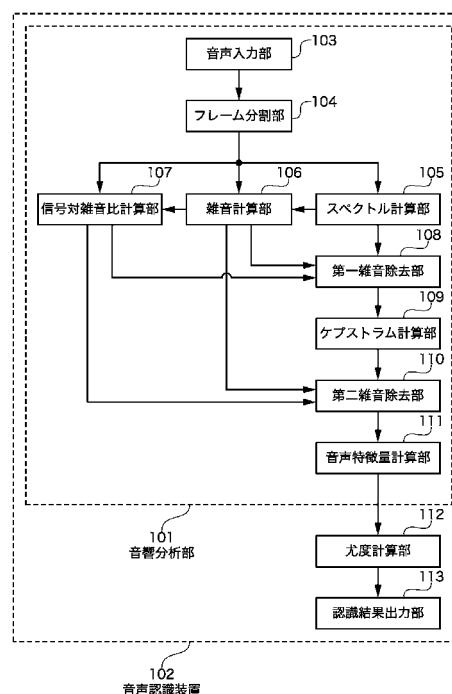




【特許請求の範囲】**【請求項 1】**

入力した音声の雑音成分を抑圧する雑音抑圧装置であって、
前記音声に含まれる雑音を推定する手段と、
前記音声と前記雑音とから信号対雑音比を求める手段と、
スペクトル次元において、前記音声から、前記信号対雑音比に応じて設定される第 1 の減算係数と前記雑音とを乗じて得た値を減算するスペクトル減算を行う手段と、
ケプストラム次元において、前記スペクトル減算の結果から、前記信号対雑音比に応じて設定される第 2 の減算係数と前記雑音とを乗じて得た値を減算するケプストラム減算を行う手段と、
を有することを特徴とする雑音抑圧装置。

10

【請求項 2】

前記信号対雑音比が低いほど、第 1 の減算係数が小さな値に設定される一方、第 2 の減算係数が大きな値に設定されることを特徴とする請求項 1 に記載の雑音抑圧装置。

【請求項 3】

第 1 の減算係数と第 2 の減算係数とは、両者の和が一定値となる関係を有することを特徴とする請求項 2 に記載の雑音抑圧装置。

【請求項 4】

第 1 の減算係数は、周波数帯域別に設定されることを特徴とする請求項 1 に記載の雑音抑圧装置。

20

【請求項 5】

請求項 1 から 4 までのいずれかに記載の雑音抑圧装置を有する音声認識装置。

【請求項 6】

入力した音声の雑音成分を抑圧する雑音抑圧方法であって、
前記音声に含まれる雑音を推定するステップと、
前記音声と前記雑音とから信号対雑音比を求めるステップと、
スペクトル次元において、前記音声から、前記信号対雑音比に応じて設定される第 1 の減算係数と前記雑音とを乗じて得た値を減算するスペクトル減算を行うステップと、
ケプストラム次元において、前記スペクトル減算の結果から、前記信号対雑音比に応じて設定される第 2 の減算係数と前記雑音とを乗じて得た値を減算するケプストラム減算を行うステップと、
を有することを特徴とする雑音抑圧方法。

30

【請求項 7】

前記信号対雑音比が低いほど、第 1 の減算係数が小さな値に設定される一方、第 2 の減算係数が大きな値に設定されることを特徴とする請求項 6 に記載の雑音抑圧方法。

【請求項 8】

第 1 の減算係数と第 2 の減算係数とは、両者の和が一定値となる関係を有することを特徴とする請求項 7 に記載の雑音抑圧方法。

【請求項 9】

第 1 の減算係数は、周波数帯域別に設定されることを特徴とする請求項 6 に記載の雑音抑圧方法。

40

【請求項 10】

コンピュータ実行可能なプログラムであって、入力した音声の雑音成分を抑圧するために、
前記音声に含まれる雑音を推定するためのコードと、
前記音声と前記雑音とから信号対雑音比を求めるためのコードと、
スペクトル次元において、前記音声から、前記信号対雑音比に応じて設定される第 1 の減算係数と前記雑音とを乗じて得た値を減算するスペクトル減算を実行させるためのコードと、
ケプストラム次元において、前記スペクトル減算の結果から、前記信号対雑音比に応じ

50

て設定される第2の減算係数と前記雑音とを乗じて得た値を減算するケプストラム減算を実行させるためのコードと、

を含むことを特徴とするプログラム。

【請求項11】

請求項10に記載のプログラムを格納したコンピュータ読み取り可能な記憶媒体。

【請求項12】

入力した音声の雑音成分を抑圧する雑音抑圧装置であって、

前記音声に含まれる雑音を推定する手段と、

前記音声と前記雑音とから信号対雑音比を求める手段と、

スペクトル次元において、前記信号対雑音比に応じた数だけ隣接する周波数成分毎の前記雑音のスペクトルの平均を計算して平均化雑音スペクトルを得る手段と、

スペクトル次元において、前記音声から前記平均化雑音スペクトルを減算する手段と、

を有することを特徴とする雑音抑圧装置。

【請求項13】

請求項12に記載の雑音抑圧装置を有する音声認識装置。

【請求項14】

入力した音声の雑音成分を抑圧する雑音抑圧方法であって、

前記音声に含まれる雑音を推定するステップと、

前記音声と前記雑音とから信号対雑音比を求めるステップと、

スペクトル次元において、前記信号対雑音比に応じた数だけ隣接する周波数成分毎の前記雑音のスペクトルの平均を計算して平均化雑音スペクトルを得るステップと、

スペクトル次元において、前記音声から前記平均化雑音スペクトルを減算するステップと、

を有することを特徴とする雑音抑圧方法。

【請求項15】

コンピュータ実行可能なプログラムであって、入力した音声の雑音成分を抑圧するために、

前記音声に含まれる雑音を推定するためのコードと、

前記音声と前記雑音とから信号対雑音比を求めるためのコードと、

スペクトル次元において、前記信号対雑音比に応じた数だけ隣接する周波数成分毎の前記雑音のスペクトルの平均を計算して平均化雑音スペクトルを得るためのコードと、

スペクトル次元において、前記音声から前記平均化雑音スペクトルを減算するためのコードと、

を含むことを特徴とするプログラム。

【請求項16】

請求項15に記載のプログラムを格納したコンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声処理技術に関し、特に、入力した音声データから雑音成分を除去する技術に関する。

【背景技術】

【0002】

音声認識は、例えば、誰でも簡単に音声でコマンドを入力することができるユーザーインターフェースとして注目されている。近年では、隠れマルコフモデル(HMM)を音響モデルに用いた音声認識が主流となり、雑音の少ない環境下においては高い認識率で不特定話者の音声を認識できるようになっている。しかし、実環境において音声認識を行う場合は、雑音が大きな問題となる。

【0003】

雑音に対処する方法としては、雑音を含んだ原音声から雑音成分を推定し、その雑音成

10

20

30

40

50

分を減じることにより雑音を含んでいない音声を推定して認識する方法が一般的である。その代表的なものとしてスペクトルサブトラクション (Spectral Subtraction) 法 (以下、「SS法」という。) が提案されている (例えば、非特許文献 1 を参照。)。SS法は、非音声区間などで求めた雑音成分を原音声からスペクトル領域において減じることによって雑音を除去する方法である。SS法に代表される雑音適応化法では、雑音を推定して、これを原音声の特徴量から減ずることにより、雑音のない音声の特徴量に近づけて認識率の向上を狙っている。

【0004】

図 8 は、SS法を説明する図である。同図において、801 は原音声、802 は推定した雑音成分、803 は SS法により雑音成分を除去した音声である。また、804 が SS法を表す式である。通常はスペクトル領域において、804 式の係数 α を固定にして、原音声スペクトル 801 (X_{S+N}) から推定した雑音スペクトル 802 (X_N) を減ずる。

10

【0005】

【非特許文献 1】Steven F. BOLL: "Suppression of Acoustic Noise Speech Using Spectral Subtraction", IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-27, No.2, (1979), p113-120

【発明の開示】

【発明が解決しようとする課題】

【0006】

しかしながら、推定した雑音成分が正しく求まっていない場合には、雑音成分の引きすぎなどが起こる。特に、音声信号と雑音のレベルに大きな差がない場合では、減算の結果がマイナスになり、これを補正するために所定の値で埋めるフロアリング処理を施すことになる。そうすると、フロアリング処理を行う場所と行わない場所で値が大きく変動し、雑音除去性能に悪影響を与える。また、各周波数サンプルで減算を行うため、隣り合う周波数サンプルでフロアリングされたりされなかったりと、減算後の値がばらつき、不連続になって雑音除去性能に悪影響を与える。そのため、なるべくフロアリング処理をなくすことと、なるべく減算による値のばらつきを抑えることが、SS法を効果的に働かせるために有効である。

20

【0007】

本発明はこのような技術的課題に鑑みてなされたものであり、スペクトルサブトラクションのフロアリング処理によるひずみや値のばらつき、不連続性を低減させ、もって雑音抑圧性能を向上させることを目的とする。

30

【課題を解決するための手段】

【0008】

上記した課題は本発明の雑音抑圧装置および方法によって解決される。本発明の一側面に係る雑音抑圧装置および方法においては、次のような処理が行われる。まず、入力した音声に含まれる雑音を推定し、前記音声と前記雑音とから信号対雑音比を求める。次に、スペクトル次元において、前記音声から、前記信号対雑音比に応じて設定される第 1 の減算係数と前記雑音とを乗じて得た値を減算するスペクトル減算を行い、さらに、ケプストラム次元において、前記スペクトル減算の結果から、前記信号対雑音比に応じて設定される第 2 の減算係数と前記雑音とを乗じて得た値を減算するケプストラム減算を行う。

40

【0009】

また、本発明の別の側面に係る雑音抑圧装置および方法によれば、上記した課題は次のようにして解決される。まず、入力した音声に含まれる雑音を推定し、前記音声と前記雑音とから信号対雑音比を求める。次に、スペクトル次元において、前記信号対雑音比に応じた数だけ隣接する周波数成分毎の前記雑音のスペクトルの平均を計算して平均化雑音スペクトルを得た後、スペクトル次元において、前記音声から前記平均化雑音スペクトルを減算する。

【発明の効果】

【0010】

50

本発明によれば、スペクトルサブトラクションのフロアリング処理によるひずみや値のばらつき、不連続性を低減させ、もって雑音抑圧性能を向上させることができる。

【発明を実施するための最良の形態】

【0011】

以下、図面を参照して本発明の好適な実施形態について詳細に説明する。

【0012】

(第1の実施形態)

図1は、本発明の雑音抑圧装置が適用される音声認識装置の機能構成を示すブロック図である。

【0013】

同図において、101は、音声の特徴を分析する音響分析部である。102は、音響分析部101を含み、音響分析部101の出力である音声特徴量を用いて音声を認識する音声認識装置である。103は、音声認識装置102において、音声を入力する音声入力部である。104は、入力音声から所定の区間毎にフレームとして区切るフレーム分割部である。105は、音声の各フレームでスペクトルを求めるスペクトル計算部である。106は、音声から雑音成分を推定する雑音計算部である。

【0014】

107は、各フレームにおける信号と雑音とのレベルの比を求める信号対雑音比計算部である。108は、各フレームにおいて、音声から雑音成分をスペクトル次元で除去する第一雑音除去部である。109はスペクトル次元で雑音除去した音声から、LPCケプストラムやMFCCなどのケプストラムを求めるケプストラム計算部である。110は、雑音成分をケプストラム次元で除去する第二雑音除去部である。111は、雑音除去したケプストラムや、音声パワーや、これらの回帰係数などの情報により音声特徴量を形成する音声特徴量計算部である。112は、音声特徴量を入力として、各認識語彙の尤度計算を行う尤度計算部である。113は、尤度計算の結果から得られる音声認識結果を出力する認識結果出力部である。

【0015】

図2は、本実施形態における認識処理の流れを表すフローチャートである。図1、図2を用いて、本実施形態の認識処理の流れを以下に説明する。

【0016】

まず、ステップS201で音声入力部103が音声を入力する。この入力音声を、ステップS202において、フレーム分割部104が所定の時間区間のフレームに分割する。ここでは、1フレームの時間長を例えば25msecとするが、各フレームは例えば10msecの時間だけシフトしながら得られるものとする。次に、ステップS203において、スペクトル計算部105は音声フレームからスペクトル(パワースペクトルもしくは振幅スペクトルのいずれでも良い)を計算する。次に、ステップS204において、雑音計算部106は音声に含まれる雑音成分を計算する。ここで、雑音成分は各フレームで更新しても構わないし、非音声区間と推定できる所定のフレームで計算した後は推定を止め、同じ雑音成分の値を保持しつづけても良い。

【0017】

次に、ステップS205において、信号対雑音比計算部107は現フレームの音声と雑音計算部106が推定した雑音とのレベル(パワーなど)の比を計算し、信号対雑音比(SNR)を求める。次に、ステップS206において、第一雑音除去部108は、信号対雑音比計算部107が決定した信号対雑音比によって第一減算係数を決定し、雑音計算部106が推定した雑音のスペクトルにこの第一減算係数を乗じ、音声スペクトルから引く。

【0018】

次に、ステップS207において、ケプストラム計算部109が第一雑音除去部108の出力からケプストラムを計算する。ケプストラムの計算については、スペクトルにメル周波数窓をかけてから求めるMFCCでもよいし、線形予測分析をしてから求めるLPC

10

20

30

40

50

ケプストラムでもよく、あるいはその他の方法でも構わない。次に、ステップ S 2 0 8 において、第二雑音除去部 1 1 0 は、信号対雑音比計算部 1 0 7 が決定した信号対雑音比によって第二減算係数を決定し、雑音計算部 1 0 6 が推定した雑音のケプストラムにこの第二雑音係数を乗じ、音声のケプストラムから引く。

【 0 0 1 9 】

次に、ステップ S 2 0 9 において音声特徴量計算部 1 1 1 が第二雑音除去部 1 1 0 の出力のケプストラムや、このケプストラムの回帰係数、音声パワーなどの情報を含む音声特徴量を形成する。次に、ステップ S 2 1 0 において、尤度計算部 1 1 2 がこの特徴量を入力し、尤度計算を行う。ステップ S 2 1 1 において、入力音声の最終フレームまで解析が済むか、尤度計算により解が決定した場合には、ステップ S 2 1 2 において認識結果出力部 1 1 3 が音声認識結果を出力する。そうでない場合にはステップ S 2 0 2 へ戻り、次のフレームに進む。

10

【 0 0 2 0 】

ここで、第一雑音除去部 1 0 8 及び第二雑音除去部 1 1 0 における雑音除去について詳細に説明する。第一雑音除去部 1 0 8 はスペクトル次元での減算、第二雑音除去部 1 1 0 はケプストラム次元での減算である。ここで、第一雑音除去部 1 0 8 の入力となる音声スペクトルを X_{S+N} 、推定雑音のスペクトルを X_N 、雑音を減算したスペクトルを X_S と表すと、第一雑音除去部の処理は次式で表される。

【 0 0 2 1 】

$$X_S = X_{S+N} - (1 - k(SNR)) X_N \quad (\text{if } X_{S+N} - (1 - k(SNR)) X_N \geq \alpha X_{S+N}) \quad 20$$

$$X_S = \alpha X_{S+N} \quad (\text{else})$$

【 0 0 2 2 】

ここで、上式の第 2 式がスペクトルが負にならないように補正をするフロアリング処理である。なお、 α は従来の SS 法で適用される減算係数である（図 8 を参照。）。

【 0 0 2 3 】

また、第二雑音除去部 1 1 0 の入力となる音声ケプストラムを C_{S+N} 、推定雑音のケプストラムを C_N 、雑音を減算したケプストラムを C_S と表すと、第二雑音除去部の処理は次式のように計算される。ケプストラムについては減算の値が負になっても良いのでフロアリング処理は必要ない。

30

【 0 0 2 4 】

$$C_S = C_{S+N} - k(SNR) C_N$$

【 0 0 2 5 】

これらの式で、上式の第二減算係数 $k(SNR)$ は、信号対雑音比計算部 1 0 7 で計算される信号対雑音比に応じて設定される。図 3 に、信号対雑音比 SNR に対する第二減算係数 $k(SNR)$ をプロットしたグラフを示す。同図のように、第二減算係数 $k(SNR)$ は信号対雑音比が高くなると例えば 0 に近づき、信号対雑音比が低くなるほど例えば 1 に近づきよう設計する。

【 0 0 2 6 】

これにより、非音声区間など信号対雑音比が低いフレームにおいては、第二減算係数 $k(SNR)$ は大きな値に設定される一方、第一減算係数は $1 - k(SNR)$ で表されるから、その値は逆に小さくなる。典型的には、第一減算係数と第二減算係数との和が一定値（例えば、1）になる関係を有することが好ましい。したがって信号対雑音比が低い場合には、スペクトル次元では雑音は大きく減算されず、これによりスペクトルの過減算によるフロアリング処理の発生を防止する。その一方でこの場合には、第二減算係数が大きな値に設定され、ケプストラム次元で雑音が大きく減算されることになる。

40

【 0 0 2 7 】

この効果として、スペクトル次元の減算でフロアリングが起き難くなる。また、スペクトルからケプストラムを求めるにはスペクトルの周波数成分をメル窓や線形予測係数などにより代表点にならしてから求めるので、ケプストラム次元での減算ではスペクトルの各

50

周波数成分で引く時のような値の不連続なばらつきを抑えられる。

【0028】

その一方で、音声区間など信号対雑音比が高いフレームにおいては第一雑音除去部の第一減算係数が大きくなり、スペクトル次元で雑音は大きく減算され、第二雑音除去部の第二減算係数が小さくなり、ケプストラム次元で雑音は大きく減算されなくなる。これにより、信号が雑音に対して大きい場合には、SS法と同様にスペクトル次元で積極的に雑音を減算する。

【0029】

このように本実施形態によれば、音響分析部において、スペクトル次元とケプストラム次元の両方で雑音の減算を行い、信号対雑音比に応じてそれぞれの雑音減算係数が調整される。 10

【0030】

なお、上述の実施形態では第一雑音除去部108で使用する第一減算係数 $1 - k(SNR)$ を全ての周波数帯域に対して同じ値に設定したが、異なる複数の周波数帯域(サブバンド)毎に信号対雑音比を測定し、サブバンド毎に異なる第一減算係数を設定しても良い。この場合、ケプストラム次元における第二雑音除去部110の第二減算係数は、サブバンド毎の第一減算係数にそれぞれ対応する $k(SNR)$ を平均して得た値を用いてもよい。

【0031】

(第2の実施形態)

上述の第1の実施形態では、スペクトル次元とケプストラム次元の両方で雑音の減算を行い、信号対雑音比によってそれぞれの雑音減算係数を変えるようにしたが、本実施形態は、スペクトル次元のみで雑音の減算を行い、そのときに信号対雑音比に応じて雑音の周波数方向の平滑化を行う別の実施形態である。 20

【0032】

図4は、本実施形態に係る音声認識装置の機能構成を示すブロック図を示す。

【0033】

同図において、401は、音声の特徴を分析する音響分析部である。402は、音響分析部401を有し、音響分析部401の出力である音声特徴量を用いて音声を認識する音声認識装置である。403は、音声認識装置402において、音声を入力する音声入力部 30
である。404は、入力音声から所定の区間毎にフレームとして区切るフレーム分割部である。405は、音声の各フレームでスペクトルを求めるスペクトル計算部である。406は、音声から雑音成分を推定する雑音計算部である。

【0034】

407は、各フレームにおける信号と雑音とのレベルの比を求める信号対雑音比計算部である。408は、雑音計算部407で計算された雑音成分のスペクトルを、各周波数成分において隣接する周波数におけるスペクトルの値との平均を取る雑音平均部である。409は、各フレームにおいて、雑音平均部408で求めた平滑化された雑音を入力音声フレームからスペクトル次元で除去する雑音除去部である。410は、ケプストラムや音声のパワーやこれらの回帰係数などの音声特徴量を計算する音声特徴量計算部である。4 40
11は、音声特徴量を入力として、各認識語彙の尤度計算を行う尤度計算部である。412は、尤度計算の結果から得られる音声認識結果を出力する認識結果出力部である。

【0035】

図5は、本実施形態における認識処理の流れを表すフローチャートである。図4、図5を用いて、本実施形態の認識処理の流れを以下に説明する。

【0036】

まず、ステップS501で音声入力部403が音声を入力する。この入力音声を、ステップS502において、フレーム分割部404が所定の時間区間のフレームに分割する。ここでは、1フレームの時間長を例えば25msecとするが、各フレームは例えば10msecの時間だけシフトしながら得られるものとする。次に、ステップS503において、スペクト 50

ル計算部405は音声フレームからスペクトルを計算する。次に、ステップS504において、雑音計算部406は音声に含まれる雑音成分を計算する。ここで、雑音成分は各フレームで更新しても構わないし、非音声区間と推定できる所定のフレームで計算した後は推定を止め、同じ雑音成分の値を保持しつづけても良い。

【0037】

次に、ステップS505において、信号対雑音比計算部407は現フレームの音声と雑音計算部406が推定した雑音とのレベル（パワーなど）の比を計算し、信号対雑音比（SNR）を求める。次に、ステップS506において、雑音平均部408は、雑音成分のスペクトルを各周波数成分において隣接する周波数におけるスペクトルの値との平均を取る。ここで、平均を取るサンプル数は、信号対雑音比計算部407によって求めた信号対雑音比に応じて決定する。図6に、信号対雑音比に応じて隣接サンプル数を変化させる対応グラフを示す。同図において横軸が信号対雑音比、縦軸が平均を取るための隣接サンプル数Lであり、各周波数のサンプル及び前後Lサンプル、計 $(2L+1)$ サンプルで平均を取ることを意味する。信号対雑音比が高い場合には雑音に対して信号が大きいため、減算によるフロアリングや値のばらつき、不連続性が起こりにくいので、平均を取らずそのまま雑音を用いる（隣接サンプル数を0に近づける）。一方、信号対雑音比が低い場合には、信号に対する雑音の影響が大きいため減算によるフロアリングや値のばらつき、不連続性が起こりやすいので、大きめの隣接サンプル数で雑音を平均化する。このとき、隣接サンプル数は全ての周波数帯で同じ値を使っても良いし、異なる複数の周波数帯域（サブバンド）毎に信号対雑音比を求めて異なる隣接サンプル数を用いても良い。

10

20

【0038】

図7は、各サンプルで周波数方向に平均を取った様子を表す図である。同図のように、それぞれの周波数成分で平均をとることにより、スペクトル形状が滑らかになり、減算の際の不連続性が起こりにくくなる。

【0039】

次に、ステップS507において、雑音除去部409は、雑音平均部408が求めた平滑化された雑音スペクトルを音声スペクトルから引く。本実施形態において雑音を減算する際の減算係数は固定値を用い、減算によって負になったサンプルではフロアリング処理を施す。

【0040】

次に、ステップS508において、音声特徴量計算部410が雑音除去部409の出力となるスペクトルからケプストラムあるいは線形予測係数を計算したり、パワーを計算したり、これらの回帰係数などを計算したりして、音声特徴量を形成する。次に、ステップS509において、尤度計算部411がこの特徴量を入力し、尤度計算を行う。ステップS510において、入力音声の最終フレームまで解析が済むか、尤度計算により解が決定した場合には、ステップS511において認識結果出力部412が音声認識結果を出力する。そうでない場合にはステップS502へ戻り、次のフレームに進む。

30

【0041】

（その他の実施形態）

以上、本発明の実施形態を詳述した。上述した実施形態では、音声認識装置を説明したが、本発明はとりわけ、音声分析部に含まれる雑音抑圧機能にその特徴があり、これを単体の装置とした雑音抑圧装置自体が本発明を構成しうる。この雑音抑圧装置は音声認識装置のみならず音声録音再生装置や話者認識装置などさまざまな音声処理装置に適用することができる。

40

【0042】

また、本発明は、複数の機器から構成されるシステムに適用してもよいし、また、一つの機器からなる装置に適用してもよい。

【0043】

なお、本発明は、前述した実施形態の機能を実現するソフトウェアのプログラムを、システムあるいは装置に直接あるいは遠隔から供給し、そのシステムあるいは装置のコンピ

50

ュータがその供給されたプログラムコードを読み出して実行することによっても実現される。その場合、プログラムの機能を有していれば、その形態はプログラムである必要はない。

【 0 0 4 4 】

従って、本発明の機能処理をコンピュータで実現するために、そのコンピュータにインストールされるプログラムコード自体およびそのプログラムを格納した記憶媒体も本発明を構成することになる。つまり、本発明の特許請求の範囲には、本発明の機能処理を実現するためのコンピュータプログラム自体、およびそのプログラムを格納した記憶媒体も含まれる。

【 0 0 4 5 】

その場合、プログラムの機能を有していれば、オブジェクトコード、インタプリタにより実行されるプログラム、OSに供給するスクリプトデータ等、プログラムの形態を問わない。

【 0 0 4 6 】

プログラムを供給するための記憶媒体としては、例えば、フレキシブルディスク、ハードディスク、光ディスク、光磁気ディスク、MO、CD-ROM、CD-R、CD-RW、磁気テープ、不揮発性のメモリカード、ROM、DVD(DVD-ROM, DVD-R)

【 0 0 4 7 】

その他、プログラムの供給方法としては、クライアントコンピュータのブラウザを用いてインターネットのホームページに接続し、そのホームページから本発明のコンピュータプログラムそのもの、もしくは圧縮され自動インストール機能を含むファイルをハードディスク等の記憶媒体にダウンロードすることによっても供給できる。また、本発明のプログラムを構成するプログラムコードを複数のファイルに分割し、それぞれのファイルを異なるホームページからダウンロードすることによっても実現可能である。つまり、本発明の機能処理をコンピュータで実現するためのプログラムファイルを複数のユーザに対してダウンロードさせるWWWサーバも、本発明のクレームに含まれるものである。

【 0 0 4 8 】

また、本発明のプログラムを暗号化してCD-ROM等の記憶媒体に格納してユーザに配布し、所定の条件をクリアしたユーザに対し、インターネットを介してホームページから暗号化を解く鍵情報をダウンロードさせ、その鍵情報を使用することにより暗号化されたプログラムを実行してコンピュータにインストールさせて実現することも可能である。

【 0 0 4 9 】

また、コンピュータが、読み出したプログラムを実行することによって、前述した実施形態の機能が実現される他、そのプログラムの指示に基づき、コンピュータ上で稼動しているOSなどが、実際の処理の一部または全部を行い、その処理によっても前述した実施形態の機能が実現され得る。

【 0 0 5 0 】

さらに、記憶媒体から読み出されたプログラムが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によっても前述した実施形態の機能が実現される。

【 図面の簡単な説明 】

【 0 0 5 1 】

【 図 1 】 本発明の第 1 の実施形態に係る音声認識装置の機能構成を示すブロック図である。

【 図 2 】 本発明の第 1 の実施形態における音声認識処理を示すフローチャートである。

【 図 3 】 本発明の第 1 の実施形態における信号対雑音比とk(SNR)の関係を示すグラフである。

10

20

30

40

50

【図 4】本発明の第 2 の実施形態に係る音声認識装置の機能構成を示すブロック図である。

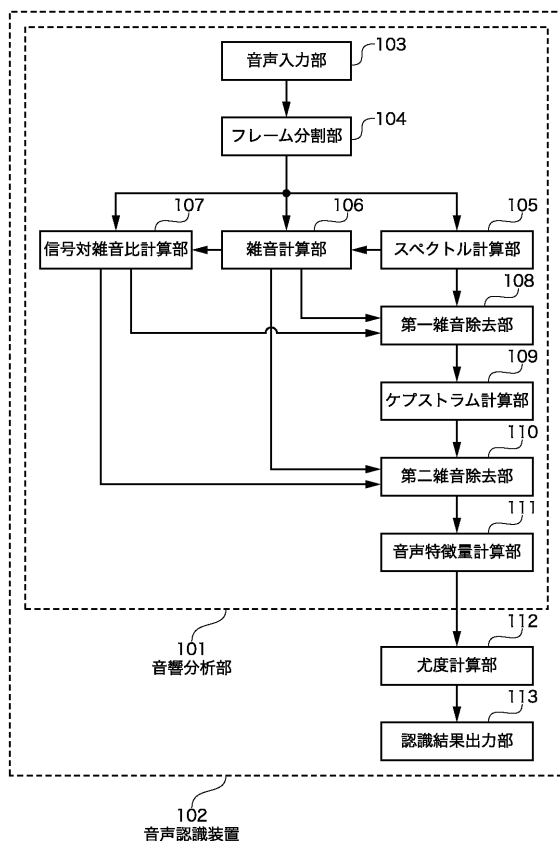
【図 5】本発明の第 2 の実施形態における音声認識処理を示すフローチャートである。

【図 6】本発明の第 2 の実施形態における信号対雑音比と隣接サンプル数 L の関係を示すグラフである。

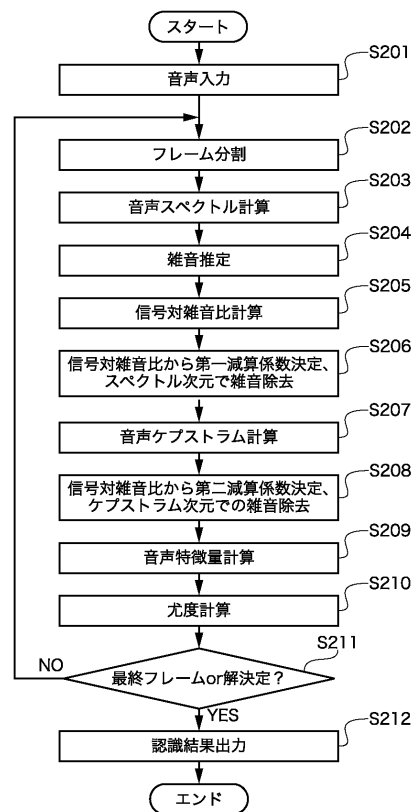
【図 7】本発明の第 2 の実施形態における雑音スペクトルの各周波数成分を平滑化する様子を表す図である。

【図 8】従来の SS 法を説明する図である。

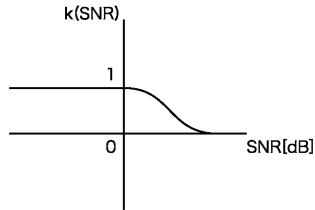
【図 1】



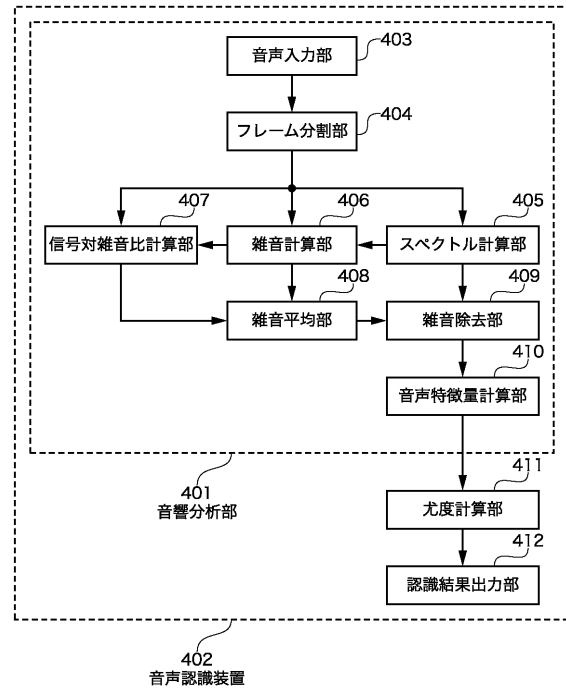
【図 2】



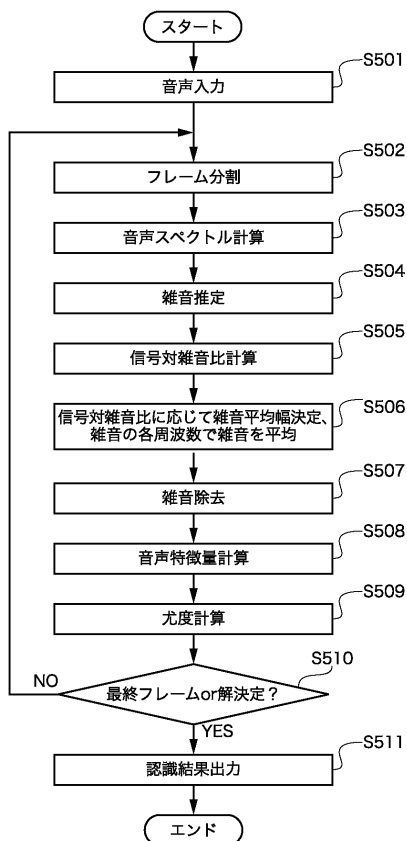
【図 3】



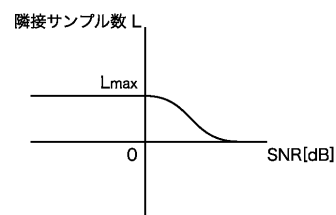
【図 4】



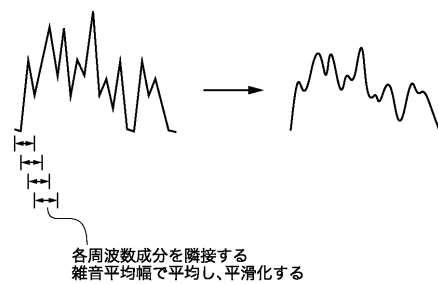
【図 5】



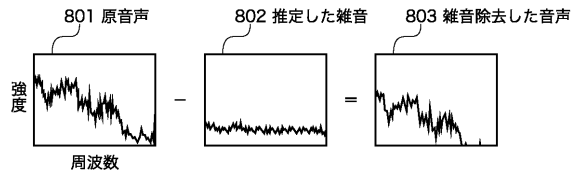
【図 6】



【図 7】



【 図 8 】



$$X_{S+N} - \alpha X_N = X_S$$

804

$\left[\begin{array}{l} X_{S+N} : \text{原音声のスペクトル} \\ X_N : \text{推定した雑音のスペクトル} \\ X_S : \text{雑音除去した音声のスペクトル} \\ \alpha : \text{減算係数} \end{array} \right]$
--

フロントページの続き

(72)発明者 深田 俊明

東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

Fターム(参考) 5D015 AA01 BB02 CC03 CC11 EE05