04

請求項の数 12　(全 20 頁)

(21)出願番号　特願2000-74262(P2000-74262)
(22)出願日　平成12年3月16日(2000.3.16)
(65)公開番号　特開2001-265367(P2001-265367A)
(43)公開日　平成13年9月28日(2001.9.28)
審査請求日　平成14年8月26日(2002.8.26)

(73)特許権者　000006013
三菱電機株式会社
東京都千代田区丸の内二丁目２番３号
(74)代理人　100066474
弁理士　田澤　博昭
(74)代理人　100088605
弁理士　加藤　公延
(72)発明者　池田　卓史
東京都千代田区丸の内二丁目２番３号　三菱電機株式会社内

審査官　渡邊　聡

最終頁に続く

(54)【発明の名称】音声区間判定装置及び音声区間判定方法

(57)【特許請求の範囲】

【請求項１】

所定の区間ごとに入力される入力信号を時間軸信号から周波数軸信号に直交変換して現在の区間の入力信号の周波数スペクトルを求める直交変換手段と、上記直交変換手段により求められた周波数スペクトルを複数の周波数帯域に分割して現在の区間の帯域スペクトルを求める帯域分割手段と、上記帯域分割手段により求められた現在の区間の帯域スペクトルと雑音区間と判定された過去の区間の帯域スペクトルとの比である帯域別Ｓ／Ｎ比を算出するＳ／Ｎ比算出手段と、上記Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比に基づいて入力信号の現在の区間が音声区間又は雑音区間の何れであるかを判定する判定手段とを備えた音声区間判定装置において、上記判定手段は上記Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比の正変動量を算出し、判定処理を実行するに際して、その正変動量を考慮することを特徴とする音声区間判定装置。

【請求項２】

所定の区間ごとに入力される入力信号を時間軸信号から周波数軸信号に直交変換して現在の区間の入力信号の周波数スペクトルを求める直交変換手段と、上記直交変換手段により求められた周波数スペクトルを複数の周波数帯域に分割して現在の区間の帯域スペクトルを求める帯域分割手段と、上記帯域分割手段により求められた現在の区間の帯域スペクトルと雑音区間と判定された過去の区間の帯域スペクトルとの比である帯域別Ｓ／Ｎ比を算出するＳ／Ｎ比算出手段と、上記Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比に基づいて入力信号の現在の区間が音声区間又は雑音区間の何れであるかを判定する判定手段

とを備えた音声区間判定装置において、上記判定手段は上記Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比の負変動量を算出し、判定処理を実行するに際して、その負変動量を考慮することを特徴とする音声区間判定装置。

【請求項３】

判定手段は、入力信号の現在の区間が雑音区間であると判定する場合、入力信号の現在の区間の帯域スペクトルにより過去の雑音区間の帯域スペクトルを更新することを特徴とする請求項１または請求項２記載の音声区間判定装置。

【請求項４】

判定手段は、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比の平均値を算出し、判定処理を実行するに際して、その平均値を考慮することを特徴とする請求項１から請求項３のうちのいずれか１項記載の音声区間判定装置。

【請求項５】

判定手段は、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比のうち上位Ｎ位（Ｎは自然数）の帯域別Ｓ／Ｎ比を検索し、判定処理を実行するに際して、その帯域別Ｓ／Ｎ比を考慮することを特徴とする請求項１から請求項３のうちのいずれか１項記載の音声区間判定装置。

【請求項６】

判定手段は、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比の分散値を算出し、判定処理を実行するに際して、その分散値を考慮することを特徴とする請求項１から請求項５のうちのいずれか１項記載の音声区間判定装置。

【請求項７】

判定手段は、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比が基準値以上である帯域の個数又は割合を求め、判定処理を実行するに際して、その個数又は割合を考慮することを特徴とする請求項１から請求項５のうちのいずれか１項記載の音声区間判定装置。

【請求項８】

判定手段は、判定処理を実行するに際して、入力信号の現在の区間の信号エネルギーを考慮することを特徴とする請求項１から請求項７のうちのいずれか１項記載の音声区間判定装置。

【請求項９】

判定手段は、判定処理を実行するに際して、入力信号の現在の区間の差分信号エネルギーを考慮することを特徴とする請求項１から請求項８のうちのいずれか１項記載の音声区間判定装置。

【請求項１０】

判定手段は、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比にしたがって判定処理に用いる閾値を更新することを特徴とする請求項１から請求項９のうちのいずれか１項記載の音声区間判定装置。

【請求項１１】

所定の区間ごとに入力される入力信号を時間軸信号から周波数軸信号に直交変換して現在の区間の入力信号の周波数スペクトルを求めると、その周波数スペクトルを複数の周波数帯域に分割して現在の区間の帯域スペクトルを求め、その現在の区間の帯域スペクトルと雑音区間と判定された過去の区間の帯域スペクトルとの比である帯域別Ｓ／Ｎ比を算出し、その帯域別Ｓ／Ｎ比に基づいて入力信号の現在の区間が音声区間又は雑音区間の何れであるかを判定する音声区間判定方法において、その帯域別Ｓ／Ｎ比の正変動量を算出し、判定処理を実行するに際して、その正変動量を考慮することを特徴とする音声区間判定方法。

【請求項１２】

所定の区間ごとに入力される入力信号を時間軸信号から周波数軸信号に直交変換して現在の区間の入力信号の周波数スペクトルを求めると、その周波数スペクトルを複数の周波数帯域に分割して現在の区間の帯域スペクトルを求め、その現在の区間の帯域スペクトルと雑音区間と判定された過去の区間の帯域スペクトルとの比である帯域別Ｓ／Ｎ比を算出し

、その帯域別Ｓ／Ｎ比に基づいて入力信号の現在の区間が音声区間又は雑音区間の何れであるかを判定する音声区間判定方法において、その帯域別Ｓ／Ｎ比の負変動量を算出し、判定処理を実行するに際して、その負変動量を考慮することを特徴とする音声区間判定方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

この発明は、種々の雑音環境下で用いられる音声通信システムや音声認識システム等において、所定の区間ごとに入力される入力信号が音声区間であるか又は、雑音区間かを判別する音声区間判定装置及び音声区間判定方法に関するものである。

【０００２】

【従来の技術】

雑音のある環境下において、正確に音声認識を行う必要がある場合、音声区間判定装置が用いられる。

例えば、特開昭５９－９９４９７号公報には、入力信号の直交変換を実施し、その変換値から求めたスペクトルに基づいて入力信号パワーと雑音信号パワーとの比であるＳ／Ｎ比を算出し、そのＳ／Ｎ比と固定閾値を大小比較することにより、音声区間を検出する音声区間判定装置が開示されている。

【０００３】

また、背景雑音と音声信号が混在する入力信号から、音質を損なうことなく背景雑音を除去する背景雑音除去装置においては、入力信号の様態を検出する様態検出手段が用いられ、その様態検出手段は、例えば、特開平１０－１７１４９７号公報に開示されている。

その様態検出手段は、文献（Ｓｔｅｖｅｎ Ｆ．Ｂｏｌｌ，“Ｓｕｐｐｒｅｓｓｉｏｎ ｏｆ Ａｃｏｕｓｔｉｃ ｎｏｉｓｅ ｉｎ ｓｐｅｅｃｈ ｕｓｉｎｇ ｓｐｅｃｔｒａｌ ｓｕｂｔｒａｃｔｉｏｎ”，ＩＥＥＥ Ｔｒａｎｓ．ＡＳＳＰ，Ｖｏｌ．ＡＳＳＰ－２７，Ｎｏ．２，Ａｐｒｉｌ １９７９）に示された振幅スペクトルに着目した方法、即ち、スペクトルサブトラクション（Ｓｐｅｃｔｒａｌ Ｓｕｂｔｒａｃｔｉｏｎ：ＳＳ）法に基づく方法である。

【０００４】

具体的には、周波数軸上において、背景雑音と音声信号を含む入力信号から、過去の背景雑音を平均して求めた背景雑音の推定値を減算することにより、入力信号中に含まれる背景雑音の定常成分を取り除き、減算後の信号と背景雑音の推定値から、全周波数帯域のＳ／Ｎ比（ＳＮ＿ａｌｌ）と、全周波数帯域を複数の帯域に分割した小帯域毎のＳ／Ｎ比（ＳＮ＿ｃｈ（ｉ），ｉはチャネル番号））とを計算する。

そして、これらの差分（ＳＮ＿ｃｈ（ｉ）－ＳＮ＿ａｌｌ）が所定値以下の帯域では、背景雑音の急峻な変化に対応する背景雑音を推定し、この推定値を用いてさらに減算処理を行うことにより、背景雑音の非定常成分を取り除くものである。

【０００５】

また、音声区間の検出を実行する際に固定閾値との比較を行うのではなく、入力信号の様態に応じて、閾値を随時更新し、更新後の閾値との比較を行う手段が、例えば、特開平２－２７２８３６号公報に開示されている。

これは、入力信号のパワーを計算し、このパワー値，零交差回数，最大零交差間隔及び過去の音声区間の判定結果に基づいて閾値の更新を随時実行し、現フレームのパワー値と更新後の閾値とを比較することにより、音声区間の検出を行うものである。

【０００６】

【発明が解決しようとする課題】

従来の音声区間判定装置は以上のように構成されているので、音声区間の判別精度が悪く、音声区間を雑音区間と誤判定する課題があった。

具体的には、下記に示す理由から誤判定を招いている。

【０００７】

特開昭５９－９９４９７号公報に記載されている音声区間判定装置は、入力信号パワーと雑音信号パワーとの比であるＳ／Ｎ比（全周波数帯域のＳ／Ｎ比）を用いて、音声区間判定を実施している。図16は帯域別Ｓ／Ｎ比（全周波数が複数の小帯域に分割された小帯域毎のＳ／Ｎ比）の一例を表しており、（ａ）は雑音区間、（ｂ）は音声区間である。図１６において、帯域別Ｓ／Ｎ比の平均値は（ａ），（ｂ）ともに等しく（$\mu_1 = \mu_2$）、閾値ＴＨとの大小比較により音声区間判定を行うと、帯域別Ｓ／Ｎ比の平均値μ_1，μ_2が閾値ＴＨよりも小さいため、どちらの場合においても「雑音区間（非音声区間）」と判定される。

【０００８】

しかし、帯域別Ｓ／Ｎ比が図１６（ｂ）に示すものである場合、低域において、帯域別Ｓ／Ｎ比が平均値μ_2よりも高い帯域の数が、図１６（ａ）に示すものよりも多く、また、これらの帯域では平均値μ_2との差が大きく、本来なら「音声区間」であると判定されるべきである。ところが、従来の音声区間判定装置では、この「音声区間」を「雑音区間」と誤判定してしまう。

【０００９】

また、従来の音声区間判定装置では、入力信号が音声区間であるか否かを判定する際に使用する閾値ＴＨが固定値であるため、背景雑音の定常性を仮定して、閾値ＴＨを設定する場合が多い。このようにして、閾値ＴＨを設定すると、入力信号中の雑音レベルが急に大きく変動する場合（例えば、音声通話中に乗用車が側を通り過ぎる場合には、雑音レベルが急に大きく変動する）には、全周波数帯域のＳ／Ｎ比が小さくなり、閾値ＴＨ以下となることがある。この場合、音声区間であっても、雑音区間であると誤判定される。

【００１０】

上記課題を解決するため、特開平２－２７２８３６号公報には、フレームパワー，零交差回数，最大零交差間隔及び有音／無音の判定結果に基づいて閾値を適応的に算出するものが開示されている。

即ち、零交差回数，最大零交差間隔及び有音／無音の判定出力の間にある条件を満足する場合に限り、過去フレームのパワーの平均値と現フレームのパワーの重み付け平均（ＡＲ平滑）を実施し、その重み付け平均を実施した値を定数倍した値を新たな閾値とすることにより、閾値の更新を行っている。

しかし、高雑音下では入力信号に占める雑音成分の割合が高いため、音声パワーとの判別が困難になり、また、零交差回数や最大零交差間隔の雑音区間での分布が音声区間での分布に類似した分布となるため、閾値の更新が適切に行われず、常に音声区間と判定されるため、雑音区間を音声区間と誤判定する課題があった。

【００１１】

また、特開平１０－１７１４９７号公報では、小帯域毎の帯域別Ｓ／Ｎ比と全周波数帯域のＳ／Ｎ比との差分値に応じて、スペクトル減算後の入力信号スペクトルに含まれる背景雑音の非定常成分を取り除いている。この方法は、帯域間の帯域別Ｓ／Ｎ比のばらつきを考慮に入れた方法であり、雑音の引き去り処理に対する効果は見られるが、音声区間の判定精度を高めることはできない。

【００１２】

この発明は上記のような課題を解決するためになされたもので、入力信号のＳ／Ｎ比が低い場合でも、的確に音声区間を判定することができる音声区間判定装置及び音声区間判定方法を得ることを目的とする。

【００１３】

【課題を解決するための手段】

この発明に係る音声区間判定装置は、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比の正変動量を算出し、判定処理を実行するに際して、その正変動量を考慮するようにしたものである。

【００１４】

この発明に係る音声区間判定装置は、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比

の負変動量を算出し、判定処理を実行するに際して、その負変動量を考慮するようにしたものである。

【００１５】

この発明に係る音声区間判定装置は、入力信号の現在の区間が雑音区間であると判定する場合、入力信号の現在の区間の帯域スペクトルにより過去の雑音区間の帯域スペクトルを更新するようにしたものである。

【００１６】

この発明に係る音声区間判定装置は、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比の平均値を算出し、判定処理を実行するに際して、その平均値を考慮するようにしたものである。

【００１７】

この発明に係る音声区間判定装置は、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比のうち上位Ｎ位（Ｎは自然数）の帯域別Ｓ／Ｎ比を検索し、判定処理を実行するに際して、その帯域別Ｓ／Ｎ比を考慮するようにしたものである。

【００１８】

この発明に係る音声区間判定装置は、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比の分散値を算出し、判定処理を実行するに際して、その分散値を考慮するようにしたものである。

【００１９】

この発明に係る音声区間判定装置は、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比が基準値以上である帯域の個数又は割合を求め、判定処理を実行するに際して、その個数又は割合を考慮するようにしたものである。

【００２０】

この発明に係る音声区間判定装置は、判定処理を実行するに際して、入力信号の現在の信号エネルギーを考慮するようにしたものである。

【００２１】

この発明に係る音声区間判定装置は、判定処理を実行するに際して、入力信号の差分信号エネルギーを考慮するようにしたものである。

【００２２】

この発明に係る音声区間判定装置は、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比にしたがって判定処理に用いる閾値を更新するようにしたものである。

【００２３】

この発明に係る音声区間判定方法は、帯域別Ｓ／Ｎ比の正変動量を算出し、判定処理を実行するに際して、その正変動量を考慮するようにしたものである。

【００２４】

この発明に係る音声区間判定方法は、帯域別Ｓ／Ｎ比の負変動量を算出し、判定処理を実行するに際して、その負変動量を考慮するようにしたものである。

【００２５】

【発明の実施の形態】

以下、この発明の実施の一形態を説明する。

実施の形態１．

図１はこの発明の実施の形態１による音声区間判定装置を示す構成図であり、図において、１は例えばマイクロフォンにより取り込まれた信号（入力信号）を入力する入力端子、２は入力端子１から入力された信号を所定の周波数でサンプリングを実施し、そのサンプリング結果であるフレーム信号（所定のサンプル数毎に区切られたフレーム信号）をアナログ／ディジタル変換するＡ／Ｄ変換器、３はＡ／Ｄ変換器２が出力するフレーム信号に対して窓関数演算を実行する窓関数演算器、４は窓関数演算器３の演算結果を一次のディジタルフィルタに通すことにより高域強調処理を実施する高域強調器である。

【００２６】

５は高域強調器４により高域強調されたフレーム信号を直交変換して、そのフレーム信号

の周波数スペクトルを求める直交変換器（直交変換手段）、６は直交変換器５により求められた周波数スペクトルを複数の帯域に分割して、そのフレーム信号の帯域スペクトルを求める帯域分割器（帯域分割手段）、７は総合判定器１３により制御される切換スイッチ、８は総合判定器１３が現在の入力信号が雑音区間であると判定する場合、入力信号の帯域スペクトルにより過去の雑音区間の帯域スペクトルを更新する雑音更新器、９は過去の雑音区間の帯域スペクトルを保持する雑音保持器である。

【００２７】

１０は帯域分割器６により求められたフレーム信号の帯域スペクトルと過去の雑音区間の帯域スペクトルとの比である帯域別Ｓ／Ｎ比を算出する帯域別Ｓ／Ｎ比算出器（Ｓ／Ｎ比算出手段）、１１は帯域別Ｓ／Ｎ比算出器１０により算出された帯域別Ｓ／Ｎ比の平均値を算出する平均値算出器、１２は帯域別Ｓ／Ｎ比算出器１０により算出された帯域別Ｓ／Ｎ比の正変動量を算出する正変動量算出器、１３は平均値算出器１１により算出された帯域別Ｓ／Ｎ比の平均値と正変動量算出器１２により算出された帯域別Ｓ／Ｎ比の正変動量とに基づいて現在の入力信号が音声区間又は雑音区間の何れであるかを判定する総合判定器、１４は総合判定器１３の判定結果を出力する出力端子である。

なお、雑音更新器８，雑音保持器９，平均値算出器１１，正変動量算出器１２及び総合判定器１３から判定手段が構成されている。

【００２８】

次に動作について説明する。

例えば、マイクロフォンにより取り込まれた信号が入力端子１から入力信号として入力されると、Ａ／Ｄ変換器２は、その入力信号に対して所定の周波数（例えば、８ｋＨｚの周波数）でサンプリングを実施し、そのサンプリング結果であるフレーム信号をアナログ／ディジタル変換する。

【００２９】

窓関数演算器３は、Ａ／Ｄ変換器２がフレーム信号をアナログ／ディジタル変換すると、変換後のフレーム信号に対して、例えば、ハニング窓などの窓関数を乗じることにより窓関数演算を実行する。

高域強調器４は、窓関数演算器３の演算結果を下記に示すような伝達関数Ｈ（ｚ）を有する一次のディジタルフィルタに通すことにより高域強調処理を実施する。

【００３０】

$H(z) = 1 - \alpha \cdot z^{-1}$

$z = \exp(j\omega)$

ただし、ωは角周波数、ｊは虚数単位、αは高域強調係数であり、通常１に近い値（例えば、０．８）が設定される。

なお、高域強調器４を通さず、窓関数演算器３の出力を直接直交変換器５に入力してもよい。

【００３１】

直交変換器５は、高域強調器４により高域強調されたフレーム信号又は窓関数演算器３により窓関数演算されたフレーム信号を入力すると、そのフレーム信号に対して、例えば、高速フーリエ変換（ＦＦＴ）などの直交変換処理を実行し、時間tの関数として表される時間軸信号を周波数成分（周波数スペクトル）に変換する。

【００３２】

帯域分割器６は、直交変換器５が周波数スペクトルを求めると、その周波数スペクトルを複数の帯域に分割することにより、そのフレーム信号の帯域スペクトルを求める。例えば、周波数スペクトルを１６の帯域に分割し、それぞれの帯域毎にスペクトルの重み付け平均をとり、これを帯域スペクトルＸ（ｉ，ｋ）とする。ただし、ｋは帯域番号、ｉはフレーム番号とする。

この他、直交変換された周波数スペクトルに対して、「バークフィルタ」又は「メルフィルタ」等の聴覚特性を模擬するフィルタを通した後、重み付け平均をとってもよい。

【００３３】

帯域別Ｓ／Ｎ比算出器１０は、帯域分割器６がフレーム信号の帯域スペクトルＸ（ｉ，ｋ）を求めると、雑音保持器９に保持されている前フレームまでに推定された雑音信号の帯域スペクトルＮ（ｉ－１，ｋ）との比、即ち、帯域別Ｓ／Ｎ比を帯域毎に算出する。

【００３４】

平均値算出器１１は、帯域別Ｓ／Ｎ比算出器１０が帯域別Ｓ／Ｎ比を出力すると、現フレームｉにおける帯域別Ｓ／Ｎ比の平均値ＳＮ＿ｍｅａｎ（ｉ）（以下、平均Ｓ／Ｎ比という）を算出する。

正変動量算出器１２は、帯域別Ｓ／Ｎ比算出器１０が帯域別Ｓ／Ｎ比を出力し、平均値算出器１１が平均Ｓ／Ｎ比ＳＮ＿ｍｅａｎ（ｉ）を出力すると、これらから下式にしたがって帯域別Ｓ／Ｎ比の正変動量ＳＮ＿ｐｌｕｓ＿ｖａｒｉａｔｉｏｎ（ｉ）を算出する。ここで、帯域別Ｓ／Ｎ比の正変動量とは、帯域別Ｓ／Ｎ比が平均Ｓ／Ｎ比より高い帯域において、帯域別Ｓ／Ｎ比と平均Ｓ／Ｎ比の差を加算したものである。ただし、下式において、Ｋは帯域数、ＳＮＲ（ｉ，ｋ）はフレーム番号ｉ，帯域番号ｋにおける帯域別Ｓ／Ｎ比である。

【００３５】

【数１】

$$SN_plus_variation(i) = \sum_{k=0}^{K-1} \delta(i,k)$$

ただし、

$$\delta(i,k) = \begin{cases} SNR(i,k) - SN_mean(i) & SNR(i,k) > SN_mean(i) \\ 0 & SNR(i,k) \le SN_mean(i) \end{cases}$$

【００３６】

この実施の形態１では、帯域別Ｓ／Ｎ比が平均Ｓ／Ｎ比より高い帯域において、帯域別Ｓ／Ｎ比と平均Ｓ／Ｎ比の差を加算して正変動量を算出するものについて示したが、帯域別Ｓ／Ｎ比の最大値と平均Ｓ／Ｎ比の差を正変動量としてもよい。また、帯域別Ｓ／Ｎ比の上位Ｎ位までの帯域において、帯域別Ｓ／Ｎ比と平均Ｓ／Ｎ比の差を加算して正変動量を算出してもよい。

【００３７】

総合判定器１３は、平均値算出器１１により算出された平均Ｓ／Ｎ比ＳＮ＿ｍｅａｎ（ｉ）と、正変動量算出器１２により算出された帯域別Ｓ／Ｎ比の正変動量ＳＮ＿ｐｌｕｓ＿ｖａｒｉａｔｉｏｎ（ｉ）と、前フレームにおける総合判定器１３の出力値（音声らしさレベルｓｐｅｅｃｈ＿ｌｅｖｅｌ（ｉ－１））とを用いて、図２の条件判定式１を実行することにより、現フレームｉにおける音声らしさレベルｓｐｅｅｃｈ＿ｌｅｖｅｌ（ｉ）を求める。ここで、音声らしさレベルは、フレーム信号が音声区間である可能性を示す尺度であり、音声らしさレベルが高いほど、音声区間である可能性が高い。

【００３８】

ただし、ＳＮ＿ｍｅａｎ＿ＴＨは平均Ｓ／Ｎ比に対する固定閾値、ＳＮ＿ｐｌｕｓ＿ｖａｒｉａｔｉｏｎ＿ＴＨ１，ＳＮ＿ｐｌｕｓ＿ｖａｒｉａｔｉｏｎ＿ＴＨ２は帯域別Ｓ／Ｎ比の正変動量に対する固定閾値である。また、ｓｐｅｅｃｈ＿ｔｍｐはフレームｉにおける音声らしさレベルを求めるための一時変数である。

【００３９】

なお、この実施の形態１では、閾値との大小比較により、音声らしさレベルが離散的に変化する場合について示しているが、平均Ｓ／Ｎ比及び正変動量を線形変換し、その変換値を加算した値を音声らしさレベルとすることにより、音声らしさレベルを連続的に変化させてもよい。ここで得られる音声らしさレベルの値は、とりうる値が複数であるため、音声らしさの程度を出力する場合であれば、音声らしさレベルの値をそのまま出力すればよいし、「音声区間」であるか「雑音区間」であるかを判定する場合には、音声らしさレベ

ルの値が所定の閾値以上であれば「音声区間」、そうでなければ「雑音区間」と判定すればよい。そして、その判定結果を出力端子１４から出力する。

【００４０】

総合判定器１３は、上記判定処理の実行後、次フレームの判定精度を高めるため、現フレームにおける入力信号が雑音区間であると判定した場合、または、音声らしさレベルが低い場合、切換スイッチ７をａ側に切り換えて、雑音保持器９に保持されている雑音信号の帯域スペクトルＮ（ｉ－１，ｋ）を更新させる。

【００４１】

雑音更新器８は、切換スイッチ７がａ側に切り換えられると、帯域分割器６が出力する現フレームにおける入力信号の帯域スペクトルＸ（ｉ，ｋ）と、雑音保持器９に保持されている前フレームまでに推定された雑音信号の帯域スペクトルＮ（ｉ－１，ｋ）とを下式に代入して、現フレームにおける雑音信号の帯域スペクトルＮ（ｉ，ｋ）を算出し、雑音保持器９に保持されている雑音信号を更新する。即ち、雑音保持器９の保持内容を現フレームにおける雑音信号の帯域スペクトルＮ（ｉ，ｋ）に変更する。

【００４２】

Ｎ（ｉ，ｋ）＝β・Ｎ（ｉ－１，ｋ）＋（１－β）・Ｘ（ｉ，ｋ）

ただし、βは雑音信号の推定速度を決定する係数であり、例えば、０．８という値をとるが、他にも帯域によって値を変えたり、音声らしさレベルが低い程、βを小さくすることにより、音声らしさレベルに応じた値をとるようにしてもよい。

【００４３】

以上で明らかなように、この実施の形態１によれば、帯域分割器６により求められたフレーム信号の帯域スペクトルと過去の雑音区間の帯域スペクトルとの比である帯域別Ｓ／Ｎ比を算出し、その帯域別Ｓ／Ｎ比に基づいて現在のフレーム信号が音声区間又は雑音区間の何れであるかを判定するように構成したので、現在のフレームのＳ／Ｎ比が低い等の場合でも、的確に音声区間を判定することができる効果を奏する。

【００４４】

即ち、帯域別Ｓ／Ｎ比の正変動量を算出する正変動量算出器１２を設けることにより、平均Ｓ／Ｎ比だけではなく、帯域別Ｓ／Ｎ比の帯域毎の大きさの違いを捉えることが可能となる。具体的には、平均Ｓ／Ｎ比により音声らしさレベルの初期値を決定し（条件判定式１のア，エ）、さらに帯域別Ｓ／Ｎ比の帯域毎の大きさの違い（帯域別Ｓ／Ｎ比の正変動量）により音声らしさレベルの初期値からの変化量を決定する（条件判定式１のイ，ウ，オ，カ）。

【００４５】

例えば、図１６に示す二つの帯域別Ｓ／Ｎ比の平均Ｓ／Ｎ比は、両者とも閾値より小さいため、音声らしさレベルの初期値が小さくなる。そして、図１６（ａ）の場合、正変動量が小さいので、音声らしさレベルをさらに小さくし、図１６（ｂ）の場合、正変動量が大きいので、音声らしさレベルを大きくする。これにより、帯域別Ｓ／Ｎ比の帯域毎の大きさの違いを考慮に入れた精度のよい音声区間判定を行うことができる。

【００４６】

実施の形態２．

上記実施の形態１では、帯域別Ｓ／Ｎ比の平均値を用いて音声区間であるか否かを判定するものについて示したが、帯域別Ｓ／Ｎ比のうち上位Ｎ位（Ｎは自然数）の帯域別Ｓ／Ｎ比を検索し、判定処理を実行するに際して、その帯域別Ｓ／Ｎ比を考慮するようにしてもよい。

なお、帯域別Ｓ／Ｎ比の最大値を判定処理に用いる場合には、帯域別Ｓ／Ｎ比の最大値を正変動量とすることにより、同様の音声区間判定を実施することができる。

【００４７】

実施の形態３．

上記実施の形態１，２では、帯域別Ｓ／Ｎ比の正変動量を算出し、その正変動量を用いて音声区間であるか否かを判定するものについて示したが、帯域別Ｓ／Ｎ比の負変動量を算

出し（帯域別Ｓ／Ｎ比の負変動量とは、帯域別Ｓ／Ｎ比が平均Ｓ／Ｎ比より低い帯域において、帯域別Ｓ／Ｎ比と平均Ｓ／Ｎ比の差を加算したものである）、その負変動量を用いて音声区間であるか否かを判定するようにしてもよい。
なお、正変動量と負変動量は極性のみの違いである。
【００４８】
実施の形態４.
図３はこの発明の実施の形態４による音声区間判定装置を示す構成図であり、図において、図１と同一符号は同一または相当部分を示すので説明を省略する。
１５は帯域別Ｓ／Ｎ比算出器１０により算出された帯域別Ｓ／Ｎ比の分散値を算出する分散算出器、１６は平均値算出器１１により算出された帯域別Ｓ／Ｎ比の平均値と分散算出器１５により算出された帯域別Ｓ／Ｎ比の分散値とに基づいて現在の入力信号が音声区間又は雑音区間の何れであるかを判定する総合判定器である。なお、分散算出器１５及び総合判定器１６は判定手段を構成する。
【００４９】
次に動作について説明する。
上記実施の形態１，２では、帯域別Ｓ／Ｎ比の正変動量を算出し、その正変動量を用いて音声区間であるか否かを判定するものについて示したが、帯域別Ｓ／Ｎ比の分散値を算出し、その分散値を用いて音声区間であるか否かを判定するようにしてもよい。
【００５０】
具体的には、分散算出器１５が帯域別Ｓ／Ｎ比の分散値ＳＮ＿ｖａｒｉａｎｃｅ（ｉ）を算出すると、総合判定器１６が平均値算出器１１により算出された平均Ｓ／Ｎ比ＳＮ＿ｍｅａｎ（ｉ）と、分散算出器１５により算出された帯域別Ｓ／Ｎ比の分散値ＳＮ＿ｖａｒｉａｎｃｅ（ｉ）と、前フレームにおける総合判定器１６の出力値（音声らしさレベルｓｐｅｅｃｈ＿ｌｅｖｅｌ（ｉ－１））とを用いて、図４の条件判定式２を実行することにより、現フレームｉにおける音声らしさレベルｓｐｅｅｃｈ＿ｌｅｖｅｌ（ｉ）を求める。
ここで、ＳＮ＿ｖａｒｉａｎｃｅ＿ＴＨ１，ＳＮ＿ｖａｒｉａｎｃｅ＿ＴＨ２は、帯域別Ｓ／Ｎ比の分散値に対する固定閾値である。
【００５１】
この実施の形態４によれば、帯域別Ｓ／Ｎ比の平均値からの広がりを表す分散値を用いることにより、帯域別Ｓ／Ｎ比の平均値からの相対的な変動量を捉えることが可能になり、上記実施の形態１と同様の効果を奏することができる。
【００５２】
実施の形態５.
図５はこの発明の実施の形態５による音声区間判定装置を示す構成図であり、図において、図１と同一符号は同一または相当部分を示すので説明を省略する。
１７は帯域別Ｓ／Ｎ比算出器１０により算出された帯域別Ｓ／Ｎ比が基準値以上である帯域の個数（高Ｓ／Ｎ比帯域数）を求める高Ｓ／Ｎ比帯域数算出器、１８は平均値算出器１１により算出された帯域別Ｓ／Ｎ比の平均値と高Ｓ／Ｎ比帯域数算出器１７により求められた高Ｓ／Ｎ比帯域数とに基づいて現在の入力信号が音声区間又は雑音区間の何れであるかを判定する総合判定器である。なお、高Ｓ／Ｎ比帯域数算出器１７及び総合判定器１８は判定手段を構成する。
【００５３】
次に動作について説明する。
上記実施の形態１，２では、帯域別Ｓ／Ｎ比の正変動量を算出し、その正変動量を用いて音声区間であるか否かを判定するものについて示したが、帯域別Ｓ／Ｎ比が基準値以上である帯域の個数（高Ｓ／Ｎ比帯域数）を求め、その高Ｓ／Ｎ比帯域数を用いて音声区間であるか否かを判定するようにしてもよい。
【００５４】
具体的には、高Ｓ／Ｎ比帯域数算出器１７は、帯域別Ｓ／Ｎ比算出器１０が帯域別Ｓ／Ｎ

比を出力すると、下式にしたがって高Ｓ／Ｎ比帯域数Ｈｉｇｈ＿ＳＮＲ＿ｎｕｍｂｅｒ（ｉ）を求める。ただし、下式において、ＳＮＲ＿ＴＨは帯域別Ｓ／Ｎ比との比較に用いる固定閾値、ε（ｉ，ｋ）はフレーム番号ｉ，帯域番号ｋにおける帯域別Ｓ／Ｎ比が閾値ＳＮＲ＿ＴＨより大きいか否かを示し、帯域別Ｓ／Ｎ比が閾値より大きければε（ｉ，ｋ）＝１、そうでなければ、ε（ｉ，ｋ）＝０である。

【００５５】

【数２】

$$\mathrm{High_SNR_number}(i) = \sum_{k=0}^{K-1} \varepsilon(i,k)$$

ただし、$$\varepsilon(i,k) = \begin{cases} 1 & \mathrm{SNR}(i,k) > \mathrm{SNR_TH} \\ 0 & \mathrm{SNR}(i,k) \le \mathrm{SNR_TH} \end{cases}$$

【００５６】

総合判定器１８が平均値算出器１１により算出された平均Ｓ／Ｎ比ＳＮ＿ｍｅａｎ（ｉ）と、高Ｓ／Ｎ比帯域数算出器１７により求められた高Ｓ／Ｎ比帯域数Ｈｉｇｈ＿ＳＮＲ＿ｎｕｍｂｅｒ（ｉ）と、前フレームにおける総合判定器１８の出力値（音声らしさレベルｓｐｅｅｃｈ＿ｌｅｖｅｌ（ｉ－１））とを用いて、図６の条件判定式３を実行することにより、現フレームｉにおける音声らしさレベルｓｐｅｅｃｈ＿ｌｅｖｅｌ（ｉ）を求める。

ここで、Ｈｉｇｈ＿ＳＮＲ＿ｎｕｍｂｅｒ＿ＴＨ１，Ｈｉｇｈ＿ＳＮＲ＿ｎｕｍｂｅｒ＿ＴＨ２は、高Ｓ／Ｎ比帯域数に対する固定閾値である。

【００５７】

この実施の形態５では、高Ｓ／Ｎ比帯域数を用いて音声区間であるか否かを判定するものについて示したが、帯域当たりの高Ｓ／Ｎ比帯域数（Ｈｉｇｈ＿ＳＮＲ＿ｎｕｍｂｅｒ（ｉ）／Ｋ）を算出し、帯域当たりの高Ｓ／Ｎ比帯域数を用いて音声区間であるか否かを判定するようにしてもよい。

この際、帯域当たりの高Ｓ／Ｎ比帯域数に対する閾値としては、Ｈｉｇｈ＿ＳＮＲ＿ｎｕｍｂｅｒ＿ＴＨ１／Ｋ，Ｈｉｇｈ＿ＳＮＲ＿ｎｕｍｂｅｒ＿ＴＨ２／Ｋを使用する。

【００５８】

また、高Ｓ／Ｎ比帯域数の代わりに、帯域Ｓ／Ｎ比が閾値未満である帯域の数（低Ｓ／Ｎ比帯域数）を算出し、帯域数Ｋから低Ｓ／Ｎ比帯域数を差し引いた値や、低Ｓ／Ｎ比帯域数を帯域数Ｋで割った値（１－帯域当たりの低Ｓ／Ｎ比帯域数）をＨｉｇｈ＿ＳＮＲ＿ｎｕｍｂｅｒ（ｉ）として、上記の条件判定式３を実行するなど、帯域別Ｓ／Ｎ比が基準値未満である帯域の数又は割合を用いて判定することも可能である。

【００５９】

この実施の形態５によれば、帯域別Ｓ／Ｎ比が基準値以上である帯域の個数又は割合を用いることにより、相対的に帯域別Ｓ／Ｎ比の高い帯域を捉えることが可能になり、上記実施の形態１と同様の効果を奏することができる。

【００６０】

実施の形態６．

図７はこの発明の実施の形態６による音声区間判定装置を示す構成図であり、図において、図１と同一符号は同一または相当部分を示すので説明を省略する。

１９は現フレームにおける入力信号の信号エネルギーを算出するエネルギー算出器、２０は現フレームにおける入力信号の差分信号エネルギーを算出する差分エネルギー算出器、２１は入力信号の信号エネルギーと差分信号エネルギーを考慮して現在の入力信号が音声区間又は雑音区間の何れであるかを判定する総合判定器である。なお、エネルギー算出器１９，差分エネルギー算出器２０及び総合判定器２１は判定手段を構成する。

【００６１】

次に動作について説明する。
上記実施の形態１では、平均Ｓ／Ｎ比と正変動量を考慮して、現在の入力信号が音声区間又は雑音区間の何れであるかを判定するものについて示したが、さらに、入力信号の信号エネルギーと差分信号エネルギーを考慮して判定するようにしてもよい。
なお、この例では、時間信号からエネルギーを算出するものについて示しているが、これに限るものではなく、周波数スペクトルから得られるスペクトルパワーを用いてもよいし、両者を用いてもよい。
【００６２】
具体的には、エネルギー算出器１９は、フレーム毎に入力信号の信号エネルギーＥｎｅｒｇｙ（ｉ）を下式にしたがって算出する。
ただし、下式において、Ｎはフレーム長をサンプル数で表現した場合の値であり、例えば、サンプリング周波数が８ｋＨｚ、フレーム長２０ｍｓｅｃの場合、Ｎ＝１６０となる。また、Ｘ（ｉ，ｎ）はフレーム（フレーム番号＝ｉ）内のｎ番目の入力信号である。
【００６３】
【数３】

$$\mathrm{Energy}(i) = 10 \cdot \log_{10}\left\{\frac{1}{N}\sum_{n=0}^{N-1} X(i,n)^2\right\}$$

【００６４】
差分エネルギー算出器２０は、前フレームの信号エネルギーＥｎｅｒｇｙ（ｉ－１）と現フレームの信号エネルギーＥｎｅｒｇｙ（ｉ）との差分、即ち、Ｅｎｅｒｇｙ（ｉ）－Ｅｎｅｒｇｙ（ｉ－１）を計算することにより、入力信号の差分信号エネルギーｄｅｌｔａ＿Ｅｎｅｒｇｙ（ｉ）を算出する。
【００６５】
総合判定器２１は、平均Ｓ／Ｎ比ＳＮ＿ｍｅａｎ（ｉ）と、帯域別Ｓ／Ｎ比の正変動量ＳＮ＿ｐｌｕｓ＿ｖａｒｉａｔｉｏｎ（ｉ）と、現フレームの信号エネルギーＥｎｅｒｇｙ（ｉ）と、差分信号エネルギーｄｅｌｔａ＿Ｅｎｅｒｇｙ（ｉ）と、前フレームにおける総合判定器２１の出力値（音声らしさレベルｓｐｅｅｃｈ＿ｌｅｖｅｌ（ｉ－１））とを用いて、図８の条件判定式４を実行することにより、現フレームｉにおける音声らしさレベルｓｐｅｅｃｈ＿ｌｅｖｅｌ（ｉ）を求める。
【００６６】
ただし、Ｅｎｅｒｇｙ＿ＴＨは現フレームの信号エネルギーＥｎｅｒｇｙ（ｉ）に対する固定閾値、ｄｅｌｔａ＿Ｅｎｅｒｇｙ＿ＴＨは差分信号エネルギーｄｅｌｔａ＿Ｅｎｅｒｇｙ（ｉ）に対する固定閾値である。
【００６７】
この実施の形態６によれば、背景雑音の推定精度に依存する帯域別Ｓ／Ｎ比の平均値及び正変動量に加えて、背景雑音の推定精度に依存しない入力信号の信号エネルギー及び差分信号エネルギーを用いることにより、背景雑音の推定精度が劣化しても、精度よく音声区間を判定することができる効果を奏する。
【００６８】
実施の形態７．
上記実施の形態６では、上記実施の形態１に対してエネルギー算出器１９と差分エネルギー算出器２０を追加することにより、入力信号の信号エネルギーと差分信号エネルギーを考慮して判定処理を実施するものについて示したが、上記実施の形態２，３に対してエネルギー算出器１９と差分エネルギー算出器２０を追加することにより、入力信号の信号エネルギーと差分信号エネルギーを考慮して判定処理を実施するようにしてもよく、上記実施の形態６と同様の効果を奏することができる。
【００６９】
実施の形態８．

上記実施の形態６では、上記実施の形態１に対してエネルギー算出器１９と差分エネルギー算出器２０を追加することにより、入力信号の信号エネルギーと差分信号エネルギーを考慮して判定処理を実施するものについて示したが、図９に示すように、上記実施の形態４に対してエネルギー算出器１９と差分エネルギー算出器２０を追加することにより、総合判定器（判定手段）２２が入力信号の信号エネルギーと差分信号エネルギーを考慮して判定処理を実施するようにしてもよく、上記実施の形態６と同様の効果を奏することができる。

この場合、総合判定器２２は、図１０の条件判定式５を実行することにより、現フレームｉにおける音声らしさレベルｓｐｅｅｃｈ＿ｌｅｖｅｌ（ｉ）を求める。

【００７０】

この実施の形態８によれば、背景雑音の推定精度に依存する帯域別Ｓ／Ｎ比の平均値及び分散値に加えて、背景雑音の推定精度に依存しない入力信号の信号エネルギー及び差分信号エネルギーを用いることにより、背景雑音の推定精度が劣化しても、精度よく音声区間を判定することができる効果を奏する。

【００７１】

実施の形態９．

上記実施の形態６では、上記実施の形態１に対してエネルギー算出器１９と差分エネルギー算出器２０を追加することにより、入力信号の信号エネルギーと差分信号エネルギーを考慮して判定処理を実施するものについて示したが、図１１に示すように、上記実施の形態５に対してエネルギー算出器１９と差分エネルギー算出器２０を追加することにより、総合判定器（判定手段）２３が入力信号の信号エネルギーと差分信号エネルギーを考慮して判定処理を実施するようにしてもよく、上記実施の形態６と同様の効果を奏することができる。

この場合、総合判定器２３は、図１２の条件判定式６を実行することにより、現フレームｉにおける音声らしさレベルｓｐｅｅｃｈ＿ｌｅｖｅｌ（ｉ）を求める。

【００７２】

この実施の形態９によれば、背景雑音の推定精度に依存する帯域別Ｓ／Ｎ比の平均値及び高Ｓ／Ｎ比帯域数に加えて、背景雑音の推定精度に依存しない入力信号の信号エネルギー及び差分信号エネルギーを用いることにより、背景雑音の推定精度が劣化しても、精度よく音声区間を判定することができる効果を奏する。

【００７３】

実施の形態１０．

上記実施の形態１～９では、総合判定器１３等が音声区間の判定処理に用いる判定パラメータ（例えば、帯域別Ｓ／Ｎ比の平均値、上位Ｎ位の帯域別Ｓ／Ｎ比、正変動量、負変動量、分散値、高Ｓ／Ｎ比帯域数、高Ｓ／Ｎ比帯域の割合）に対する閾値が固定のものを示したが、帯域別Ｓ／Ｎ比算出部１０により算出された帯域別Ｓ／Ｎ比にしたがって、これらの閾値を更新するようにしてもよい。

【００７４】

以下、閾値の判定処理を具体的に説明する。

図１３は総合判定器１３の内部を示す構成図であり、図において、１３ａは帯域別Ｓ／Ｎ比の平均値を入力して判定処理を実行する個別判定器、１３ｂは帯域別Ｓ／Ｎ比の正変動量を入力して判定処理を実行する個別判定器である。

図１４は個別判定器１３ａ，１３ｂの内部を示す構成図であり、図において、３１は比較演算器、３２は閾値保持器、３３は閾値更新器である。

【００７５】

次に動作について説明する。

比較演算器３１は、判定パラメータと閾値保持器３２に保持されている閾値との大小比較を実施し、判定パラメータの値が閾値よりも大きければ音声らしさレベルを上げる一方、小さければ音声らしさレベルを下げ、この結果、得られた音声らしさレベルを出力する。

【００７６】

閾値更新器３３は、現フレームの判定パラメータの値と過去の判定パラメータの値を用いて閾値を更新する。以下、判定パラメータの一例として、帯域別Ｓ／Ｎ比の平均値（平均Ｓ／Ｎ比）に対する閾値の更新処理を説明する。ただし、他の閾値も同様にして更新する。

図１５は平均Ｓ／Ｎ比から過去数フレーム（図では５フレーム）内における平均Ｓ／Ｎ比の最大値を求める様子を示している。

【００７７】

図１５の上段において、例えば、時刻t_1に着目すると、過去５フレーム（Ａの範囲）の平均Ｓ／Ｎ比の最大値は、時刻$t_{1,max}$のときの値であり、時刻$t_{1,max}$の値を図１５の下段にプロットする。時刻t_1以外の各時刻についても同様にして、図１５の下段に最大値をプロットする。

そして、閾値更新器３３は、図１５の下段に示される平均Ｓ／Ｎ比の最大値を下式に代入して、新たな閾値を計算する。

$$SN_mean_TH(i) = \gamma \cdot SN_mean_TH(i-1) + (1-\gamma) \cdot SN_mean_max(i)$$

【００７８】

ただし、ＳＮ＿ｍｅａｎ＿ＴＨ（ｉ）はフレームｉの平均Ｓ／Ｎ比に対する更新後の閾値、ＳＮ＿ｍｅａｎ＿ｍａｘ（ｉ）は過去５フレーム内の平均Ｓ／Ｎ比の最大値、γは平均Ｓ／Ｎ比に対する閾値の収束速度であり、例えば、０．７という値をとるが、他にも帯域や音声らしさレベルによって値を変えたりしてもよい。

【００７９】

これにより、この実施の形態１０によれば、上記実施の形態１～５の効果に加えて、例えば、雑音レベルが大きく上昇した場合、雑音レベルの上昇に追従して平均Ｓ／Ｎ比の最大値が小さくなり、これに連動して閾値も小さくなることから、固定閾値を用いる場合に比べて、追従性よく適切な閾値を設定することが可能となる。その結果、音声区間の判定精度が向上し、さらに雑音区間の帯域スペクトルの推定精度が向上する効果を奏する。

【００８０】

また、閾値の更新に過去数フレーム内の最大値を用いることにより、従来から一般的に用いられているＡＲ平滑を行う場合と比べて、閾値が過渡に変化せず、また、背景雑音レベルの変化量の大小に拘わらず、安定した音声区間の判定を実施することができる効果を奏する。なお、ＡＲ平滑を用いて、各種の閾値を更新してもよいことは言うまでもない。

【００８１】

【発明の効果】

以上のように、この発明によれば、帯域分割手段により求められた入力信号の現在の区間の帯域スペクトルと雑音区間と判定された過去の区間の帯域スペクトルとの比である帯域別Ｓ／Ｎ比を算出し、その帯域別Ｓ／Ｎ比に基づいて入力信号の現在の区間が音声区間又は雑音区間の何れであるかを判定するように構成したので、現在の入力信号のＳ／Ｎ比が低い等の場合でも、的確に音声区間を判定することができる効果がある。

この発明によれば、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比の正変動量を算出し、判定処理を実行するに際して、その正変動量を考慮するように構成したので、音声区間の判定精度を高めることができる効果がある。

【００８２】

この発明によれば、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比の負変動量を算出し、判定処理を実行するに際して、その負変動量を考慮するように構成したので、音声区間の判定精度を高めることができる効果がある。

【００８３】

この発明によれば、入力信号の現在の区間が雑音区間であると判定する場合、入力信号の帯域スペクトルにより過去の雑音区間の帯域スペクトルを更新するように構成したので、過去の雑音区間の帯域スペクトルの確度を高めることができる効果がある。

【００８４】

この発明によれば、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比の平均値を算出し、判定処理を実行するに際して、その平均値を考慮するように構成したので、音声区間の判定精度を高めることができる効果がある。

【００８５】

この発明によれば、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比のうち上位Ｎ位（Ｎは自然数）の帯域別Ｓ／Ｎ比を検索し、判定処理を実行するに際して、その帯域別Ｓ／Ｎ比を考慮するように構成したので、音声区間の判定精度を高めることができる効果がある。

【００８６】

この発明によれば、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比の分散値を算出し、判定処理を実行するに際して、その分散値を考慮するように構成したので、音声区間の判定精度を高めることができる効果がある。

【００８７】

この発明によれば、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比が基準値以上である帯域の個数又は割合を求め、判定処理を実行するに際して、その個数又は割合を考慮するように構成したので、音声区間の判定精度を高めることができる効果がある。

【００８８】

この発明によれば、判定処理を実行するに際して、入力信号の現在の信号エネルギーを考慮するように構成したので、背景雑音の推定精度が劣化しても、精度よく音声区間を判定することができる効果がある。

【００８９】

この発明によれば、判定処理を実行するに際して、入力信号の差分信号エネルギーを考慮するように構成したので、背景雑音の推定精度が劣化しても、精度よく音声区間を判定することができる効果がある。

【００９０】

この発明によれば、Ｓ／Ｎ比算出手段により算出された帯域別Ｓ／Ｎ比にしたがって判定処理に用いる閾値を更新するように構成したので、音声区間の判定精度がさらに向上する効果がある。

【図面の簡単な説明】

【図１】この発明の実施の形態１による音声区間判定装置を示す構成図である。

【図２】条件判定式１を示す説明図である。

【図３】この発明の実施の形態４による音声区間判定装置を示す構成図である。

【図４】条件判定式２を示す説明図である。

【図５】この発明の実施の形態５による音声区間判定装置を示す構成図である。

【図６】条件判定式３を示す説明図である。

【図７】この発明の実施の形態６による音声区間判定装置を示す構成図である。

【図８】条件判定式４を示す説明図である。

【図９】この発明の実施の形態８による音声区間判定装置を示す構成図である。

【図１０】条件判定式５を示す説明図である。

【図１１】この発明の実施の形態９による音声区間判定装置を示す構成図である。

【図１２】条件判定式６を示す説明図である。

【図１３】総合判定器１３の内部を示す構成図である。

【図１４】個別判定器１３ａ，１３ｂの内部を示す構成図である。

【図１５】平均Ｓ／Ｎ比から過去数フレーム内における平均Ｓ／Ｎ比の最大値を求める様子を示す説明図である。

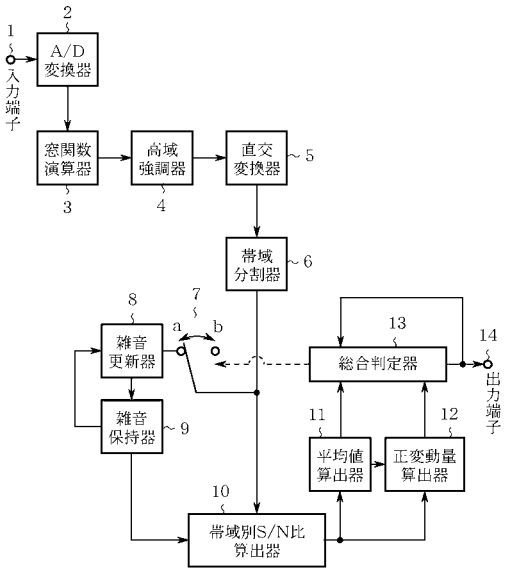
【図１６】帯域別Ｓ／Ｎ比の一例を表す説明図である。

【符号の説明】

１　入力端子、２　Ａ／Ｄ変換器、３　窓関数演算器、４　高域強調器、５　直交変換器（直交変換手段）、６　帯域分割器（帯域分割手段）、７　切換スイッチ、８　雑音更新器（判定手段）、９　雑音保持器（判定手段）、１０　帯域別Ｓ／Ｎ比算出器（Ｓ／Ｎ比

算出手段）、１１　平均値算出器（判定手段）、１２　正変動量算出器（判定手段）、１３　総合判定器（判定手段）、１３ａ　個別判定器、１３ｂ　個別判定器、１４　出力端子、１５　分散算出器（判定手段）、１６　総合判定器（判定手段）、１７　高Ｓ／Ｎ比帯域数算出器（判定手段）、１８　総合判定器（判定手段）、１９　エネルギー算出器（判定手段）、２０　差分エネルギー算出器（判定手段）、２１　総合判定器（判定手段）、２２　総合判定器（判定手段）、２３　総合判定器（判定手段）、３１　比較演算器、３２　閾値保持器、３３　閾値更新器。

【図１】

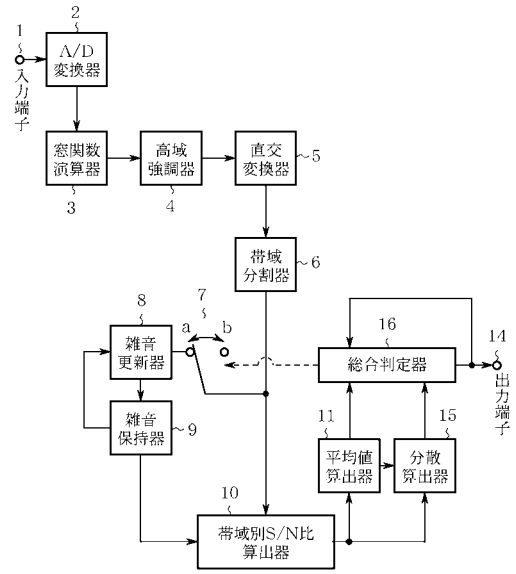


【図２】

```
[条件判定式1]

speech_tmp = speech_level ( i-1) ;
if ( SN_mean (i) >  SN_mean_TH1 ) {
   speech_tmp = speech_tmp + 2;        -------- ア
   if ( SN_plus_variation (i) > SN_plus_variation_TH1 )
      speech_tmp = speech_tmp + 1;     -------- イ
else
      speech_tmp = speech_tmp - 1;     -------- ウ
}else{
   speech_tmp = speech_tmp -2;         -------- エ
   if ( SN_plus_variation (i) > SN_plus_variation_TH2 )
      speech_tmp = speech_tmp + 1;     -------- オ
else
      speech_tmp = speech_tmp - 1;     -------- カ
}
speech_level (i) = speech_tmp;
```

【図３】

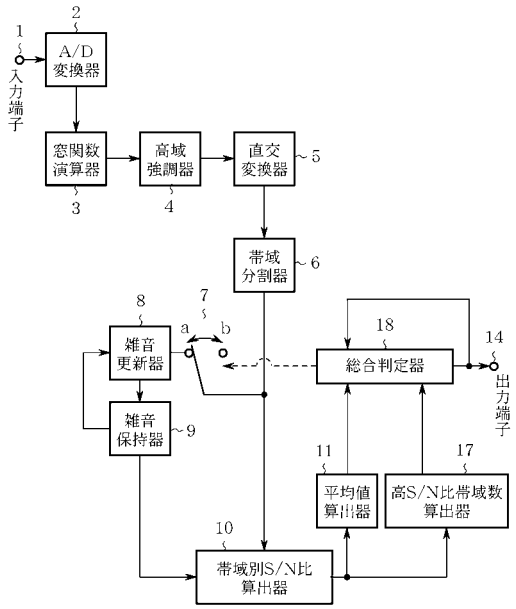


【図４】

```
[条件判定式2]

speech_tmp = speech_level ( i-1) ;
if ( SN_mean (i) >  SN_mean_TH ) {
   speech_tmp = speech_tmp + 1;
   if ( SN_variance (i) >  SN_variance_TH1 )
      speech_tmp = speech_tmp + 1;
else
      speech_tmp = speech_tmp - 1;
}else{
   speech_tmp = speech_tmp -1;
   if ( SN_variance (i) >  SN_variance_TH2 )
      speech_tmp = speech_tmp + 1;
else
      speech_tmp = speech_tmp - 1;
}
speech_level (i) = speech_tmp;
```

【図５】

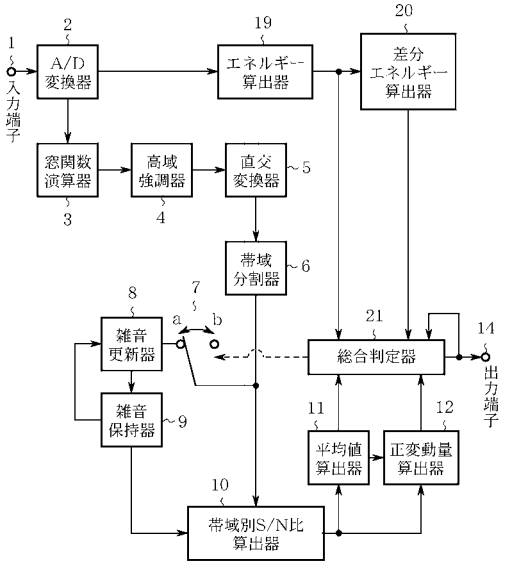


【図６】

```
[条件判定式3]

speech_tmp = speech_level ( i-1) ;
if ( SN_mean (i) >  SN_mean_TH ) {
   speech_tmp = speech_tmp + 1;
   if ( High_SNR_number (i) >  High_SNR_number_TH1 )
      speech_tmp = speech_tmp + 1;
else
      speech_tmp = speech_tmp - 1;
}else{
   speech_tmp = speech_tmp -1;
   if ( High_SNR_number (i) >  High_SNR_number_TH2 )
      speech_tmp = speech_tmp + 1;
else
      speech_tmp = speech_tmp - 1;
}
speech_level (i) = speech_tmp;
```

【図７】

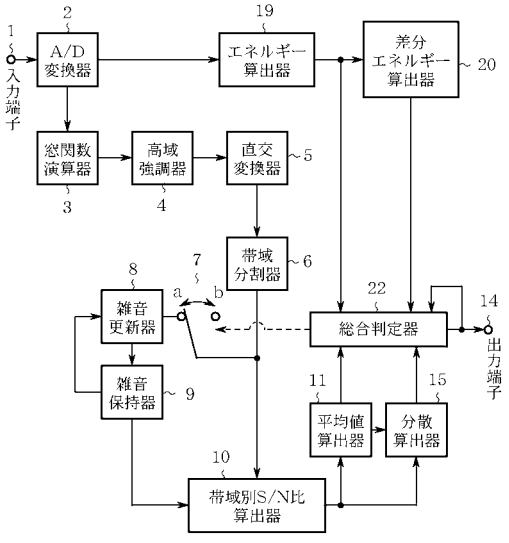


【図８】

[条件判定式4]

```
speech_tmp = speech_level ( i-1) ;
if ( SN_mean (i) >  SN_mean_TH ) {
  if ( SN_plus_variation (i) >  SN_plus_variation_TH1 )
     speech_tmp = speech_tmp + 2;
else
     speech_tmp = speech_tmp + 1;
}else{
  if ( SN_plus_variation (i) >  SN_plus_variation_TH2 )
     speech_tmp = speech_tmp - 1;
else
     speech_tmp = speech_tmp - 2;
}
if ( delta_Energy (i) > delta_Energy_TH )
     speech_tmp = speech_tmp + 1;
if ( Energy (i) < Energy_TH )
     speech_tmp = speech_tmp - 2;
speech_level (i) = speech_tmp;
```

【図９】

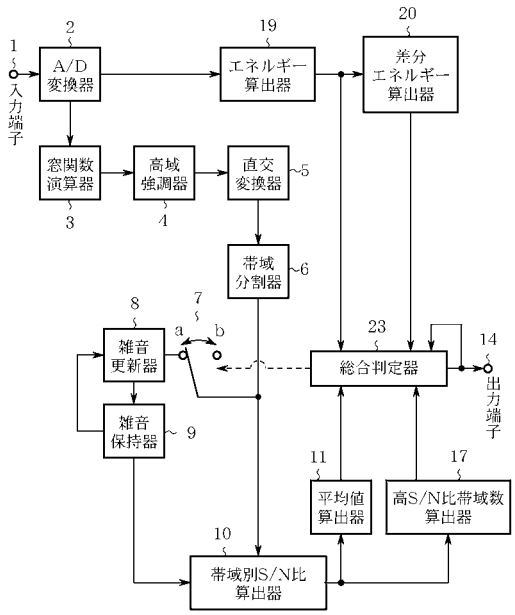


【図１０】

[条件判定式5]

```
speech_tmp = speech_level ( i-1) ;
if ( SN_mean (i) >  SN_mean_TH ) {
     speech_tmp = speech_tmp + 1;
     if ( SN_variance (i) >  SN_variance_TH1 )
     speech_tmp = speech_tmp + 1;
else
     speech_tmp = speech_tmp - 1;
}else{
  speech_tmp = speech_tmp - 1;
  if ( SN_plus_variance (i) >  SN_plus_variance_TH2 )
     speech_tmp = speech_tmp + 1;
else
     speech_tmp = speech_tmp - 1;
}
if ( delta_Energy (i) > delta_Energy_TH )
     speech_tmp = speech_tmp + 2;
if ( Energy (i) < Energy_TH )
     speech_tmp = speech_tmp - 1;
speech_level (i) = speech_tmp;
```

【図１１】

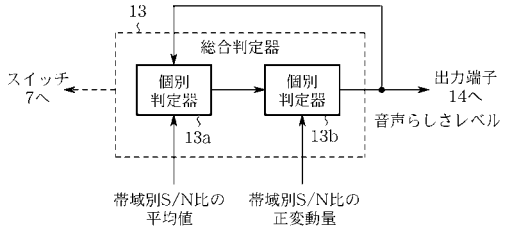


【図１２】

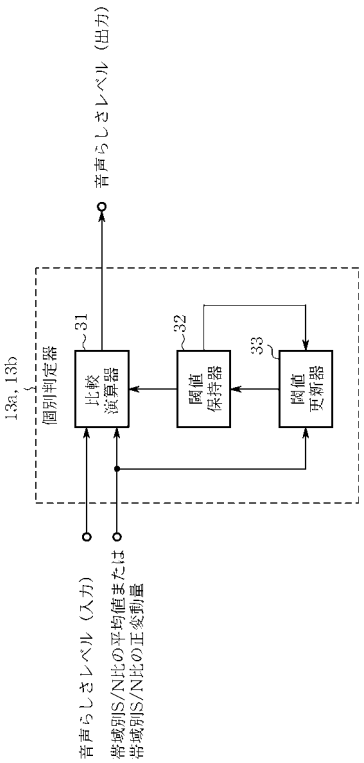
[条件判定式6]

```
speech_tmp = speech_level ( i-1) ;
if ( SN_mean (i) >  SN_mean_TH ) {
      speech_tmp = speech_tmp + 1;
      if ( High_SNR_number (i) >  High_SNR_number_TH1 )
     speech_tmp = speech_tmp + 1;
else
     speech_tmp = speech_tmp - 1;
}else{
   speech_tmp = speech_tmp - 1;
   if ( High_SNR_number (i) >  High_SNR_number_TH2 )
     speech_tmp = speech_tmp + 1;
else
     speech_tmp = speech_tmp - 1;
}
if ( delta_Energy (i) > delta_Energy_TH )
     speech_tmp = speech_tmp + 2;
if ( Energy (i) < Energy_TH )
     speech_tmp = speech_tmp - 1;
speech_level (i) = speech_tmp;
```

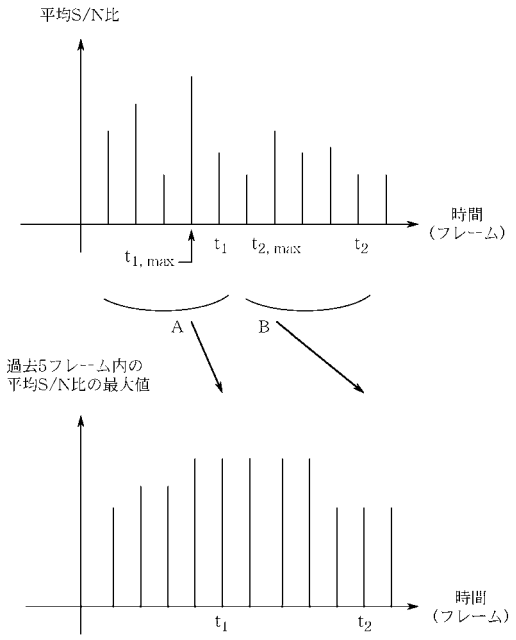
【図１３】



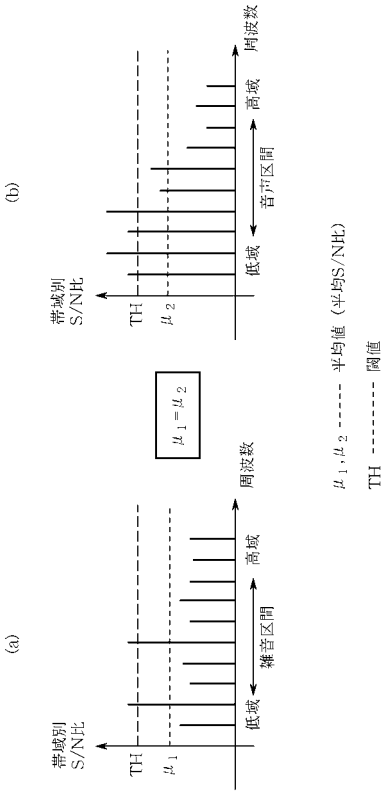
【図１４】



【図１５】


平均S/N比
時間
(フレーム)
$t_{1,max}$
t_1
$t_{2,max}$
t_2
A
B
過去5フレーム内の
平均S/N比の最大値
時間
(フレーム)
t_1
t_2

【図１６】


(b)
周波数
高域
音声区間
低域
帯域別
S/N比
TH
μ_2
$\mu_1 = \mu_2$
(a)
周波数
高域
雑音区間
低域
帯域別
S/N比
TH
μ_1
μ_1, μ_2 ---- 平均値（平均S/NHL）
TH ------ 閾値

フロントページの続き

(56)参考文献 特公平０２－０３９７９９（ＪＰ，Ｂ２）
特開平０９－２１２１９５（ＪＰ，Ａ）
特開平０９－１６０５８５（ＪＰ，Ａ）
特開平０３－１１４１００（ＪＰ，Ａ）
特開平０３－２６６８９９（ＪＰ，Ａ）
特開平０３－２５３８９９（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.[7]，ＤＢ名)
G10L 11/00
G10L 11/02