

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4512848号  
(P4512848)

(45) 発行日 平成22年7月28日(2010.7.28)

(24) 登録日 平成22年5月21日(2010.5.21)

(51) Int.Cl. F I  
**G 1 0 L 21/02 (2006.01)** G 1 0 L 21/02 1 0 2 B  
**G 1 0 L 15/20 (2006.01)** G 1 0 L 15/20 3 7 0 D

請求項の数 6 (全 22 頁)

<p>(21) 出願番号 特願2005-10639(P2005-10639)                  (22) 出願日 平成17年1月18日(2005.1.18)                  (65) 公開番号 特開2006-201287(P2006-201287A)                  (43) 公開日 平成18年8月3日(2006.8.3)                  審査請求日 平成18年9月26日(2006.9.26)</p> <p>特許法第30条第1項適用 平成16年9月21日 社団法人日本音響学会発行の「日本音響学会2004年秋季研究発表会講演論文集-1-」に発表</p> <p>(出願人による申告)平成16年度独立行政法人情報通信研究機構、研究テーマ「大規模コーパスベース音声対話翻訳技術の研究開発」に関する委託研究、産業活力再生特別措置法第30条の適用を受ける特許出願</p>	<p>(73) 特許権者 393031586                  株式会社国際電気通信基礎技術研究所                  京都府相楽郡精華町光台二丁目2番地2</p> <p>(74) 代理人 100099933                  弁理士 清水 敏</p> <p>(72) 発明者 藤本 雅清                  京都府相楽郡精華町光台二丁目2番地2                  株式会社国際電気通信基礎技術研究所内</p> <p>(72) 発明者 中村 哲                  京都府相楽郡精華町光台二丁目2番地2                  株式会社国際電気通信基礎技術研究所内</p> <p>審査官 涌井 智則</p>
---	---

最終頁に続く

(54) 【発明の名称】 雑音抑圧装置及び音声認識システム

(57) 【特許請求の範囲】

【請求項1】

目的音声と雑音とが重畳した雑音重畳音声について所定周期ごとにフレーム化された所定時間長のフレームよりそれぞれ抽出される特徴量を受けて、前記雑音重畳音声に関する所定の状態空間モデルと、前記目的音声に関する所定の音響モデルとを基に、複数のパーティクルを有するパーティクルフィルタを用いて、前記雑音の推定パラメータを前記フレームごとに逐次生成するための雑音推定手段と、

前記雑音重畳音声の特徴量と、前記雑音の推定パラメータと、前記音響モデルとを用いて、前記フレームごとに前記目的音声の推定特徴量を算出するための目的音声推定手段とを含む、雑音抑圧装置であって、

前記雑音推定手段は、

前記雑音重畳音声の特徴量を基に前記雑音の初期分布を推定し、当該初期分布にしたがった確率で、前記複数のパーティクルの各々における雑音の初期パラメータをそれぞれサンプリングするための初期パラメータ推定手段と、

前記音響モデルと前記雑音重畳音声の特徴量とを基に、各パーティクルにおいて、前記音響モデルからの出力パラメータのサンプリング結果に対して拡張カルマンフィルタを適用して、各パーティクルにおける先行する第1のフレームの雑音の推定パラメータをそれぞれ前記第1のフレームに後続する第2のフレームに対応するものに更新するための更新手段と、

前記第2のフレームにおける前記複数のパーティクルの各々の重みを算出するための重

み算出手段と、

前記更新手段により更新された雑音の推定パラメータと前記重み算出手段により算出された重みとにより前記複数のパーティクルの各々について表現される雑音パラメータの確率分布に基づき、前記複数のパーティクルにおける前記第2のフレームの雑音の推定パラメータをそれぞれ再サンプリングするための再サンプリング手段とを含む、雑音抑圧装置

。

【請求項2】

前記雑音推定手段はさらに、

前記再サンプリング手段により再サンプリングされたパラメータを基に、前記複数のパーティクルの各々において、前記第1のフレームに対応する前記雑音の推定パラメータをそれぞれ、前記第2のフレームに対応するものに再更新するための再更新手段と、

10

前記複数のパーティクルの各々において、前記再更新手段により再更新された前記雑音の推定パラメータと、前記再サンプリング手段により再サンプリングされた前記雑音の推定パラメータとの一方を、所定の判定基準にしたがい前記第2のフレームの雑音の推定パラメータとして選択するための選択手段とを含む、請求項1に記載の雑音抑圧装置。

【請求項3】

前記目的音声推定手段は、

前記雑音重畳音声の特徴量と、前記雑音の推定パラメータと、前記音響モデルとを基に、前記フレームに対応する前記雑音重畳音声の確率モデルを合成するための雑音重畳音声モデル合成手段と、

20

前記雑音重畳音声の特徴量、前記雑音の推定パラメータ、前記音響モデル、及び前記雑音重畳音声の確率モデルを基に、前記フレームごとに、前記目的音声の推定特徴量を算出するための推定特徴量算出手段とを含む、請求項1に記載の雑音抑圧装置。

【請求項4】

前記雑音重畳音声モデル合成手段は、前記複数のパーティクルの各々に対して、前記雑音の推定パラメータと、前記音響モデルとを基に、当該パーティクルにおける雑音重畳音声の確率モデルのパラメータを推定するためのパラメータ推定手段を含む、請求項3に記載の雑音抑圧装置。

【請求項5】

前記推定特徴量算出手段は、

30

前記フレームごとに、前記複数のパーティクルの各々の前記目的音声の推定パラメータを、前記雑音重畳音声の特徴量、前記音響モデル、前記雑音の推定パラメータ、及び前記雑音重畳音声の確率モデルを基に算出するための手段と、

前記複数のパーティクルの各々における前記目的音声の推定パラメータを基に、当該フレームにおける目的音声の推定特徴量を算出するための手段とを含む、請求項3又は請求項4に記載の雑音抑圧装置。

【請求項6】

請求項1～請求項5のいずれかに記載の雑音抑圧装置と、

前記雑音抑圧装置により算出される前記目的音声の推定特徴量を受けて、前記目的音声に関する前記所定の音響モデルと、認識対象言語に関する所定の言語モデルとを用いて、前記目的音声に関する音声認識を行なうための音声認識手段とを含む、音声認識システム

40

。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、雑音下での音声認識技術に関し、特に、非定常雑音下での音声認識精度を改善するための雑音抑圧装置及びそれを使用した音声認識システムに関する。

【背景技術】

【0002】

人間にとって容易かつ自然なヒューマンマシンインタフェースを実現するための技術と

50

して、音声認識技術が研究されている。近年では、大規模な音声・テキストデータベースと統計確率的な音声認識手法とにより、高い認識率での音声認識が実現されるようになった。今日ではさらに、人間が機械と接する実環境において、高速にかつ高い認識率で音声認識を実現するための応用技術開発が進められている。

【0003】

実環境が実験室等の環境と大きく異なる点のひとつに、雑音の存在がある。雑音は、無視できない音量で絶え間なく不規則に発生する。雑音下での音声認識性能の改善は、音声認識の応用技術開発を行なう上で、早急に解決されるべき問題である。

【0004】

この問題を解決するための技術のひとつに、音声認識の前処理の段階で雑音を推定し抑圧する技術がある。後掲の非特許文献1には、雑音抑圧の一般的な方法であるスペクトルサブトラクション法が開示されている。この方法では、発話の前の区間において観測された雑音の振幅スペクトルが発話中の区間における雑音の振幅スペクトルと同じであると仮定し、発話から得られた音声信号の振幅スペクトルから、発話直前に観測された雑音の振幅スペクトルを減算することで雑音を抑圧する。

10

【0005】

音声認識の前処理の段階において雑音を逐次的に推定し抑圧する技術もある。非特許文献2には、逐次EM (Expectation Maximization) アルゴリズムを適用して雑音の最尤推定値を逐次的に求める手法が開示されている。逐次EMアルゴリズムを用いて逐次的に雑音を推定する手法では、雑音の時間変動に対処しつつ高精度に雑音の推定及び抑圧を行なうことができる。

20

【0006】

非特許文献3及び非特許文献4に開示された、カルマンフィルタを用いて雑音の推定値を逐次的に求める手法も一般的に用いられている。この手法では、一期先予測とフィルタリングとを交互に行なうことによって、雑音を逐次的に推定し抑圧する。

【0007】

また、この問題を解決するための技術として、雑音を考慮した確率モデルを用いて適応的に音声認識を行なう技術がある。例えば後掲の特許文献1には、パーティクルフィルタと呼ばれる逐次推定法を用いて、雑音パラメータの推定と、隠れマルコフモデル (HMM: Hidden Markov Model) を構成する隠れ状態の時間的成長とを行ない、当該隠れマルコフモデルに基づく音声認識を行なう音声認識システムが開示されている。

30

【0008】

【特許文献1】特開2002-251198号公報

【非特許文献1】S. F. ボル: 「スペクトルサブトラクションを用いた、音声内の音響ノイズの抑圧」、IEEE Trans. ASSP, Vol. 27, No. 2, 113-120頁、1979年 (S.F. Boll: "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," IEEE Trans. ASSP, Vol. 27, No. 2, pp. 113-120, 1979)

【非特許文献2】M. アフィフィ、O. シオアン: 「ロバスト音声認識のための最適な忘却による逐次推定」、IEEE Trans. SAP, Vol. 12, No. 1, 19-26頁、2004年 (M. Afify, O. Siohan: "Sequential Estimation with Optimal Forgetting for Robust Speech Recognition," IEEE Trans. SAP, Vol. 12, No.1, pp. 19-26, 2004)

40

【非特許文献3】有本 卓: 「カルマンフィルタ」、産業図書

【非特許文献4】中野 道雄 監修、西山 清 著: 「パソコンで解くカルマンフィルタ」、丸善

【発明の開示】

【発明が解決しようとする課題】

【0009】

実環境において雑音の多くは非定常な雑音である。すなわち、雑音の音響的特徴は時間

50

の経過に伴い変動する。非特許文献 1 に記載のスペクトルサブトラクション法のように、雑音が定常的なものであるという前提のもとで雑音の推定及び抑圧を行なう技術では、雑音の時間変動に対応できず、高精度に雑音を抑圧することができない。

【 0 0 1 0 】

非特許文献 2 に記載された逐次 E M アルゴリズムを用いた手法は、尤度関数の局所最適値に値が収束するまで反復計算を行なう。そのため、雑音の変動する度に膨大な量の計算が必要となり、計算に時間を要する。よって、この手法により実時間で雑音を推定し抑圧するのは困難である。

【 0 0 1 1 】

非特許文献 3 及び非特許文献 4 に開示されたカルマンフィルタを用いる推定方法は、一期先予測とフィルタリングとを交互に行ない逐次推定を行なう。そのため、逐次 E M アルゴリズムのような反復計算を必要とはしない。しかし、カルマンフィルタを用いた手法は、雑音の事後確率分布が単一正規分布であるものとして確率分布を推定する。真の確率分布が混合分布であった場合には、単一正規分布で近似される。そのため、精度が劣化する。

10

【 0 0 1 2 】

特許文献 1 に記載の音声認識システムのように、雑音を考慮したモデルを用いて音声認識を行なう技術では、雑音が重畳された音声と確率モデルとのマッチングが行なわれる。そのため、例えば音響モデル適応など、雑音のない音声に基づいて行なうべき前処理を実行できない。

20

【 0 0 1 3 】

それゆえに、本発明の目的は、非定常雑音下での音声認識性能を改善し、かつ雑音の抑圧を短時間で行なうことができる雑音抑圧装置を提供することである。

【課題を解決するための手段】

【 0 0 1 4 】

本発明の第 1 の局面に係る雑音抑圧装置は、目的音声と雑音とが重畳した雑音重畳音声について所定周期ごとにフレーム化された所定時間長のフレームよりそれぞれ抽出される特徴量を受けて、雑音重畳音声に関する所定の状態空間モデルと、目的音声に関する所定の音響モデルとを基に、複数のパーティクルを有するパーティクルフィルタを用いて、雑音の推定パラメータをフレームごとに逐次生成するための雑音推定手段と、雑音重畳音声の特徴量と雑音の推定パラメータ、及び音響モデルを用いて、フレームごとに目的音声の推定特徴量を算出するための目的音声推定手段とを含む。

30

【 0 0 1 5 】

好ましくは、雑音推定手段は、雑音重畳音声の特徴量を基に雑音の初期分布を推定し、初期分布にしたがった確率で、複数のパーティクルの各々における雑音の初期パラメータをそれぞれサンプリングするための初期パラメータ推定手段と、音響モデルと雑音重畳音声の特徴量とを基に、拡張カルマンフィルタを用いて、各パーティクルにおける先行する第 1 のフレームの雑音の推定パラメータをそれぞれ第 1 のフレームに後続する第 2 のフレームに対応するものに更新するための更新手段と、第 2 のフレームにおける複数のパーティクルの各々の重みを算出するための重み算出手段と、更新手段により更新された雑音の推定パラメータと重み算出手段により算出された重みとにより複数のパーティクルの各々について表現される雑音パラメータの確率分布に基づき、複数のパーティクルにおける第 2 のフレームの雑音の推定パラメータをそれぞれ再サンプリングするための再サンプリング手段とを含む。

40

【 0 0 1 6 】

より好ましくは、雑音推定手段はさらに、再サンプリング手段により再サンプリングされたパラメータを基に、複数のパーティクルの各々において、第 1 のフレームに対応する雑音の推定パラメータをそれぞれ、第 2 のフレームに対応するものに再更新するための再更新手段と、複数のパーティクルの各々において、再更新手段により再更新された雑音の推定パラメータと、再サンプリング手段により再サンプリングされた雑音の推定パラメータ

50

タとの一方を、所定の判定基準にしたがい第2フレームの雑音の推定パラメータとして選択するための選択手段とを含む。

【0017】

好ましくは、目的音声推定手段は、雑音重畳音声の特徴量と、雑音の推定パラメータと、音響モデルとを基に、フレームに対応する雑音重畳音声の確率モデルを合成するための雑音重畳音声モデル合成手段と、雑音重畳音声の特徴量、雑音の推定パラメータ、音響モデル、及び雑音重畳音声の確率モデルを基に、フレームごとに、目的音声の推定特徴量を算出するための推定特徴量算出手段とを含む。

【0018】

雑音重畳音声モデル合成手段は、複数のパーティクルの各々に対して、雑音の推定パラメータと音響モデルとを基に、当該パーティクルにおける雑音重畳音声の確率モデルのパラメータを推定するための手段を含んでもよい。

10

【0019】

推定特徴量算出手段は、フレームごとに、複数のパーティクルの各々の目的音声の推定パラメータを、雑音重畳音声の特徴量、音響モデル、雑音の推定パラメータ、及び雑音重畳音声の確率モデルを基に算出するための手段と、複数のパーティクルの各々における目的音声の推定パラメータを基に、当該フレームにおける目的音声の推定特徴量を算出するための手段とを含んでもよい。

【0020】

本発明の第2の局面に係る音声認識システムは、本発明の第1の局面に係る上記したいずれかの雑音抑圧装置と、雑音抑圧装置により算出される目的音声の推定特徴量を受けて、目的音声に関する所定の音響モデルと、認識対象言語に関する所定の言語モデルとを用いて、目的音声に関する音声認識を行なうための認識手段とを含む。

20

【発明を実施するための最良の形態】

【0021】

以下、図面を参照しつつ、本発明の一実施の形態について説明する。なお、以下の説明に用いる図面では、同一の部品には同一の符号を付してある。それらの名称及び機能も同一である。したがって、それらについての説明は繰返さない。以下の説明のテキスト中で使用する記号「 $\wedge$ 」等は、本来はその直後の文字の直上に記載すべきものであるが、テキスト記法の制限により当該文字の直前に記載する。式中では、これらの記号等は本来の位置に記載してある。また以下の説明のテキスト中では、ベクトル又は行列については例えば「ベクトル $X_t$ 」、「行列 $W$ 」等のように直前に「ベクトル」、「行列」等を付した通常のテキストの形で記載するが、式中ではいずれも太字で記載する。

30

【0022】

[構成]

図1に、本実施の形態に係る音声認識システム100全体の構成を示す。図1を参照して、この音声認識システム100は、音源102が発生する音を収集し、収集した音から認識に用いる特徴量を抽出するための前処理部104と、音声と音素との関係を表す確率モデル(音響モデル)を準備するための音響モデル部106と、認識対象の言語における単語の接続確率等を表す確率モデル(言語モデル)を準備するための言語モデル部108と、前処理部104から出力された特徴量に対応する単語等を、音響モデル部106の音響モデルと言語モデル部108の言語モデルとを用いて探索するための探索部110とを含む。

40

【0023】

音源102は、認識されるべき音声(目的音声)を発話する話者116と、話者116の周囲の雑音源118とを含む。音源102が発生する音は、話者116の発話により発生する音声120と、雑音源118が発生する雑音122とが重畳した音である。本明細書では、話者116の発話により発生する雑音のない目的音声120を、「クリーン音声」と呼ぶ。クリーン音声120と雑音122とが重畳した音を「雑音重畳音声」と呼ぶ。

【0024】

50

音響モデル部 106 は、クリーン音声 120 に対する HMM からなる音響モデルを準備し保持する。音響モデル部 106 は、予め用意された大量の学習データを記憶するための学習データ記憶部 132 と、学習データ記憶部 132 に記憶された学習データを用いて HMM に対する学習処理を行なうためのモデル学習部 134 と、モデル学習部 134 による学習で得られるクリーン音声 HMM 130 を記憶するためのクリーン音声 HMM 記憶部 136 とを含む。

【0025】

図 2 に、クリーン音声 HMM 130 の概念を模式的に示す。図 2 を参照して、クリーン音声 HMM 130 は、複数の定常信号源（状態）の間の遷移により時系列信号をモデル化した確率モデルである。このクリーン音声 HMM 130 は、複数の状態 142A, 142B, 142C, 142D, ..., 142M を含む。クリーン音声 HMM 130 ではさらに、各状態に対してその状態から他の状態への遷移確率が定義される。例えば状態 142A に対して、自己すなわち状態 142A への遷移 144 に関する遷移確率と、状態 142B への遷移 146 に関する遷移確率とが定義されている。また、このクリーン音声 HMM 130 の各状態においては、各遷移に伴う出力の出力確率が定義される。本実施の形態においては、出力確率はそれぞれ混合正規分布によって表現される。例えば状態 142A は単一正規分布 148A, 148B, ..., 148K からなる混合正規分布を持つ。

【0026】

再び図 1 を参照して、前処理部 104 は、音源 102 が発生する雑音重畳音声を計測し、得られる計測信号に所定の信号処理を施すことにより雑音重畳音声に関する所定の特徴量ベクトル（以下単に「特徴量」と呼ぶ。）124 を抽出するための計測部 112 と、計測部 112 が抽出する特徴量 124 に含まれる雑音成分を、クリーン音声 HMM 130 を用いて抑圧する雑音抑圧処理を行なうための雑音抑圧部 114 とを含む。

【0027】

具体的には、計測部 112 は、雑音重畳音声を時間長数 10 ミリ秒のフレームごとに対数メルフィルタバンク分析し、得られる対数メル周波数スペクトルを要素とするベクトルを特徴量 124 として出力する。

【0028】

雑音抑圧部 114 は、クリーン音声 HMM 130 を用いて、雑音重畳音声の特徴量 124 からクリーン音声 120 の特徴量を推定する。そして推定により得られた特徴量 126 を探索部 110 に出力する。本明細書では、推定により得られる特徴量 126 によって表される音声を「推定クリーン音声」と呼ぶ。

【0029】

図 3 に、雑音重畳音声の観測過程及び雑音の状態変化過程を表現する状態空間モデル 160 を示す。図 3 を参照して、時刻  $t$  のフレームにおける雑音重畳音声の特徴量 124 を特徴量ベクトル  $X_t$  とする。特徴量ベクトル  $X_t$  は、計測により得られる既知のパラメータである。特徴量ベクトル  $X_t$  は、クリーン音声 120 からの成分であるベクトル  $S_t$  と、雑音 122 からの成分である雑音成分ベクトル  $N_t$  とからなる。ベクトル  $S_t$  及び  $N_t$  は、いずれも未知のパラメータである。

【0030】

上記したベクトル  $X_t$ 、 $S_t$  及び  $N_t$  の次元数は同一である。なお、以下に説明する処理はこれらベクトルの各次元の要素についてそれぞれ行なわれるが、以下の説明では、簡単のために各次元の要素を特に区別して言及することはしない。

【0031】

状態空間モデル 160 において、クリーン音声 120 の出力過程は HMM でモデル化できるものと仮定する。すなわち、時刻  $t$  におけるクリーン音声 120 の成分であるベクトル  $S_t$  は、クリーン音声 HMM 130 におけるある状態内のある確率分布にしたがって出力されるものと仮定する。

【0032】

クリーン音声 HMM 130 において、時刻  $t$  に対応する状態を  $s_t$  とし、状態  $s_t$  内の混

10

20

30

40

50

合確率分布を構成するある単一正規分布を  $k_t$  とする。また、分布  $k_t$  から出力されるパラメータのベクトルをベクトル  $S_{s_t, k_t, t}$  とする。以下、クリーン音声 HMM 130 から出力されるパラメータベクトル  $S_{s_t, k_t, t}$  を、「(クリーン音声 HMM 130 の) 出力パラメータ」と呼ぶ。クリーン音声 120 の特徴量ベクトル  $S_t$  と、出力パラメータベクトル  $S_{s_t, k_t, t}$  との間には誤差が存在する。この誤差もまたベクトルであり、このベクトルをベクトル  $V_t$  とする。雑音重畳音声の特徴量ベクトル  $X_t$  の観測過程は、クリーン音声 HMM 130 の出力パラメータベクトル  $S_{s_t, k_t, t}$ 、誤差ベクトル  $V_t$ 、及び雑音 122 の成分ベクトル  $N_t$  を用いて、次の式 (1) により表現される。

【0033】

【数1】

$$\begin{aligned} \mathbf{X}_t &= \mathbf{S}_{s_t, k_t, t} + \log(\mathbf{I} + \exp(\mathbf{N}_t - \mathbf{S}_{s_t, k_t, t})) + \mathbf{V}_t \\ &= f(\mathbf{S}_{s_t, k_t, t}, \mathbf{N}_t) + \mathbf{V}_t \end{aligned} \quad (1)$$

10

誤差ベクトル  $V_t$  は、次の式 (2) のように単一正規分布にしたがうものとする。

【0034】

【数2】

$$\mathbf{V}_t \sim \mathcal{N}(\mathbf{0}, \Sigma_{s_t, k_t}) \quad (2)$$

ただしこの式において  $S_{s_t, k_t}$  はクリーン音声 HMM 130 における状態  $s_t$ 、分布  $k_t$  より得られるパラメータの共分散行列を表わし、記号「 $\sim$ 」は左辺の値が右辺に示される確率分布にしたがうことを示し、「 $\mathcal{N}(\mu, \Sigma)$ 」は、平均値ベクトル  $\mu$ 、分散  $\Sigma$  の正規分布を表す。

20

【0035】

また状態空間モデル 160 において、雑音の成分ベクトル  $N_t$  は、ランダムウォーク過程にしたがって変化するものと仮定する。すなわち、時刻  $t-1$  における雑音 122 の成分ベクトル  $N_{t-1}$  と時刻  $t$  における雑音 122 の成分ベクトル  $N_t$  との間に誤差が生じるものと仮定する。この誤差をベクトル  $W_t$  とする。すると、ベクトル  $N_t$  の状態変化過程は、次の式 (3) により表現される。

【0036】

【数3】

$$\mathbf{N}_t = \mathbf{N}_{t-1} + \mathbf{W}_t \quad (3)$$

30

誤差ベクトル  $W_t$  は、次の式 (4) のように単一正規分布にしたがうものとする。

【0037】

【数4】

$$\mathbf{W}_t \sim \mathcal{N}(\mathbf{0}, \Sigma_w) \quad (4)$$

ただし、式 (4) において  $\Sigma_w$  は、誤差ベクトル  $W_t$  の共分散行列を表す。

【0038】

図1に示す雑音抑圧部 114 は、上記の式 (1) ~ 式 (4) により表現される状態空間モデル 160 を用いて、時刻  $t$  に対応するフレームごとに、クリーン音声の特徴量ベクトルを逐次推定する。

40

【0039】

図4に、雑音抑圧部 114 の構成をブロック図で示す。図4を参照して、雑音抑圧部 114 は、雑音重畳音声の特徴量  $X_t$  (124) を受けて、クリーン音声 HMM 130 を用いて状態空間モデル 160 における雑音 122 の確率分布 (以下、「雑音確率分布」と呼ぶ。) を推定するための雑音確率分布推定部 200 と、雑音確率分布推定部 200 により推定された雑音確率分布とクリーン音声 HMM 130 とから雑音重畳音声の確率モデルの平均ベクトルと共分散行列とを生成するためのパラメータ生成部 202 と、雑音確率分布、雑音重畳音声の平均ベクトル及び共分散行列、並びにクリーン音声 HMM 130 を用いて、推定クリーン音声の特徴量 126 を算出するためのクリーン音声推定部 204 とを含

50

む。

【 0 0 4 0 】

雑音確率分布推定部 2 0 0 は、雑音確率分布をフレームごとに逐次推定し、雑音確率分布を表すパラメータ 2 0 6 を出力する機能を持つ。ここに、雑音 1 2 2 の成分ベクトル  $N_0, \dots, N_t$  からなる系列をベクトル  $N_{0:t} = \{N_0, \dots, N_t\}$  とする。ベクトル  $N_{0:t}$  の事後確率分布  $p(N_{0:t} | X_{0:t})$  は、1 次マルコフ連鎖を用いて、次の式 (5) のように表される。

【 0 0 4 1 】

【数 5】

$$p(N_{0:t} | X_{0:t}) = p(N_0 | X_0) \prod_{t'=1}^t p(N_{t'} | N_{t'-1}) p(X_{t'} | N_{t'}) \quad (5)$$

10

式 (5) の  $p(N_t | N_{t-1})$  は、単一正規分布を用いて次の式 (6) のようにモデル化される。

【 0 0 4 2 】

【数 6】

$$p(N_t | N_{t-1}) = \mathcal{N}(N_t; N_{t-1}, \Sigma_w) \quad (6)$$

また、式 (5) の  $p(X_t | N_t)$  は、単一正規分布を用いて次の式 (7) のようにモデル化される。

20

【 0 0 4 3 】

【数 7】

$$p(X_t | N_t) = \mathcal{N}(X_t; f(S_{s_t, k_t, t}, N_t), \Sigma_{S, s_t, k_t}) \quad (7)$$

【 0 0 4 4 】

したがって、状態空間モデル 1 6 0 を基に雑音ベクトル  $N_t$  の確率分布を逐次推定する問題は、雑音重畳音声ベクトル  $X_t$  が与えられた時の出力確率を最大にするような事後確率分布  $p(N_t | N_{t-1})$  を推定する問題に帰着する。雑音確率分布推定部 2 0 0 は、雑音重畳音声ベクトル  $X_t$  と状態空間モデル 1 6 0 とに基づき、この推定を行なう。

【 0 0 4 5 】

雑音確率分布推定部 2 0 0 は、雑音の成分ベクトル  $N_t$  の確率分布を逐次的に推定する際に、パーティクルフィルタと呼ばれる手法を用いる。この推定法は、状態空間内に、局限された状態空間 (パーティクル) を多数生成して、各パーティクルにおいてパラメータの確率分布を推定し、状態空間内におけるパラメータの確率分布を、各パーティクルにおいて推定された確率分布を用いて近似的に表現する手法である。この手法では、多数のパーティクルにおける初期的なパラメータを、ランダムなサンプリングにより、又は当該パラメータの初期分布からのサンプリングにより決定する。そして、以下の処理をフレームごとに行なう。すなわち、あるフレームに対応して各パーティクルにおいてパラメータが決定されると、各パーティクルのパラメータを当該フレームに後続するフレームに対応するものに更新し、その更新の尤度に応じて各パーティクルに対して重みを付与する。そして、更新後のパーティクルにおけるパラメータの確率分布にしたがい、当該後続のフレームに対応する各パーティクルのパラメータを再サンプリングする。再サンプリングされたパラメータを基に、当該後続のフレームに対応する各パーティクルのパラメータを決定する。以上の処理をフレームごとに行なうことにより、逐次的に各パーティクルにおけるパラメータを決定する。状態空間におけるパラメータは、パーティクルにおけるパラメータの重み付き和によって近似的に表現される。すなわち、パーティクルの数を  $J$ 、 $j$  番目のパーティクルにおける雑音のパラメータをベクトル  $N_t^{(j)}$  とし、当該パーティクルに対する重みを  $w_t^{(j)}$  とすると、式 (5) に示す雑音の事後確率分布  $p(N_{0:t} | X_{0:t})$  は、次の式 (8) によって近似的に表現される。

30

40

【 0 0 4 6 】

【数 8】

$$p(\mathbf{N}_{0:t}|\mathbf{X}_{0:t}) \simeq \sum_{j=1}^J w_t^{(j)} p(\mathbf{N}_{0:t}^{(j)}|\mathbf{X}_{0:t}) \quad (8)$$

【0047】

パラメータ生成部 202 は、具体的には V T S (Vector Taylor Series) 法と呼ばれる H M M 合成法によって、パーティクルフィルタにより推定された雑音の分布を用い、複数のパーティクルにおける雑音重畳音声の特徴量ベクトル  $\mathbf{X}_t$  の平均ベクトル及び共分散行列 (208) をそれぞれ算出する機能を持つ。

【0048】

クリーン音声推定部 204 は、最小 2 乗誤差 (Minimum Mean Square Error : M M S E) 推定法で、フレームごとに、複数のパーティクルにおけるクリーン音声のパラメータをそれぞれ推定し、それら推定されたパラメータの重み付き和によって推定クリーン音声の特徴量 126 を算出する機能を持つ。クリーン音声推定部 204 はさらに、雑音確率分布推定部 200 に、次のフレームへの移行に関する要求 210 を発行する機能を持つ。

【0049】

図 5 に、雑音確率分布推定部 200 の構成をブロック図で示す。図 5 を参照して、雑音確率分布推定部 200 は、雑音重畳音声の特徴量 124 とクリーン音声推定部 204 からの要求 210 とを受けて、処理対象となるフレームを選択し、当該フレームにおける雑音重畳音声の特徴量 124 をフレームに応じた出力先に出力するためのフレーム選択部 220 と、フレーム選択部 220 から最初の所定フレーム分の雑音重畳音声の特徴量 124 を受けて雑音の初期分布を推定し、各パーティクルにおける雑音の初期的なパラメータを決定するための雑音初期分布推定部 222 と、フレーム選択部 220 から時刻  $t$  ( $t-1$ ) のフレームにおける雑音重畳音声の特徴量 124 を受けて、逐次的に、パーティクルにおける雑音のパラメータと当該パーティクルに対する重みとを算出するための逐次計算部 224 とを含む。

【0050】

雑音初期分布推定部 222 は、雑音の初期分布  $p(\mathbf{N}_0)$  が単一正規分布であるものとみなして雑音の初期分布を推定する。雑音の初期分布における平均ベクトルを  $\mu_N$  とし、共分散行列を行列  $\Sigma_N$  とすると、雑音の初期分布  $p(\mathbf{N}_0)$  は次の式 (9) のように表される。

【0051】

【数 9】

$$p(\mathbf{N}_0) = \mathcal{N}(\mathbf{N}_0; \mu_N, \Sigma_N) \quad (9)$$

雑音初期分布推定部 222 は、最初の所定フレーム分の区間の雑音重畳音声の特徴量ベクトル  $\mathbf{X}_t$  が雑音の成分のみからなるものとみなし、雑音の初期分布の平均ベクトル  $\mu_N$ 、及び共分散行列  $\Sigma_N$  を推定する。例えば 0 ~ 9 の 10 フレーム分の区間がこの区間に該当する場合、雑音初期分布推定部 222 は、平均ベクトル  $\mu_N$  及び共分散行列  $\Sigma_N$  をそれぞれ、次の式 (10) 及び式 (11) によって算出する。ただし、ベクトルの右肩に付した「T」は転置を表す。

【0052】

【数 10】

$$\mu_N = \frac{1}{10} \sum_{t=0}^9 \mathbf{X}_t \quad (10)$$

$$\Sigma_N = \frac{1}{10} \sum_{t=0}^9 (\mathbf{X}_t - \mu_N) (\mathbf{X}_t - \mu_N)^T \quad (11)$$

雑音初期分布推定部 222 はさらに、各パーティクルにおける雑音の初期パラメータベクトル  $\mathbf{N}_0^{(j)}$  を、初期分布  $p(\mathbf{N}_0)$  からのサンプリングによって生成する。すなわち、式 (12) と式 (13) とに示すサンプリングを、パーティクル  $j$  ( $1 \leq j \leq J$ ) ごとに行

10

20

30

40

50

なう。

【 0 0 5 3 】

【 数 1 1 】

$$\mathbf{N}_0^{(j)} \sim p(\mathbf{N}_0) \quad (12)$$

$$\Sigma_{\mathbf{N}_0}^{(j)} = \Sigma_{\mathbf{N}} \quad (13)$$

【 0 0 5 4 】

逐次計算部 2 2 4 は、クリーン音声 HMM 1 3 0 の出力パラメータ 1 4 0 をサンプリングするためのクリーン音声 HMM サンプリング部 2 2 6 と、時刻  $t$  における雑音重畳音声の特徴量 1 2 4 を受け、各パーティクルにおける雑音のパラメータを更新するための更新部 2 3 0 と、更新後のパーティクルに対する重みをそれぞれ算出するための重み算出部 2 3 2 と、重み算出部 2 3 2 により算出された重みに基づき、パーティクルにおける雑音のパラメータを再サンプリングするための再サンプリング部 2 3 4 と、再サンプリングされたパーティクルにおける雑音のパラメータと時刻  $t - 1$  における各パーティクルにおける雑音のパラメータとに基づき、各パーティクルにおける雑音のパラメータを決定し、推定雑音分布 2 0 6 を生成するための推定雑音分布生成部 2 3 6 とを含む。

10

【 0 0 5 5 】

更新部 2 3 0 は、状態空間モデル 1 6 0 ( 図 3 ) を基に構成される拡張カルマンフィルタを用いて、時刻  $t - 1$  に対応するパーティクルにおける雑音のパラメータを、時刻  $t$  に対応するものに更新する機能を持つ。拡張カルマンフィルタは、式 ( 1 ) に示すように非線形項を含む状態空間モデルに対応したカルマンフィルタである。本実施の形態における拡張カルマンフィルタの分布更新式を、以下の式 ( 1 4 ) ~ 式 ( 1 9 ) に示す。なお、これらの数式において時刻  $t - 1$  に対応するパラメータから予測される時刻  $t$  におけるパラメータについては添え字として「 $_{t|t-1}$ 」を付してある。

20

【 0 0 5 6 】

【 数 1 2 】

$$\mathbf{N}_{t|t-1}^{(j)} = \hat{\mathbf{N}}_{t-1}^{(j)} \quad (14)$$

$$\Sigma_{\mathbf{N}_{t|t-1}}^{(j)} = \Sigma_{\mathbf{N}_{t-1}}^{(j)} + \Sigma_{\mathbf{w}} \quad (15)$$

$$\mathbf{K}_t^{(j)} = \Sigma_{\mathbf{N}_{t|t-1}}^{(j)} \mathbf{F}_t^{(j)T} \left[ \mathbf{F}_t^{(j)} \Sigma_{\mathbf{N}_{t|t-1}}^{(j)} \mathbf{F}_t^{(j)T} + \Sigma_{\mathbf{s}, \mathbf{s}_t^{(j)}, \mathbf{k}_t^{(j)}} \right]^{-1} \quad (16)$$

30

$$\mathbf{F}_t^{(j)} = \partial f \left( \mathbf{s}_{\mathbf{s}_t^{(j)}, \mathbf{k}_t^{(j)}, t}^{(j)}, \mathbf{N}_{t|t-1}^{(j)} \right) / \partial \mathbf{N}_{t|t-1}^{(j)} \quad (17)$$

$$\hat{\mathbf{N}}_t^{(j)} = \mathbf{N}_{t|t-1}^{(j)} + \mathbf{K}_t^{(j)} \left( \mathbf{X}_t - f \left( \mathbf{s}_{\mathbf{s}_t^{(j)}, \mathbf{k}_t^{(j)}, t}^{(j)}, \mathbf{N}_{t|t-1}^{(j)} \right) \right) \quad (18)$$

$$\Sigma_{\mathbf{N}_t}^{(j)} = \Sigma_{\mathbf{N}_{t|t-1}}^{(j)} - \mathbf{K}_t^{(j)} \mathbf{F}_t^{(j)} \Sigma_{\mathbf{N}_{t|t-1}}^{(j)} \quad (19)$$

ただし、式 ( 1 6 ) ~ 式 ( 1 8 ) のベクトル  $\mathbf{s}_{\mathbf{s}_t^{(j)}, \mathbf{k}_t^{(j)}, t}^{(j)}$  は、 $j$  番目のパーティクルにおいてクリーン音声 HMM 1 3 0 ( 図 2 参照 ) の出力パラメータベクトル  $\mathbf{s}_{\mathbf{s}_t, \mathbf{k}_t, t}$  に相当するパラメータである。また前述した通り、行列  $\mathbf{w}$  は、時刻  $t - 1$  から時刻  $t$  への状態変化の際に雑音ベクトル  $\mathbf{N}_t$  に生じる誤差ベクトル  $\mathbf{W}_t$  の共分散行列を表す。

40

【 0 0 5 7 】

クリーン音声 HMM サンプリング部 2 2 6 は、クリーン音声 HMM 1 3 0 ( 図 2 参照 ) 内の状態  $\mathbf{s}_{t-1}^{(j)}$  を、状態  $\mathbf{s}_{t-1}^{(j)}$  からの遷移確率にしたがってサンプリングし、サンプリングされた状態  $\mathbf{s}_t^{(j)}$  内の混合分布から、単一正規分布  $\mathbf{k}_t^{(j)}$  をその混合重みに基づいてサンプリングする。クリーン音声 HMM サンプリング部 2 2 6 はさらに、サンプリングされた単一正規分布  $\mathbf{k}_t^{(j)}$  から出力パラメータベクトル  $\mathbf{s}_{\mathbf{s}_t^{(j)}, \mathbf{k}_t^{(j)}, t}^{(j)}$  を確率分布にしたがってサンプリングして、更新部 2 3 0 に与える。

【 0 0 5 8 】

クリーン音声 HMM 1 3 0 内の状態  $\mathbf{s}_{t-1}^{(j)}$  から状態  $\mathbf{s}_t$  への遷移確率をベクトル  $\mathbf{a}_{\mathbf{s}_{t-1}^{(j)}, \mathbf{s}_t}$  とすると、状態  $\mathbf{s}_t^{(j)}$  は、クリーン音声 HMM 1 3 0 から次の式 ( 2 0 ) に

50

示すサンプリングによって得られる。

【 0 0 5 9 】

【 数 1 3 】

$$s_t^{(j)} \sim a_{\mathbf{S}, s_{t-1}^{(j)}, s_t} \quad (20)$$

状態  $s_t^{(j)}$  の混合分布における単一正規分布  $k_t$  の混合重みを  $P_{\mathbf{S}, s_t^{(j)}, k_t}$  とすると、分布  $k_t^{(j)}$  は、状態  $s_t^{(j)}$  から次の式 ( 2 1 ) に示すサンプリングによって得られる。

【 0 0 6 0 】

【 数 1 4 】

$$k_t^{(j)} \sim P_{\mathbf{S}, s_t^{(j)}, k_t} \quad (21)$$

10

分布  $k_t^{(j)}$  の平均ベクトルをベクトル  $\mu_{\mathbf{S}, s_t^{(j)}, k_t^{(j)}}$  とし、分布  $k_t^{(j)}$  の共分散行列を行列  $\Sigma_{\mathbf{S}, s_t^{(j)}, k_t^{(j)}}$  とすると、 $j$  番目のパーティクルにおけるクリーン音声 HMM 1 3 0 の出力パラメータベクトル  $\mathbf{S}_{s_t^{(j)}, k_t^{(j)}, t}^{(j)}$  は、分布  $k_t^{(j)}$  から、次の式 ( 2 2 ) に示すサンプリングによって得られる。

【 0 0 6 1 】

【 数 1 5 】

$$\mathbf{S}_{s_t^{(j)}, k_t^{(j)}, t}^{(j)} \sim \mathcal{N} \left( \mu_{\mathbf{S}, s_t^{(j)}, k_t^{(j)}}, \Sigma_{\mathbf{S}, s_t^{(j)}, k_t^{(j)}} \right) \quad (22)$$

なお、フレーム選択部 2 2 0 はさらに、クリーン音声 HMM サンプリング部 2 2 6 に対し、時刻  $t$  におけるクリーン音声 HMM の出力パラメータのサンプリングを要求する機能を持つ。

20

【 0 0 6 2 】

重み算出部 2 3 2 は、時刻  $t$  での雑音重畳音声の特徴量ベクトル  $\mathbf{X}_t$  と、時刻  $t$  の各パーティクルにおけるクリーン音声 1 3 0 のパラメータベクトル  $\mathbf{S}_{s_t^{(j)}, k_t^{(j)}, t}^{(j)}$ 、及び雑音のパラメータベクトル  $\mathbf{N}_t^{(j)}$  と、時刻  $t - 1$  のパーティクルに対する重み  $w_{t-1}^{(j)}$  とを基に、次の式 ( 2 3 ) 及び式 ( 2 4 ) に示す算出方法を用いて、時刻  $t$  のパーティクルに対する重み  $w_t^{(j)}$  を算出する機能を持つ。

【 0 0 6 3 】

【 数 1 6 】

$$w_t^{(j)} \propto w_{t-1}^{(j)} p(\mathbf{X}_t | \mathbf{N}_t^{(j)}) \quad (23)$$

30

$$p(\mathbf{X}_t | \mathbf{N}_t^{(j)}) = \mathcal{N} \left( \mathbf{X}_t; f \left( \mathbf{S}_{s_t^{(j)}, k_t^{(j)}, t}^{(j)}, \mathbf{N}_t^{(j)} \right), \Sigma_{\mathbf{S}, s_t^{(j)}, k_t^{(j)}} \right) \quad (24)$$

なお、重み  $w_t^{(j)}$  (  $1 \leq j \leq J$  ) は、 $\sum_{j=1}^J w_t^{(j)} = 1$  となるように正規化される。

【 0 0 6 4 】

再サンプリング部 2 3 4 は、パラメータが更新されたパーティクルにおける雑音のパラメータの確率分布にしたがい、時刻  $t$  に対応する各パーティクルにおける雑音のパラメータベクトル  $\mathbf{N}_t^{(j)}$  を再サンプリングする機能を持つ。この際、再サンプリング部 2 3 4 は、微小な重み  $w_t^{(j)}$  しか与えられていないパーティクルにおける確率分布からは、パラメータの再サンプリングを行なわない。一方、大きな重み  $w_t^{(j)}$  が与えられているパーティクルにおける確率分布からは、パラメータを重み  $w_t^{(j)}$  の大きさに応じた回数の再サンプリングを行ない、得られたパラメータをそれぞれ、当該再サンプリングの回数と同数のパーティクルに割当てて。ただし再サンプリングの全回数及びパーティクルの全数は一定 (  $J$  ) である。このようにするのは、各パーティクルに割当てられる重みが、式 ( 2 3 ) から分かるように観測された特徴量ベクトル  $\mathbf{X}_t$  の尤度に対応しているからである。

40

【 0 0 6 5 】

推定雑音分布生成部 2 3 6 は、Metropolis-Hastings アルゴリズムにより、時刻  $t$  のフレームに対応するパーティクルの再生成を行なう機能を持つ。図 6 に、推定雑音分布生成部 2 3 6 の構成をブロック図で示す。図 6 を参照して、推定雑音分布生成部 2 3 6 は、再

50

サンプリング部 234 による再サンプリングで得られた各パーティクルにおける雑音の確率分布を用いて状態空間モデル 160 における雑音の確率分布を表し、当該表した確率分布に基づき、時刻  $t-1$  に対応するパーティクルにおける雑音のパラメータを時刻  $t$  に対応するものへ、上記の式 (14) ~ 式 (19) に示す拡張カルマンフィルタを用いて再更新するための再更新部 262 と、再更新されたパーティクルに対する重み (これを以下「 $w_t^{*(j)}$ 」とする。) を上記の式 (23) 及び式 (24) に示す算出方法を用いて算出するための重み再計算部 264 と、再サンプリングされたパーティクルに対する重み  $w_t^{(j)}$  及び再更新されたパーティクルに対する重み  $w_t^{*(j)}$  から、再更新されたパラメータを許容するか否かの判定に用いる許容確率 を算出するための許容確率算出部 266 と、所定の乱数発生方法により 0 から 1 までの閉区間内の乱数  $u$  を発生させるための乱数発生部 268 と、許容確率 と乱数  $u$  とに基づき、時刻  $t$  に対応するパーティクルにおけるパラメータとして、再サンプリングされたパーティクルにおけるパラメータと、再更新されたパーティクルにおけるパラメータとの一方を選択するためのパラメータ選択部 270 とを含む。

10

【0066】

許容確率算出部 266 は、重み  $w_t^{(j)}$  及び重み  $w_t^{*(j)}$  から次の式 (25) にしたがって、許容確率 を算出する機能を持つ。

【0067】

【数 17】

$$\nu = \min \left\{ 1, w_t^{*(j)} / w_t^{(j)} \right\} \quad (25)$$

20

【0068】

パラメータ選択部 270 は、 $u$  が許容確率 以下であれば、当該パーティクルにおける雑音のパラメータを再更新で得られた新たなパラメータに変更する機能を持つ。

【0069】

[プログラム構造]

以下の説明からも明らかなように、図 1 に示す音声認識システム 100 の前処理部 104、音響モデル部 106、及び探索部 110 は、いずれもコンピュータハードウェアとその上で実行されるプログラムにより実現可能である。図 7 に、本実施の形態に係る前処理部 104 に含まれる雑音抑圧部 114 が行なう雑音抑圧処理を実現するコンピュータプログラムの制御構造をフローチャートで示す。

30

【0070】

図 7 を参照して、雑音抑圧処理が開始されると、ステップ 302 において、雑音 122 の初期分布を推定する。すなわち、上記の式 (10) 及び式 (11) に示す算出方法により、式 (9) に示す初期分布  $p(N_0)$  のパラメータ平均ベクトル  $\mu_N$  及び共分散行列  $\Sigma_N$  を算出する。さらに、初期分布  $p(N_0)$  からパラメータベクトル  $N_0^{(j)}$  ( $j = 1, \dots, J$ ) をサンプリングし、各パーティクルにおける雑音の初期的なパラメータに設定する。

【0071】

ステップ 304 では、雑音抑圧の対象となるフレームを次のフレームに移行させる。ステップ 306 では、パーティクルフィルタを用いて、処理対象のフレームにおける雑音 122 の確率分布を推定する。すなわち、各パーティクルにおける雑音のパラメータベクトル  $N_t^{(j)}$ 、及びベクトル  $N_t^{(j)}$  の共分散行列  $\Sigma_{N_t^{(j)}}$  を推定し、さらに、各パーティクルにおけるクリーン音声 HMM 130 の状態  $s_t^{(j)}$ 、並びに各パーティクルに対する重み  $w^{(j)}$  を定める。このステップでの処理については、図 8 を用いて後述する。

40

【0072】

ステップ 308 では、ステップ 306 でパーティクルごとに定めたクリーン音声 HMM 130 の状態  $s_t^{(j)}$  と、雑音のパラメータベクトル  $N_t^{(j)}$  と、共分散行列  $\Sigma_{N_t^{(j)}}$  とを用いて、各パーティクルにおける雑音重畳音声の特徴量ベクトル  $X_t$  (124) の確率分布を推定する。さらに、クリーン音声 HMM 130 を構成する分布  $k(1 \dots k \dots K)$  ごとに、パーティクルにおける雑音重畳音声の確率モデルの平均ベクトル  $\mu_{x_{kt}^{(j)}}$ 、 $t$  と、共分散

50

行列  $\mathbf{x}_{k,t}^{(j)}$  とを算出する。

【 0 0 7 3 】

ステップ 3 1 0 では、M M S E 推定法により、時刻  $t$  のフレームにおけるクリーン音声の特徴量を推定する。すなわちまず、ステップ 3 0 6 及びステップ 3 0 8 の処理で得られたパラメータを用いて、M M S E 推定法によって、各パーティクルにおけるクリーン音声 1 2 0 の M M S E 推定値ベクトル  $\hat{\mathbf{S}}_t^{(j)}$  を算出する。M M S E 推定値ベクトル  $\hat{\mathbf{S}}_t^{(j)}$  は、次の数式により得られる。

【 0 0 7 4 】

【 数 1 8 】

$$\hat{\mathbf{S}}_t^{(j)} = \mathbf{X}_t - \sum_{k=1}^K P(k|\mathbf{X}_t, (j)) \log \left( \mathbf{I} + \exp \left( \mathbf{N}_t^{(j)} - \mu_{\mathbf{S}_t^{(j)}, k} \right) \right)$$

10

この式において、 $P(k|\mathbf{X}_t, (j))$  は、 $j$  番目のパーティクルにおける、クリーン音声 H M M 1 3 0 内の分布  $k$  に対する混合重みを表す。混合重み  $P(k|\mathbf{X}_t, (j))$  は、次の数式により算出される。

【 0 0 7 5 】

【 数 1 9 】

$$P(k|\mathbf{X}_t, (j)) = \frac{P_{\mathbf{S}_t^{(j)}, k} \mathcal{N} \left( \mathbf{X}_t, \mu_{\mathbf{X}_{k,t}^{(j)}}, \Sigma_{\mathbf{X}_{k,t}^{(j)}} \right)}{\sum_{k'=1}^K P_{\mathbf{S}_t^{(j)}, k'} \mathcal{N} \left( \mathbf{X}_t, \mu_{\mathbf{X}_{k',t}^{(j)}}, \Sigma_{\mathbf{X}_{k',t}^{(j)}} \right)}$$

20

【 0 0 7 6 】

続いて、M M S E 推定値ベクトル  $\hat{\mathbf{S}}_t^{(j)}$  と重み  $w_t^{(j)}$  とを用いて、時刻  $t$  におけるクリーン音声の推定特徴量ベクトル  $\hat{\mathbf{S}}_t$  を、次の数式に示す重み付き和によって算出し、推定クリーン音声の特徴量 1 2 6 ( 図 1 参照 ) として出力する。

【 0 0 7 7 】

【 数 2 0 】

$$\hat{\mathbf{S}}_t = \sum_{j=1}^J w_t^{(j)} \hat{\mathbf{S}}_t^{(j)}$$

30

【 0 0 7 8 】

ステップ 3 1 2 では、終了判定を行なう。すなわち時刻  $t$  のフレームが最終のフレームであれば雑音抑圧処理を終了する。さもなければステップ 3 0 4 に戻る。

【 0 0 7 9 】

図 8 に、ステップ 3 0 6 ( 図 7 参照 ) において行なわれる雑音確率分布の推定処理を実現するプログラムの制御構造をフローチャートで示す。図 8 を参照して、雑音確率分布の推定処理が開始されると、ステップ 3 2 2 において、式 ( 1 4 ) ~ 式 ( 1 9 ) により示す拡張カルマンフィルタを用いて、時刻  $t - 1$  のパーティクルにおける雑音の確率分布から、時刻  $t$  のパーティクルにおける雑音の確率分布を推定する。

40

【 0 0 8 0 】

ステップ 3 2 4 では、時刻  $t$  の各パーティクルに対する重み  $w_t^{(j)}$  を、式 ( 2 3 ) 及び式 ( 2 4 ) によって算出し、正規化する。ステップ 3 2 6 では、各パーティクルに与えられた重み  $w_t^{(j)}$  に基づき、各パーティクルからの再サンプリングの回数を決し、当該パーティクルにおける雑音の確率分布に基づいてパラメータを再サンプリングする。ステップ 3 2 8 では、Metropolis-Hastings アルゴリズムを用いて時刻  $t$  におけるパーティクルを再生成する。

【 0 0 8 1 】

図 9 にステップ 3 2 8 ( 図 8 参照 ) における処理の詳細をフローチャートで示す。図 9 を参照して、ステップ 3 2 8 における処理が開始されると、ステップ 3 4 2 において、ス

50

ステップ 3 2 6 ( 図 8 参照 ) での再サンプリングで得られたパーティクルにおけるパラメータを用いて、雑音の確率分布の再更新を行なう。すなわち、時刻  $t$  のフレームのパーティクルを新たに準備し、ステップ 3 2 2 ( 図 8 参照 ) での処理と同様の処理により、時刻  $t - 1$  のフレームのパーティクルに対応するパラメータから、時刻  $t$  のフレームのパーティクルに対応するパラメータへの再更新を行ない、準備したパーティクルのパラメータに設定する。ステップ 3 4 4 では、ステップ 3 4 2 で準備したパーティクルに対する重み  $w_t^{*(j)}$  (  $j$  ) を、図 8 に示すステップ 3 2 4 の処理と同様の処理で算出し正規化する。

#### 【 0 0 8 2 】

ステップ 3 4 6 では、ステップ 3 2 4 の処理で算出された重み  $w_t^{(j)}$  と、ステップ 3 4 4 で算出された重み  $w_t^{*(j)}$  との比較により、ステップ 3 4 2 で準備されたパーティクルの許容確率  $\alpha$  を定める。ステップ 3 4 8 では、区間  $[ 0 , 1 ]$  の値からなる一様な集合  $U_{[0,1]}$  の中から任意の値を選択することにより乱数  $u$  を発生する。ステップ 3 5 0 では、ステップ 3 4 8 で発生した乱数  $u$  の値と、ステップ 3 4 6 で定めた許容確率  $\alpha$  の値とを比較する。  $u$  が許容確率の値以下であれば、ステップ 3 5 2 へ進む。さもなければステップ 3 5 4 に進む。ステップ 3 5 2 では、ステップ 3 4 2 で準備されたパーティクルを許容する。すなわち、ステップ 3 2 6 での再サンプリングで得られたパラメータを、準備されたパーティクルのパラメータで置換して処理を終了する。ステップ 3 5 4 では、ステップ 3 4 2 で準備されたパーティクルを棄却する。すなわち、準備されたパーティクル及びそのパラメータを破棄し、処理を終了する。

#### 【 0 0 8 3 】

##### [ 動作 ]

本実施の形態に係る音声認識システム 1 0 0 は以下のように動作する。まず、図 5 に示す雑音確率分布推定部 2 0 0 による雑音の初期分布の推定動作を説明する。図 1 に示す計測部 1 1 2 が、音源 1 0 2 から雑音重畳音声 ( 1 2 0 及び 1 2 2 ) を受け、特徴量  $X_t$  ( 1 2 4 ) を抽出する。抽出された特徴量  $X_t$  ( 1 2 4 ) は、雑音抑圧部 1 1 4 の図 4 に示す雑音確率分布推定部 2 0 0 に与えられる。図 5 を参照して、雑音確率分布推定部 2 0 0 のフレーム選択部 2 2 0 は、特徴量  $X_t$  ( 1 2 4 ) のうち最初の 1 0 フレーム分を、雑音初期分布推定部 2 2 2 に与える。雑音初期分布推定部 2 2 2 は、上記の式 ( 9 ) ~ 式 ( 1 1 ) に示す処理により雑音の初期分布  $p ( N_0 )$  を推定する。さらに、雑音の初期分布  $p ( N_0 )$  から、上記の式 ( 1 2 ) 及び式 ( 1 3 ) に示すサンプリングを  $J$  回行なう。このサンプリングによって、各パーティクルにおける雑音の初期的なパラメータベクトル  $N_0^{(j)}$  及び共分散行列  $\Sigma_0^{(j)}$  が決定される。雑音確率分布推定部 2 0 0 は、これらのパラメータを、時刻  $t = 0$  における推定雑音分布 2 0 6 のパラメータとして出力する。

#### 【 0 0 8 4 】

次に、雑音確率分布推定部 2 0 0 による、時刻  $t$  (  $t - 1$  ) のフレームにおける推定雑音分布 2 0 6 の推定動作を説明する。図 5 を参照して、次のフレームの処理の開始要求 2 1 0 に応答して、フレーム選択部 2 2 0 は、雑音重畳音声の特徴量  $X_t$  ( 1 2 4 ) を更新部 2 3 0 に与えるとともに、クリーン音声 HMM サンプリング部 2 2 6 に、時刻  $t$  におけるクリーン音声 HMM の出力パラメータのサンプリングを要求する。更新部 2 3 0 は、これに応答して、時刻  $t - 1$  の各パーティクルにおける推定確率分布のパラメータ 2 0 6 を取得する。

#### 【 0 0 8 5 】

クリーン音声 HMM サンプリング部 2 2 6 は、クリーン音声 HMM 1 3 0 から、出力パラメータベクトル  $S^{(j)}_{st}, k_t^{(j)}$  のサンプリングを行なう。図 1 0 に、出力パラメータベクトル  $S^{(j)}_{st}, k_t^{(j)}$  のサンプリングの概要を模式的に示す。例えば、  $j$  番目のパーティクルにおいて、時刻  $t - 1$  に対応する状態  $s_{t-1}^{(j)}$  から、遷移確率にしたがって状態遷移を行なうことにより時刻  $t$  における  $j$  番目のパーティクルに関する状態 4 0 0 を抽出する。次にクリーン音声 HMM サンプリング部 2 2 6 は、状態  $s_t^{(j)}$  の混合正規分布の中から、混合重みにしたがった確率で単一正規分布  $k_t^{(j)}$  ( 4 0 2 ) をサンプリングする。クリーン音声 HMM サンプリング部 2 2 6 はさらに、分布  $k_t^{(j)}$  ( 4 0 2 ) により表さ

れる出力確率の分布にしたがい、出力パラメータベクトル  $S^{(j)}_{st}, k_t^{(j)}, t(404)$  をサンプリングする。クリーン音声HMMサンプリング部226は、総数Jの各パーティクルにおける出力パラメータベクトル  $S^{(j)}_{st}, k_t^{(j)}, t(404)$  をそれぞれ、以上の手順でサンプリングし、図5に示す更新部230に与える。

【0086】

図11に、逐次計算部224によるパラメータの更新、及び再サンプリングの概要を模式的に示す。図11においては、雑音のパラメータが左右方向に分布し、時間が上から下に進行する。また、図11においては、パーティクルを、白抜きの丸印、及び黒塗りの丸印によって模式的に示す。例えば、白抜きの丸印で示すパーティクルが重み  $w_t^{(j)}$  の値の微小なパーティクルであり、黒塗りの丸印で示すパーティクルが重み  $w_t^{(j)}$  の値の大きなパーティクルであるものとする。

10

【0087】

図11を参照して、時刻  $t-1$  に対応するパーティクルにより状態空間420が近似的に表現されているものとする。更新部230は、式(14)~式(19)により示す拡張カルマンフィルタを用いて、状態空間420内の各パーティクルにおける雑音のパラメータベクトル  $\hat{N}_{t-1}^{(j)}$  を、時刻  $t$  に対応する推定雑音分布のパラメータベクトル  $\hat{N}_t^{(j)}$  に更新する。これにより、状態空間420内の各パーティクルは更新され、パラメータが更新されたパーティクルにより時刻  $t$  に対応する状態空間430が表現される。

【0088】

続いて重み算出部232は、状態空間430内の各パーティクルに対する重み  $w_t^{(j)}$  を、式(23)及び式(24)によって算出する。再サンプリング部234は、重み  $w_t^{(j)}$  に基づき、パーティクルにおける雑音のパラメータを再サンプリングする。この際、再サンプリング部234はまず、状態空間430内の各パーティクルからの再サンプリングの回数を、 $w_t^{(j)}$  に応じてパーティクルごとに設定する。白抜きの丸印で表される重みの微小なパーティクルからのサンプリングの回数を0に設定する。また、黒塗りの丸印で表される重みの大きなパーティクルからのサンプリングの回数を、重みの大きさに応じて1~3に設定する。続いて、状態空間430内のパーティクルにおける雑音の確率分布に基づき、設定された回数ずつ、雑音のパラメータの再サンプリングを行なう。このようにして、時刻  $t$  に対応する新たな状態空間440を表現するパーティクルがそれぞれ形成される。

20

30

【0089】

再サンプリング部234によるこのような再サンプリングが繰返し行なわれると、あるフレームに対応するパーティクルの多くにおける雑音のパラメータが、それ以前の時点のフレームに対応する少数のパーティクルにおける雑音の確率分布からサンプリングされたものとなるおそれがある。そこで、推定雑音分布生成部236は、Metropolis-Hastingsアルゴリズムを用いて、新たに時刻  $t$  に対応するパーティクルにおけるパラメータを生成することにより、このような事態を防止する。図6に示す再更新部262は、状態空間440における推定雑音分布にしたがい、時刻  $t-1$  に対応する状態空間420内のパーティクルにおける雑音のパラメータを再更新する。重み再計算部264は、再更新されたパーティクルに対する重み  $w_t^{*(j)}$  を算出する。許容確率算出部266は、再更新されたパーティクルに対する重み  $w_t^{*(j)}$  と、再サンプリングされたパーティクルに対する重み  $w_t^{(j)}$  とを基に、許容確率  $\alpha$  を算出する。パラメータ選択部270は、許容確率  $\alpha$  と、乱数発生部268が発生した  $[0, 1]$  の区間の乱数  $u$  とを比較し、乱数  $u$  が許容確率  $\alpha$  以下であれば、再サンプリングされたパーティクルにおけるパラメータを、再更新されたパーティクルにおけるパラメータで置換する。さもなければ、再更新されたパーティクルにおけるパラメータを破棄する。

40

【0090】

以上のような動作をフレームごとに繰返すことにより、各フレームに対応して、各パーティクルにおける推定雑音分布206のパラメータベクトル  $N_t^{(j)}$  及び共分散行列  $\Sigma_t^{(j)}$  が推定される。雑音確率分布推定部200は、各パーティクルにおける推定雑音分布2

50

06のパラメータベクトル $N_t^{(j)}$ 及び共分散行列 $\Sigma_t^{(j)}$ と、各パーティクルに対する重み $w_t^{(j)}$ と、当該パーティクルにおけるクリーン音声HMM130の出力パラメータ及びそのサンプリング元である状態 $s_t^{(j)}$ と、雑音重畳音声の特徴量ベクトル $X_t$ とを、フレームごとに、図4に示すパラメータ生成部202に与える。

【0091】

図4を参照して、パラメータ生成部202は、VTS法によって、時刻tに対応する各パーティクルにおける雑音重畳音声の確率モデルの平均ベクトル及び共分散行列(208)を生成する。これにより、各パーティクルにおいて雑音の確率分布と、雑音重畳音声の確率分布とが推定されたことになる。クリーン音声推定部204は、MMSE推定法により、時刻tのフレームに対応する各パーティクルにおいて、クリーン音声のMMSE推定値ベクトル $\hat{S}_t^{(j)}$ を算出する。さらに、MMSE推定値ベクトル $\hat{S}_t^{(j)}$ と重み $w_t^{(j)}$ とを用いて、時刻tにおけるクリーン音声の推定特徴量ベクトル $\hat{S}_t$ を算出し、図1に示す探索部110に出力する。

10

【0092】

図1に示す探索部110は、クリーン音声の推定特徴量ベクトル $\hat{S}_t$ を用いて、クリーン音声HMM130と、言語モデル部108に保持された言語モデルとを基に、適合する目的言語の単語等を探索し、その結果を認識出力128として出力する。

【0093】

[実験]

本実施の形態に係る音声認識システム100による効果を確認するために、雑音重畳音声からの雑音の推定実験と、雑音重畳音声の認識実験とを行なった。以下、実験方法及び結果について説明する。

20

【0094】

本実験においては、日本語の雑音下音声認識評価用共通データベースに収録されたクリーン音声1001文のデータに雑音を人工的に加算して、雑音重畳音声を生成した。加算する雑音には、それぞれ実環境で収録された工場雑音と道路工事雑音とを使用した。本実験では、雑音を加算していないクリーン音声と、クリーン音声に20dBから0dBまでのSNR(Signal-to-Noise Ratio)で雑音を加算した雑音重畳音声とを用意し、試料とした。なお、クリーン音声の各文には、CMS(Cepstrum Mean Subtraction)処理を施した。その他、音声認識の際の音響分析条件は、次の表1に示す通りであった。

30

【0095】

【表1】

標本化周波数	8kHz(16bits)
高域強調	$1 - 0.97z^{-1}$
特徴量	13th order MFCC(with zero-th) + $\Delta$ + $\Delta\Delta$
フレーム長	25ms
フレームシフト長	10ms
分析窓	Hamming window

雑音認識実験では、比較のために、上記の各試料から、本実施の形態に係る雑音抑圧処理の方法を含む次の5種の処理方法で、探索に用いる特徴量を生成した。すなわち、(A)雑音抑圧処理を行わない雑音重畳音声の特徴量(ベースライン)、(B)ETSI(European Telecommunications Standards Institute)により勧告されているETSI Advanced front-end(ES202)による雑音抑圧処理を施した特徴量、(C)雑音の確率分布を更新せずに実行される従来のMMSE推定(従来型MMSE)により得られる推定特徴量、(D)パーティクルフィルタを用いた本実施の形態に係る雑音抑圧処理により得られる推定特徴量、及び(E)雑音重畳音声への加算前の雑音(真の雑音)を用いて実行されるMMSE推定により得られる推定特徴量である。

40

【0096】

パーティクルフィルタを用いた雑音抑圧処理において、クリーン音声HMMには、1状

50

態 5 1 2 混合モデル、4 状態 1 2 8 混合モデル、8 状態 6 4 混合モデル、及び 1 6 状態 3 2 混合モデルの 4 種を用いた。

【 0 0 9 7 】

図 1 2 及び図 1 3 にそれぞれ、工場雑音及び道路工事雑音の各フレームにおける、メル対数フィルタバンクの 1 番目のフィルタによる出力値と、当該出力値の移動平均と、雑音抑圧処理において当該出力値に対応する推定結果とを示す。これらグラフの横軸は、雑音のフレーム番号を表す。グラフ上、左から右に進むにしたがい時間が推移する。図 1 2 及び図 1 3 を参照して、本実施の形態に係る雑音抑圧処理における雑音の推定結果は、工場雑音及び道路工事雑音のいずれにおいても、雑音の移動平均に近い時間軌跡となっている。よって、本実施の形態に係る雑音抑圧処理において、時間変動が推定できていることが分かる。

10

【 0 0 9 8 】

また、この雑音推定実験における処理に、クロック周波数 3 . 2 ギガヘルツ、3 2 ビットの市販の CPU (Central Processing Unit) を用いた場合、処理に要した時間は、雑音重畳音声における実時間の 0 . 8 倍であった。すなわち、雑音推定処理を実時間で処理できることが明らかとなった。

【 0 0 9 9 】

表 2 に、工場雑音を加算して得られた雑音重畳音声に対する認識実験で得られた単語認識精度を、上記 ( A ) ~ ( E ) の処理方法別に示す。また、表 3 に、道路工事雑音を加算して得られた雑音重畳音声に対する認識実験で得られた単語認識精度を、上記 ( A ) ~ ( E ) の処理方法別に示す。なお、これらの表の「 ( D ) パーティクルフィルタ」の項目においては、雑音抑圧処理に用いたクリーン音声 H M M の状態数ごとに単語認識精度を示す。

20

【 0 1 0 0 】

【表 2】

表 2: 工場雑音環境下での単語正解精度 (%)

SNR	(A)ベースライン	(B)ETSI方式	(C)従来のMMSE	(D)パーティクルフィルタ				(E)真の雑音でのMMSE
				1状態	4状態	8状態	16状態	
20 dB	93.61	92.88	96.41	96.04	95.03	95.82	96.13	98.46
15 dB	81.12	86.86	88.92	89.59	87.38	88.82	90.02	94.87
10 dB	54.81	76.73	74.27	75.41	72.95	73.90	75.87	85.11
5 dB	29.47	53.18	50.94	51.98	51.30	50.54	54.50	63.13
0 dB	18.73	23.15	24.72	26.19	28.55	26.50	28.92	37.06
平均	55.55	66.56	67.05	67.84	67.04	67.12	69.09	75.73

30

【 0 1 0 1 】

【表 3】

表 3: 道路工事雑音環境下での単語正解精度 (%)

SNR	(A)ベースライン	(B)ETSI方式	(C)従来のMMSE	(D)パーティクルフィルタ				(E)真の雑音でのMMSE
				1状態	4状態	8状態	16状態	
20 dB	96.68	96.90	99.20	98.34	97.85	98.46	98.34	99.29
15 dB	89.93	94.81	97.61	97.08	95.21	96.07	95.61	98.16
10 dB	70.28	89.81	91.77	92.39	88.79	90.33	89.84	94.87
5 dB	38.81	76.02	71.57	73.04	72.24	74.21	75.28	82.69
0 dB	22.29	48.48	43.60	43.44	48.88	46.88	49.43	54.51
平均	63.60	81.20	80.75	80.86	80.59	81.19	81.70	85.90

40

【 0 1 0 2 】

表 2 及び表 3 を参照して、( D ) パーティクルフィルタによる雑音抑圧処理を行なうことで、良好な単語認識精度が得られることが分かる。特に、S N R の小さな雑音重畳音声すなわち雑音成分の割合が大きな雑音重畳音声においては、( D ) パーティクルフィルタによる雑音抑圧処理により、高い単語認識精度が得られることが分かる。

50

## 【 0 1 0 3 】

以上の実験結果から、本実施の形態の雑音抑圧処理により、雑音下での音声認識性能が改善され、かつ実時間処理が可能になることが明らかとなった。

## 【 0 1 0 4 】

[ 変形例等 ]

なお、本実施の形態においては、パーティクルフィルタによる処理を雑音抑圧に用いている。そのため、雑音抑圧後の推定クリーン音声のパラメータを用いて探索を行なう前に、さらに音響モデル適応を行なうこともできる。音響モデル適応により、推定クリーン音声に適合した音響モデルを探索に用いることができるようになる。したがって、認識精度が向上することが期待される。

10

## 【 0 1 0 5 】

今回開示された実施の形態は単に例示であって、本発明が上記した実施の形態のみに制限されるわけではない。本発明の範囲は、発明の詳細な説明の記載を参酌した上で、特許請求の範囲の各請求項によって示され、そこに記載された文言と均等の意味及び範囲内でのすべての変更を含む。

## 【 図面の簡単な説明 】

## 【 0 1 0 6 】

【 図 1 】本発明の一実施の形態に係る音声認識システム 1 0 0 の構成を示す概略図である。

【 図 2 】クリーン音声 H M M 1 3 0 の概念を示す概略図である。

20

【 図 3 】雑音重畳音声の状態空間モデル 1 6 0 の概念を示す概略図である。

【 図 4 】雑音抑圧部 1 1 4 の構成を示すブロック図である。

【 図 5 】雑音確率分布推定部 2 0 0 の構成を示すブロック図である。

【 図 6 】推定雑音分布生成部 2 3 6 の構成を示すブロック図である。

【 図 7 】雑音抑圧処理の制御構造を示すフローチャートである。

【 図 8 】雑音確率分布推定処理の制御構造を示すフローチャートである。

【 図 9 】Metropolis-Hastingsアルゴリズムによるサンプリング処理の制御構造を示すフローチャートである。

【 図 1 0 】クリーン音声 H M M 1 3 0 からパラメータをサンプリングする動作の概要を示す図である。

30

【 図 1 1 】パーティクルフィルタによる処理の概要を示す図である。

【 図 1 2 】工場雑音のメルフィルタバンク出力値と当該雑音の推定値とのフレームごとの変化を示す図である。

【 図 1 3 】道路工事雑音のメルフィルタバンク出力値と当該雑音の推定値とのフレームごとの変化を示す図である。

## 【 符号の説明 】

## 【 0 1 0 7 】

- 1 0 0 音声認識システム
- 1 0 2 音源
- 1 0 4 前処理部
- 1 0 6 音響モデル部
- 1 0 8 言語モデル部
- 1 1 0 探索部
- 1 1 2 計測部
- 1 1 4 雑音抑圧部
- 1 1 6 話者
- 1 1 8 雑音源
- 1 2 0 クリーン音声
- 1 2 2 雑音
- 1 2 4 特徴量

40

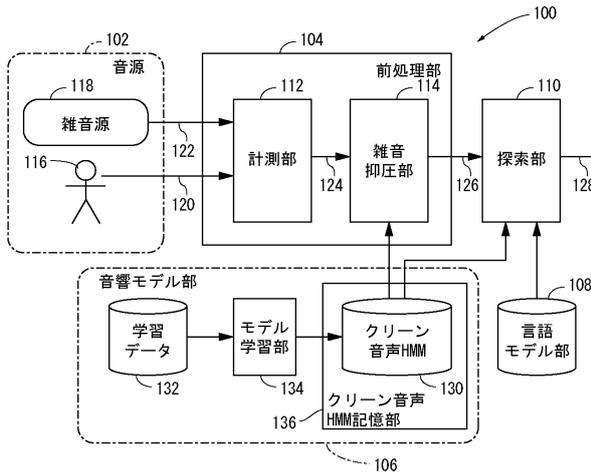
50

- 1 3 0 クリーン音声HMM
- 1 3 2 学習データ記憶部
- 1 3 4 モデル学習部
- 1 3 6 クリーン音声HMM記憶部
- 1 6 0 状態空間モデル
- 2 0 0 雑音確率分布推定部
- 2 0 2 パラメータ生成部
- 2 0 4 クリーン音声推定部
- 2 2 0 フレーム選択部
- 2 2 2 雑音初期分布推定部
- 2 2 4 逐次計算部
- 2 2 6 クリーン音声HMMサンプリング部
- 2 3 0 更新部
- 2 3 2 重み算出部
- 2 3 4 再サンプリング部
- 2 3 6 推定雑音分布生成部
- 2 6 2 再更新部
- 2 6 4 重み再計算部
- 2 6 6 許容確率算出部
- 2 6 8 乱数発生部
- 2 7 0 パラメータ選択部

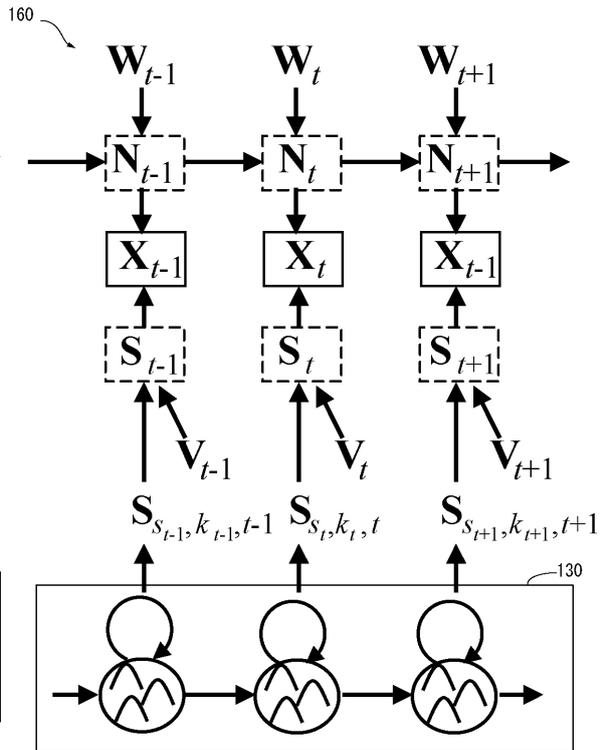
10

20

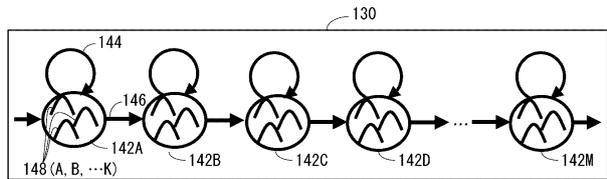
【図1】



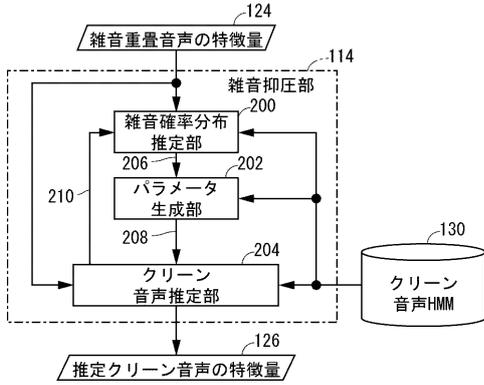
【図3】



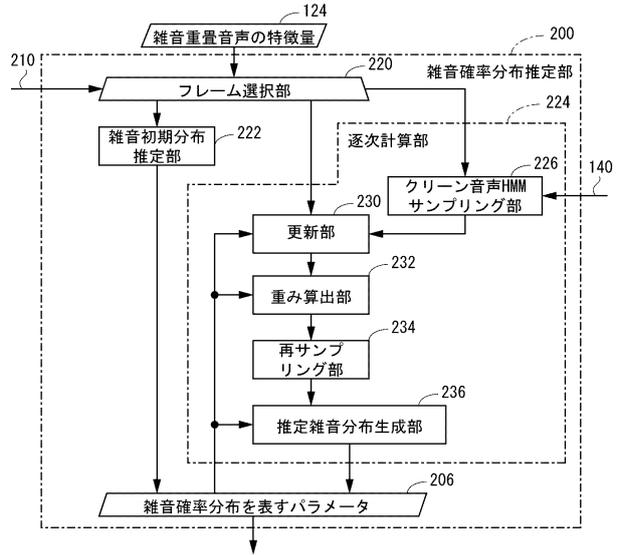
【図2】



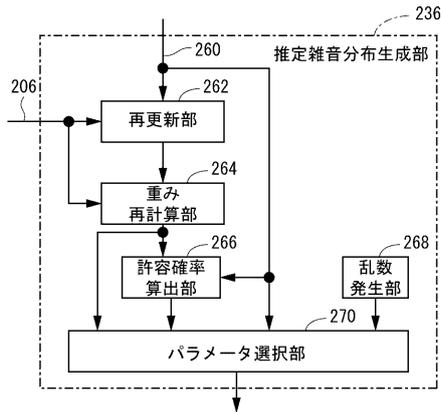
【図4】



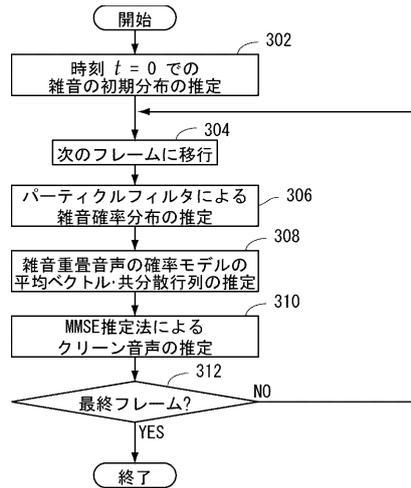
【図5】



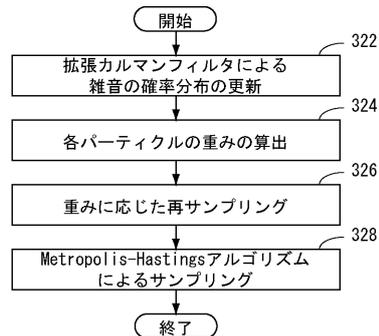
【図6】



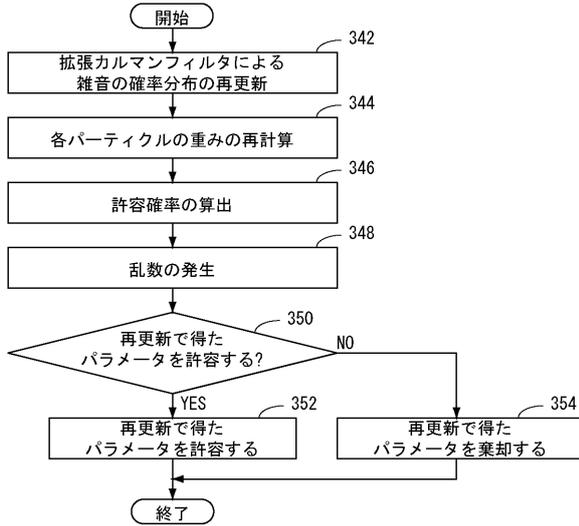
【図7】



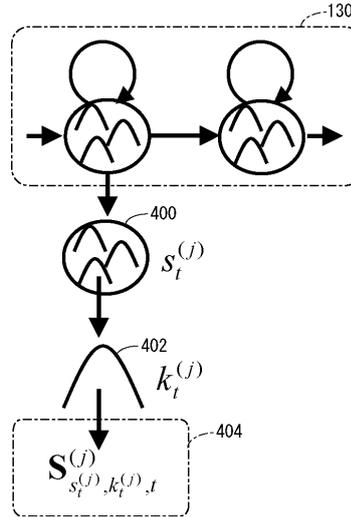
【図8】



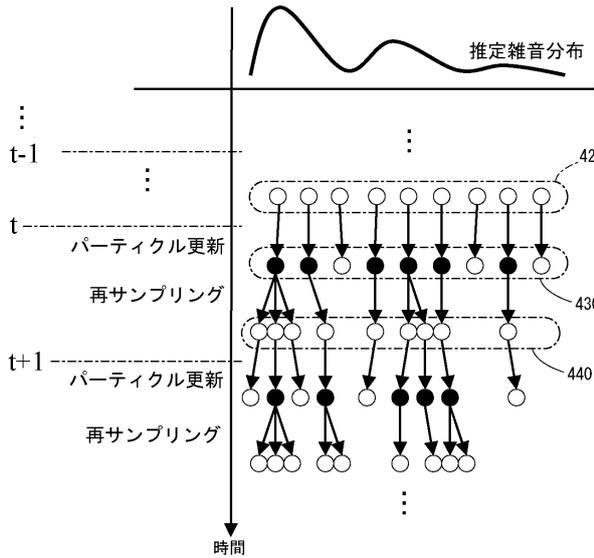
【図9】



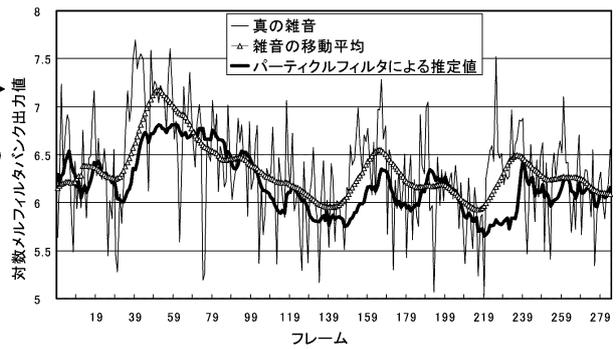
【図10】



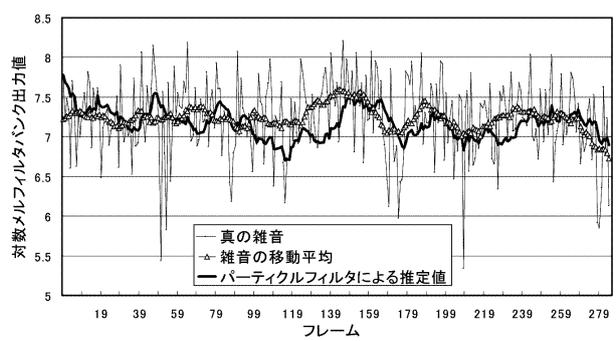
【図11】



【図12】



【図13】



## フロントページの続き

- (56)参考文献 特開2002-251198(JP,A)  
特開2005-084653(JP,A)  
特開2004-347761(JP,A)  
特開2003-216180(JP,A)  
佐宗 晃、浅野 太、他、HMMに基づいた雑音重畳音声の特徴量補正、日本音響学会2003年秋季研究発表会講演論文集、日本、日本音響学会、2003年 9月17日、I巻、pp.23-24  
Arulampalam, M.S et.al, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking", Signal Processing, IEEE Transactions on, 2002年 2月, Vol. 50, No.2, pp.174-188  
藤本ら, 「パーティクルフィルタに基づく雑音抑圧法を用いた非定常雑音及び残響環境下での音声認識」, 日本音響学会2005年春季研究発表会講演論文集 - I -, 日本, 社団法人日本音響学会, 2005年 3月, pp.113-114

## (58)調査した分野(Int.Cl., DB名)

G10L 11/00-21/06  
IEEE Xplore  
CiNii