(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3615008号
(P3615008)
(45)発行日　平成17年1月26日(2005.1.26)　(24)登録日　平成16年11月12日(2004.11.12)

(51)Int.Cl.[7]	F I		
G06T 1/00	G06T	1/00	340B
G06F 3/033	G06F	3/033	310Y
G06T 7/20	G06T	7/20	300A
G09B 21/00	G09B	21/00	F

請求項の数 7　(全 18 頁)

(21)出願番号　特願平9-11103
(22)出願日　平成9年1月24日(1997.1.24)
(65)公開番号　特開平10-208023
(43)公開日　平成10年8月7日(1998.8.7)
審査請求日　平成13年11月21日(2001.11.21)

(73)特許権者　000005108
株式会社日立製作所
東京都千代田区丸の内一丁目６番６号
(74)復代理人　100102587
弁理士　渡邊　昌幸
(74)代理人　100077274
弁理士　磯村　雅俊
(72)発明者　佐川　浩彦
東京都国分寺市東恋ケ窪一丁目２８０番地
株式会社日立製作所中央研究所内
(72)発明者　竹内　勝
東京都国分寺市東恋ケ窪一丁目２８０番地
株式会社日立製作所中央研究所内

最終頁に続く

(54)【発明の名称】手話認識装置

(57)【特許請求の範囲】
【請求項１】
手の形状や動きを電気信号に変換して、時系列手話データとして入力する手話入力手段と、
該手話入力手段から入力された時系列手話データから、動作の基本単位のうち動的な動作の基本単位、および複数の静的な動作の基本単位の時系列中に含まれる静的な動作の基本単位を認識する動的動作要素認識手段と、
該手話入力手段から入力された時系列手話データから、動作の基本単位のうち静的な動作の基本単位を認識する静的動作要素認識手段と、
上記動的動作要素認識手段から取り込まれた一つ以上の動的な動作の基本単位、あるいは複数の静的な動作の基本単位の時系列によって構成される逐次的な動作単位を認識する動的逐次要素認識手段と、
上記静的動作要素認識手段から取り込まれた一つ以上の静的な動作の基本単位によって構成される逐次的な動作単位を認識する静的逐次要素認識手段と、
上記動的逐次要素認識手段から取り込まれた動的な動作の基本単位によって構成される逐次的な動作単位と、上記静的動作要素認識手段から取り込まれた静的な動作の基本単位とを、動的な動作の基本単位によって構成される逐次的な動作単位の時間範囲に基づいて統合する静的動作要素統合手段と、
上記静的動作要素統合手段および上記静的動作要素認識手段から取り込まれた逐次要素の認識結果に対して、手話としての動作を認識する手話形態素認識手段と、

動作の基本単位の組み合わせによって表現した手話テンプレートを記憶し、上記各手段に参照される手話形態素辞書手段と、
上記手話形態素認識手段により認識した手話を、音声あるいは文字の形で出力する手段とを有することを特徴とする手話認識装置。
【請求項２】
請求項１に記載の手話認識装置において、
前記手話テンプレートを記憶する手話形態素辞書手段に記憶される動作の基本単位は、動作の種類を表す記号と連続量で表した動作の属性値の組み合わせによって表現されることを特徴とする手話認識装置。
【請求項３】
請求項２に記載の手話認識装置において、
前記動的な動作の基本単位によって構成される逐次的な動作単位と、静的な動作の基本単位を統合する静的動作要素統合手段は、動的な動作の基本単位によって構成される逐次的な動作単位によって決定された時間範囲における静的な動作の基本単位の認識結果のみを選択することを特徴とする手話認識装置。
【請求項４】
請求項３に記載の手話認識装置において、
前記動的な動作の基本単位によって構成される逐次的な動作単位と、静的な動作の基本単位を統合する静的動作要素統合手段は、まず動的な動作の基本単位によって構成される逐次的な動作単位によって決定された時間範囲を、静的な動作の基本単位を認識する静的動作要素認識手段に送り、該静的動作要素認識手段は送られた時間範囲のみについて認識処理を行い、その結果を静的動作要素統合手段に送ることを特徴とする手話認識装置。
【請求項５】
請求項１に記載の手話認識装置において、
前記動的な動作の基本単位によって構成される逐次的な動作単位を認識する動的逐次要素認識手段は、認識された動的な動作の基本単位の時間範囲の重なり度に基づいて認識を行うことを特徴とする手話認識装置。
【請求項６】
請求項１に記載の手話認識装置において、
前記静的な動作の基本単位の認識を行う静的動作要素認識手段は、時系列手話データの各時刻におけるデータを、手話テンプレートを記憶する手話形態素辞書手段に記憶されている静的な動作の基本単位の属性値に基づいて評価することを特徴とする手話認識装置。
【請求項７】
請求項１に記載の手話認識装置において、
前記静的な動作の基本単位によって構成される逐次的な動作の単位を認識する静的逐次要素認識手段は、各時刻における静的な動作の基本単位の評価結果を逐次的な動作の単位毎にまとめることにより、逐次的な動作の単位に対する評価値を求め、求めた評価値が極大値となる時間範囲を逐次的な動作の単位の認識結果とすることを特徴とする手話認識装置。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】
本発明は、手話を入力して、その結果を音声または文字の形で正確に出力することにより、聴覚障害者と健聴者とのコミュニケーションを支援する手話認識装置に関するものである。
【０００２】
【従来の技術】
従来より、手話を入力して、その入力結果を解析しその手話を認識する装置が種々提案されている。これら従来の認識方法としては、手話の動作データをパターン照合の技術（例えば、特願平５－１２５６９８号明細書および図面参照）や、ニューラルネットワークの

技術を用いて認識を行う技術と、手話動作を構成する動作の基本単位に基づいて手話の認識を行う技術（例えば、特願平６－２５３４５７号明細書および図面参照）がある。
前者（特願平５－１２５６９８号）の手法では、全体としての手動作パターンを単語辞書に格納された標準手動作パターンと比較して、一致するか否かにより認識する方法であり、後者（特願平６－２５３４５７号）の手法では、手動作パターン自体を比較するのではなく、手動作パターンを先ず動作の基本単位毎（部分パターン）に認識し、次にその動作の基本単位の結果を統合して単語を認識するものである。このように、後者の技術では、先ず認識を行う手話を構成する動作の基本単位を全て認識する。その結果と、あらかじめ手話を動作の基本単位の組み合わせとして記憶してある手話テンプレート中の動作の基本単位の時間的な関係とを比較することにより認識を行うのである。この場合、手話テンプレート中の動作の基本単位の種類や属性は全て記号によって記述されていた。

【０００３】

【発明が解決しようとする課題】

従来における動作の基本単位に基づく手話認識技術（前述した後者の技術）では、静的な動作の基本単位も動的な動作の基本単位も同時に認識を行い、認識された動作の基本単位を統合することにより手話を認識していた。このうち、静的な動作の基本単位とは、形状（手の形等）や方向（指の方向等）など、ある時間範囲においてパラメータが安定した状態を示す動作の特徴である。一方、動的な動作の基本単位は、直線運動や円運動などある時間範囲におけるパラメータの変化（移動等）を表す動作の特徴である。
静的な動作の基本単位を検出するためには、基準となるパラメータ（サンプルをとって、そこから抽出したパラメータ）と入力されたパラメータの差に対して閾値を設定し、その閾値以内の区間を検出することによって行えばよい。
しかしながら、静的な動作の基本単位のパラメータは変化しやすいため、認識したい時間範囲より大きい範囲で認識されたり、小さい範囲で検出されたりすることが多い。例えば、毎回同じような形で、かつ入力されたパラメータが閾値の境界のところで変動すると、閾値を往復してしまう結果、小間切れの区間として検出されてしまうことがある。逆に大きい区間として検出する場合も生じる。また、パラメータの変動の状態によっては、一つの動作の区間が二つ以上に分割されて認識される場合もある。
このため、静的な動作の基本単位と動的な基本単位を対等に扱っていた従来技術では、静的な動作の基本単位については正しい評価が行えないという問題があった。
また、従来の技術では、手話テンプレートを全て動作の基本単位を表す記号によって記述していたため、動作の基本単位の認識ではあらかじめ設定してある基準値に基づき認識を行い、その評価値も基準に基づいて計算を行っていた。
さらに、手話の認識結果も認識された動作の基本単位の評価値（基準値にどの程度近いかを表わす値で、近い程高い値）に基づいて計算を行っていた。このため、実際の動作におけるパラメータの範囲とあらかじめ決定していた基準値によって与えられるパラメータの範囲のずれにより、正しい評価値が得られないことが多く、認識精度が低いという問題があった。すなわち、パラメータの集中する場所は、基準値に近いとは限らず、基準値の中間位置に集中することがある。このような場合には、入力されたパラメータと基準値との差は大きくなり、評価値としては小さくなってしまう。
そこで、本発明の目的は、このような従来の課題を解決し、全ての動作の基本単位を正しく評価し、精度良く手話を認識する手話認識装置を提供することである。

【０００４】

【課題を解決するための手段】

上記目的を達成するため、本発明の手話認識装置では、動的な動作の基本単位を検出し、それらにより構成される逐次的な動作の単位を認識した後、認識した逐次的な動作の単位の時間範囲内において静的な動作の基本単位の評価結果を統合する。静的な動作の基本単位のみによって構成される逐次的な動作の単位は、各時刻における動作全体の評価値を求め、評価値が極大になる時刻を求めることにより認識を行う。また、動作の基本単位を正しく評価するために、手話テンプレートに記述される動作の基本単位は、実際の手話デー

タを用いて決定した属性値を連続量を用いて記述する。認識された動作の基本単位の評価値は、その連続量を用いて求める。
このように、動的な動作の基本単位の認識結果から決定される時間範囲に基づいて静的な動作の基本単位を評価するため、静的な動作の基本単位の認識範囲に関する問題がなくなる。
また、静的な動作の基本単位のみによって構成される逐次的な動作の単位についても、各時刻毎に逐次的な動作の単位全体の評価値を求め、評価値の極大値となる時刻を検出することにより認識を行うため、それぞれの動作の基本単位の検出時間に影響を受けることがなくなる。
さらに、動作の基本単位の特徴を表す属性値を実際の動作データから求めた連続量によって表現し、それに基づいて認識された動作の基本単位を評価するため、動作の基本単位の適切な評価値を求めることが可能となり、手話の認識精度を向上することができる。
【０００５】
【発明の実施の形態】
以下、本発明の実施例を、図面により詳細に説明する。
図２は、本発明における動作要素に基づく手話動作モデルを示す図である。
本発明を説明するために、まず手話動作のモデルを説明する。手話動作は、動作の基本単位の組み合わせにより構成される。この手話における動作の基本単位を、以後動作要素と呼ぶことにする。手話の動作を構成する動作要素間には時間的な逐次性および同時性があるため、手話の認識を行うためには、それら動作要素間の時間的な関係も記述しておく必要がある。このために、図２に示すモデルを使用する。図２において、２０１は手話形態素の動作全体を表す。手話形態素は、手話における意味の単位である。手話形態素は、まず逐次的な動作単位である逐次要素２０２，２０３，２０４に分解される。横軸は左端を基準として時間の経過を表わしているので、逐次要素２０２，２０３，２０４の順序で動作が発生したことになる。逐次要素は必ず連続的に表現され、同時に表現されることはない動作の単位である。逐次要素は、さらに複数の同じ時間範囲に表現される単位である同時要素２０５，２０６，２０７に分解される。同時要素には、動作要素２０８，２０９が含まれる。同時要素には二種類あり、一つの動作要素のみから構成される同時要素と、二つ以上の動作要素のみから構成される同時要素がある。
二つ以上の動作要素から構成される同時要素は、その中に含まれる動作要素が逐次的に表現された場合に存在するとみなされる。ただし、一つの同時要素に含まれる動作要素は、必ず同じ種類の動作要素であるとする。なお動作要素とは、方向が同じで手の形態だけ異なる等の動作の要素である。
手話動作は、このように動作要素の逐次的構造と同時的構造の組み合わせによって構成される。
【０００６】
図１は、本発明の一実施例を示す手話認識装置の概念ブロック図である。
図１において、手話入力部１０１（データグローブ）は手話における動作を電気信号に変換し、時系列データとして動的動作要素認識部１０２および静的動作要素認識部１０３に入力する。動的動作要素認識部１０２では、動作データから動的な動作要素を認識する。静的動作要素認識部１０３では、動作データの各時刻のデータに対して静的な動作要素の評価値を求める。動的逐次要素認識部１０４では、認識された動的な動作要素から構成される逐次要素を認識する。静的逐次要素認識部１０５では、静的な動作要素のみから構成される逐次要素の認識を行う。静的動作要素統合部１０６では、動的な動作要素から構成される逐次要素に静的な動作要素の認識結果を統合することにより、手話形態素の認識を行う。本発明は、この静的動作要素統合部１０６を設けたことに特徴がある。先にモデルの説明で述べたように、静的動作要素の認識範囲は基本単位のパラメータが変化し易いため、認識したい時間範囲より小さい範囲で検出されたり、大きい範囲で認識されたりするとともに、通常、静的動作要素は動的動作要素と一緒に現われるので、動的動作要素の逐次要素と静的動作要素の認識結果を統合し、評価するのである。つまり、動的動作要素は

、直線運動や円運動で時間が明確に決められるので、認識した時間範囲内において静的動作要素を評価してやればよい。なお、図２のモデルと図１の関係では、図２の動作要素２０８，２０９の中に動的動作要素と静的動作要素とが含まれると考えればよい。
手話形態素の認識結果は、出力部１０８によりモニタ１０９およびスピーカ１１０へ出力される。手話形態素辞書１１１には、手話における意味の単位である手話形態素毎に、動作要素の組み合わせによって記述した手話テンプレートが格納されている。
【０００７】
図３は、動的動作要素認識部の構造を示すブロック図であり、図４は、静的動作要素認識部の構造を示すブロック図である。
動的動作要素認識部１０２は、図３に示すように、独立したそれぞれの動作要素毎の認識部３０１，３０２，３０３から構成されている。各動作要素認識部には、それぞれの認識処理に必要な認識用パラメータ３０４，３０５，３０６が用意される。図３に示す動的動作要素認識部では、考えられる全ての動作要素を認識することになるが、手話形態素辞書中の動的動作要素のみを認識するようにしても良い。
静的動作要素認識部１０３も、図４に示すように、各動作要素毎の認識部４０１，４０２，４０３から構成されている。静的動作要素認識部１０３では、認識に必要なパラメータは全て手話形態素辞書に格納されているデータを用いる。
【０００８】
図５は、図１における手話認識装置を実現するためのハードウェアの一構成例を示す図である。
図５において、手話入力装置５０１は手話における手動作を電気信号に変換する装置であり、手袋にセンサを設置し、手の形状や動きを電気信号に変換する装置として良く知られている装置（データグローブ）を利用することができる。手話入力装置５０１によって、手話の手動作は指の曲げ角度や手の位置などからなる多次元の時系列データに変換される。演算装置５０２は、動作要素の認識や手話形態素の認識を行う装置であり、メモリ５０４，５０６，５０７，５０８，５０９，５１１からプログラムを読み込み、それらのプログラムに従って認識処理を行う。出力装置５０３は、手話形態素の認識結果を出力する装置であり、文字による出力や音声合成を用いた音声による出力装置を利用することができる。メモリ５０４は、動的動作要素を認識するためのプログラムを記憶するための記憶装置、メモリ５０５は動的動作要素を認識するために必要なパラメータを記憶するための記憶装置、メモリ５０６は動的な動作要素から構成される逐次要素を認識するためのプログラムを記憶するための記憶装置、メモリ５０７は静的動作要素を認識するために必要なプログラムを記憶するための記憶装置、メモリ５０８は動的な動作要素から構成される逐次要素と静的な動作要素の認識結果を統合するためのプログラムを記憶するための記憶装置、５０９は静的な動作要素のみで構成される逐次要素を認識するためのプログラムを記憶するための記憶装置、メモリ５１０は手話形態素の動作データである手話形態素辞書を記憶するための記憶装置、メモリ５１１は手話形態素を認識するためのプログラムを記憶するための記憶装置である。図５のハードウェア構成では、全ての認識プログラムの実行を一つの演算装置だけで行う構成であるが、この他に、複数の演算装置を用いて認識プログラムの実行をそれぞれの演算装置に分散させる構成も可能である。
【０００９】
図６は、図１の手話入力部により入力される動作データのフォーマット図である。
図６において、６０１は手の位置に関するデータであり、手の位置はさらにｘ軸のデータ６０２，ｙ軸のデータ６０３，ｚ軸のデータ６０４から構成されている。６０５は手の方向に関するデータであり、手の方向はさらにｘ軸回りの角度６０６，ｙ軸回りの角度６０７，ｚ軸回りの角度６０８から構成されている。
６０９は指の曲げに関するデータであり、指の曲げはさらに、親指の第２関節の曲げ角度６１０，親指の第３関節の曲げ角度６１１，人差し指の第１関節の曲げ角度６１２，人差し指の第２関節の曲げ角度６１３，中指の第１関節の曲げ角度６１４，中指の第２関節の曲げ角度６１５，薬指の第１関節の曲げ角度６１６，薬指の第２関節の曲げ角度６１７，

小指の第１関節の曲げ角度６１８，小指の第２関節の曲げ角度６１９から構成されている。また、６２０，６２１，６２２は、それぞれ時刻ｔ１，ｔ２，ｔｎにおける手の位置，方向，指の曲げのデータを表す。このように、手話における動作は、手の位置６０１，手の方向６０５，指の曲げ６０９からなる時系列データとして表される。

図７は、図５の動的動作要素認識用パラメータを格納するメモリ（５０５）に格納されるパラメータのフォーマット図である。

図７において、動作要素名７０１はそのパラメータを認識処理に使用する動作要素の名称，パラメータ数７０２はその動作要素の認識に使用するパラメータの数，７０３，７０４は各パラメータを表す。また、パラメータ種類７０５，７０７はそのパラメータの意味を表す名称，パラメータ７０６，７０８は実際に認識処理に利用されるパラメータの値を表す。

【００１０】

図８は、図５の手話形態素辞書メモリに格納されるフォーマット図である。

図８において、手話形態素名８０１は、それ以下に記述される動作要素の組み合わせが表す手話形態素の名称を表す。繰り返し回数８０２は、それ以下に記述される動作が繰り返される回数を表す。逐次要素数８０３は、手話動作を構成する逐次要素の数を表す。逐次要素間重なり度８０４は、それぞれの逐次要素が実際の手話動作中に表現された場合に生じる重なり、あるいはギャップに対する許容範囲を表す。すなわち、実際に認識される場合には、要素相互が重なってしまったり、あるいは要素と要素の間が空いてしまう場合があるので、その度合を登録しておく。この場合には、離れていたとき＋、重なっていたとき－となる。逐次要素間重なり度は、逐次要素数が２以上の場合に有効である。逐次要素８０５，８０６，８０７は、それぞれの逐次要素の記述を表す。同時要素数８０８は、逐次要素を構成する同時要素の数を表す。同時要素間重なり度８０９は、逐次要素を構成するそれぞれの同時要素が実際の手話動作中で表現された場合に生じる重なりに対する許容範囲を表す。同時要素間重なり度は、同時要素数が２以上の場合に有効である。繰り返し回数８１０，８１５は、それ以下に記述される動作要素の列が繰り返される回数を表す。動作要素状態数８１１，８１６は、それぞれの同時要素を構成する動作要素の数を表す。動作要素間重なり度８１２，８１７は、逐次的に表現される動作要素が実際の手話動作中で表現された場合に生じる重なりあるいはギャップに対する許容値を表す。動作要素間重なり度は、動作要素状態数が２以上の時に有効である。動作要素８１３，８１４，８１８，８１９は、それぞれの同時要素を構成する動作要素を表す。

【００１１】

図９は、動作要素の記述フォーマット図であり、図１０は、動作要素の種類およびそれぞれの属性値の種類を示す図であり、図１１は、動作要素の属性値のフォーマット図である。

図９において、９０１は動作要素の種類を、９０２はその動作要素を表現するために使用される手の部位を、９０３，９０４，９０５はその動作要素を表すために必要な属性値を表す。動作要素の種類は、図１０に示すように、１４種類の動作要素から選択する。また、図１０に示すように、動作要素の種類に応じて属性値の種類もあらかじめ決定されている。

図１１の属性値フォーマットにおいて、１１０１，１１０２，１１０３は複数の動作データから学習した属性値の平均値，１１０４，１１０５，１１０６は複数の動作データから学習した属性値の分散である。ここで属性値の平均値とは、サンプルをとって、それらの平均値ｐ１～ｐｎをとったものであり、また属性値の分散とは、平均値に対するばらつきであって、何回かとったパラメータがどのくらいばらついているかを同じデータから計算して、ｓ１～ｓｎとして表わしたものである。なお属性値の次元は、図１０に示した属性値の種類に応じてあらかじめ決定されている。

【００１２】

次に、本発明における認識処理について説明する。

図１２は、動作要素状態数が２以上の同時要素に含まれる静的動作要素の認識処理を示す

フローチャートである。
図１における動的動作要素認識部１０２では、振動や直線等の動的な動作要素および手話形態素辞書において、動作要素状態数が２以上の同時要素に含まれる形状や方法などの静的な動作要素の二種類の認識（動的動作要素および静的動作要素の認識）を行う。動的な動作要素認識の技術としては、既にある技術（例えば、特願平６－２５３４５７号明細書および図面『手話認識装置』参照）を使用することができる。動作要素状態数が２以上の同時要素に含まれる静的な動作要素の認識は、図１２に示すフローチャートに従って行うことができる。
図１２において、ステップ１２０１では、まず手話形態素辞書から動作要素状態数が２以上の同時要素に含まれる静的動作要素を抽出し、そのリストを作成する。リスト中の動作要素のフォーマットは、図９に示す動作要素のフォーマットと同じで良い。次に、ステップ１２０２において、手話入力部から１時刻分のデータを読み込む。ステップ１２０３において、動作データが最後であれば処理を終了する。最後でなければ、ステップ１２０４に移る。ステップ１２０４において、静的動作要素リストの全ての動作要素について、動作要素の属性値と読み込んだデータとからその時刻における評価値を求める。
【００１３】
評価値は、静的動作要素の属性値の種類をｎ，ｉ番目の属性値の次元数をｍ（ｉ），静的動作要素の手話形態素辞書に記述されているｉ番目の属性値の平均を（Ｐ（ｉ，１），Ｐ（ｉ，２），・・・，Ｐ（ｉ，ｍ（ｉ））），分散を（Ｓ（ｉ，１），Ｓ（ｉ，２），・・・，Ｓ（ｉ，ｍ（ｉ）），入力された時刻をｔ、入力されたデータのｉ番目の属性値を（Ｘ（ｔ，ｉ，１），Ｘ（ｔ，ｉ，２），・・・Ｘ（ｔ，ｉ，ｍ（ｉ））として、各時刻の評価値Ｅ１（ｔ）は下記（数１）の式によって求められる。
【数１】

$$E1(t) = \left(\prod_{i=1}^{n} \prod_{j=1}^{m(i)} e^{-\frac{(X(t,i,j)-P(i,j))^2}{2\times S(i,j)}} \right)^{\frac{1}{\sum_{i=1}^{n} m(i)}}$$

次に、ステップ１２０５において、各動作要素毎にそれまでに求めた評価値と新しく求めた評価値からなる評価値の時系列が極大になる時間範囲を求める。評価値の時系列が、例えば図１３に示す曲線１３０１のように求められた場合には、極大となる時刻１３０２から時刻１３０３の範囲が動作要素として検出される。
ステップ１２０６において、求めた時間範囲とそれに対応する動作要素，評価値を認識結果として出力する。ステップ１２０７において、求めた各動作要素毎の評価値を次の時刻での認識に使用するためにバッファに格納し、ステップ１２０２に戻る。
図１４は、動的動作要素および図１２のフローで検出された静的動作要素のフォーマット図である。
１４０１は動作要素の検出された時間範囲の開始時刻，１４０２は動作要素の検出された時間範囲の終了時刻，１４０３は検出された動作要素に対する評価値，１４０４は検出された動作要素の種類，１４０５はその動作要素を表現するために使用される手の部位，１４０６，１４０７は各動作要素に付属する属性値である。属性値は、動作要素が検出された範囲における動作データから求めた値である。
【００１４】
次に、図１に示す動的逐次要素認識部１０４において、動的な動作要素を含む逐次要素を認識する方法について説明する。
図１５は、動的逐次要素の認識処理のフローチャートである。
この処理では、大きく分けて同時要素の認識と逐次要素の認識の二段階の処理が行われる

。ステップ１５０１では、手話形態素辞書から動的な動作要素によって構成される同時要素と、動作要素状態が２以上の静的な動作要素によって構成される同時要素を抽出し、同時要素リストを作成する。この場合の同時要素のフォーマットを図１６に示す。図１６において、１６０１は手話形態素辞書中の同時要素の通し番号，１６０２は動作要素間の重なり度，１６０３は動作要素状態数，１６０４，１６０５は同時要素に含まれる各動作要素である。動作要素のフォーマットは図９に示すフォーマットと同じであり、また動作要素のフォーマット中の属性値は図１１のフォーマットと同じである。

ステップ１５０２では、手話形態素辞書から動的な動作要素によって構成される同時要素と、動作要素状態数が２以上の静的動作要素を含む同時要素を含む逐次要素を抽出し、動的な動作要素を含む同時要素および動作要素状態数が２以上の静的な動作要素を含む同時要素と逐次要素の対応リストを作成する。この同時要素と逐次要素の対応リスト中の逐次要素のフォーマットを図１７に示す。

図１７において、１７０１は手話形態素辞書中における逐次要素の通し番号，１７０２は同時要素間の重なり度，１７０３はその逐次要素に含まれる動的な動作要素含む同時要素および動作要素状態数が２以上の同時要素の数，１７０４，１７０５は動的な動作要素を含む同時要素および動作要素状態数が２以上の静的な動作要素を含む同時要素の手話形態素辞書中における通し番号である。

【００１５】

ステップ１５０３では、動的動作要素認識部１０２からの認識結果を一つ読み込む。次にステップ１５０４において、認識結果が最後であれば処理を終了する。そうでなければ、ステップ１５０５に進む。ステップ１５０５では、読み込んだ動作要素によって構成される同時要素の認識を行う。

この処理は、図１８に示すフローチャートに従って行われる。図１８において、ステップ１８０１では、ステップ１５０３で読み込んだ動作要素を含む同時要素を同時要素リストから検索する。ステップ１８０２において、検索された同時要素の数をカウンタｉに代入する。ステップ１８０３において、検索された同時要素のうちｉ番目の同時要素の動作要素状態数が１であればステップ１８０４に、そうでなければステップ１８０５に進む。ステップ１８０４では、ｉ番目の同時要素中の動作要素の属性値と読み込んだ動作要素の属性値とから評価値を求める。

この場合の評価値Ｅ２は、ｉ番目の同時要素の属性値の種類をｎ，ｉ番目の同時要素のｊ番目の属性値の次元数をｍ（ｊ），ｉ番目の同時要素のｊ番目の属性値を（Ｐ（ｊ，１），Ｐ（ｊ，２），…，Ｐ（ｊ，ｍ（ｊ））），分散を（Ｓ（ｊ，１），Ｓ（ｊ，２），…，Ｓ（ｊ，ｍ（ｊ））），読み込んだ動作要素のｊ番目の属性値を（Ｘ（ｊ，１），Ｘ（ｊ，２），…，Ｘ（ｊ，ｍ（ｊ））として、下記（数２）の式によって求めることができる。

【数２】

$$E2 = \left(\prod_{j=1}^{n} \prod_{k=1}^{m(j)} e^{-\frac{(X(j,k)-P(j,k))^2}{2\times S(j,k)}} \right)^{\frac{1}{\sum_{j=1}^{n} m(j)}}$$

【００１６】

次にステップ１８０５では、読み込んだ動作要素とバッファ中の動作要素からｉ番目の同時要素を構成する動作要素列と同じ動作要素列を検索する。ステップ１８０６では、検索した動作要素列中の動作要素の属性値とｉ番目の同時要素中の動作要素の属性値から同時要素の評価値を求める。この場合の評価値Ｅ３は、ｉ番目の同時要素の動作要素の数をｎ，ｉ番目の同時要素のｊ番目の動作要素の属性値の種類をｍ（ｊ），　ｉ番目の同時要素のｊ番目の動作要素のｋ番目の属性値の次元数をｑ（ｊ，ｋ），ｉ番目の同時要素のｊ番

目の動作要素のｋ番目の属性値を（Ｐ（ｊ，ｋ，１），Ｐ（ｊ，ｋ，２），…，Ｐ（ｊ，ｋ，ｑ（ｊ，ｋ））），分散を（Ｓ（ｊ，ｋ，１），Ｓ（ｊ，ｋ，２），…，Ｓ（ｊ，ｋ，ｑ（ｊ，ｋ））），　ｉ番目の同時要素のｊ番目の動作要素に対応する読み込んだ動作要素あるいはバッファ中の動作要素のｋ番目の属性値を（Ｘ（ｊ，ｋ，１），Ｘ（ｊ，ｋ，２），…，Ｘ（ｊ，ｋ，ｑ（ｊ，ｋ））），ｊ番目とｊ＋１番目の動作要素間の重なりあるいはギャップをＧ（ｊ）（重なりの場合正，ギャップの場合負），動作要素間の重なりあるいはギャップの平均値をＡ，分散をσとして，下記（数３）の式によって求めることができる。

【数３】

$$E3=\left(\prod_{j=1}^{n}\prod_{k=1}^{m(j)}\prod_{p=1}^{q(j,k)} e^{-\frac{(X(j,k,p)-P(j,k,p))^2}{2\times S(j,k,p)}}\times\prod_{j=1}^{n-1} e^{-\frac{(G(j)-A)^2}{2\times\sigma}}\right)^{\frac{1}{\sum_{j=1}^{n}\sum_{k=1}^{m(j)} q(j,k)+n-1}}$$

【００１７】

ステップ１８０７では、求めた評価値とｉ番目の同時要素を同時要素の認識結果として、認識された同時要素のバッファに格納する。バッファに格納される同時要素のフォーマットを図１９に示す。図１９において、１９０１は認識された同時要素の時間範囲の開始時刻，１９０２は認識された同時要素の時間範囲の終了時刻，１９０３は手話形態素辞書中における同時要素の通し番号，１９０４は同時要素に対する評価値である。開始時刻１９０１および終了時刻１９０２は、その同時要素が認識されたと判定されるもととなった動作要素の時間範囲に基づいて計算される。

元に戻って、図１５のステップ１５０６では、ステップ１５０５で認識された同時要素によって構成される逐次要素を認識する。

図２０は、図１５に示す逐次要素の認識処理のフローチャートである。

図２０では、まずステップ２００１において、認識された同時要素のバッファ中で新しく認識された同時要素を含む逐次要素を同時要素と逐次要素の対応リストから検索する。ステップ２００２において、検索された逐次要素の数をカウンタｉに代入する。ステップ２００３において、検索された逐次要素のｉ番目の逐次要素を構成する同時要素を、同時要素のバッファから検索する。ステップ２００４において、ｉ番目の逐次要素を構成する同時要素が全て検索された場合には、ステップ２００５に移る。必要な同時要素が全て見つからなかった場合には、ステップ２００７に移る。

【００１８】

ステップ２００５では、検索された同時要素の評価値に基づき、ｉ番目の逐次要素の評価値を求める。評価値は、ｉ番目の逐次要素を構成する同時要素の数をｎ，ｊ番目の同時要素の評価値をＥ４（ｊ），同時要素間の重なり度をＯ，同時要素間の重なり度の平均をＡ，分散をσとして下記（数４）の式によって求めることができる。なお、式中のＥ４は、上記（数１）の式から容易に求められる。

【数４】

$$E5=\left(\prod_{j=1}^{n} E4(j)\times e^{-\frac{(O-A)^2}{2\times\sigma}}\right)^{\frac{1}{n+1}}$$

また、式中の同時要素間の重なり度Ｏは、ｊ番目の同時要素の開始時刻をｓ（ｊ），終了時刻をｅ（ｊ）として下記（数５）の式によって求めることができる。

【数５】

$$O = \frac{\min_i e(i) - \max_i s(i)}{\max_i e(i) - \min_i s(i)}$$

ステップ２００６において、求めた評価値とその逐次要素を認識結果として出力する。出力される逐次要素のフォーマットを、図２１に示す。図２１において、２１０１は逐次要素の時間範囲の開始時刻，２１０２は逐次要素の時間範囲の終了時刻，２１０３は手話形態素辞書中における逐次要素の通し番号，２１０４はその逐次要素の評価値である。逐次要素の時間範囲は、その逐次要素の認識に用いられた同時要素の時間範囲に基づいて求める。例えば、全ての同時要素の重なり部分の時間範囲を用いることができる。あるいは、同時要素の時間範囲の開始時刻，終了時刻それぞれの平均を用いても良い。

【００１９】

図２２は、図１に示す静的動作要素認識部の処理を表すフローチャートである。

図１における静的動作要素認識部１０３の認識処理では、静的な動作要素の評価値を各時刻毎に求める。図２２のステップ２２０１において、手話形態素辞書から静的な動作要素を抽出し、静的動作要素のリストを作成する。この場合、動作要素状態が２以上の同時要素を構成する静的な動作要素は削除する。これは、前述の処理で終了しているためである。

動作要素リスト中の動作要素のフォーマットを図２３に示す。

図２３において、２３０１は、手話形態素辞書中における動作要素の通し番号、２３０２は動作要素の種類，２３０３はその動作要素を表現するために使用される手の部位，２３０４，２３０５はその動作要素に付属する属性値である。

図２２のステップ２２０２において、手話入力部から１時刻分のデータを読み込む。次に、ステップ２２０３において、データが最後であれば、処理を終了する。そうでなければ、ステップ２２０４に移る。ステップ２２０４では、静的動作要素リスト中の全ての動作要素について、動作要素の属性値と読み込んだデータからその時刻における評価値を求める。評価値は、前記（数１）の式によって求めることができる。ステップ２２０５では、求めた評価値と動作要素を認識結果として出力する。また、静的動作要素統合部１０６で過去の静的動作要素の認識結果を使用するため、静的動作要素のバッファを設けて、そこに認識結果を格納する。

出力される動作要素のフォーマットを図２４に示す。

図２４において、２４０１は時刻，２４０２は手話形態素辞書中における動作要素の通し番号，２４０３は動作要素の評価値である。

【００２０】

図２５は、図１に示す静的逐次要素認識部における認識処理のフローチャートである。

図１の静的逐次要素認識部１０５の認識処理では、静的動作要素認識部１０３で認識された静的動作要素の評価値から、静的動作要素のみによって構成される逐次要素の認識を行う。図２５のステップ２５０１において、手話形態素辞書から静的な動作要素のみによって構成される逐次要素を抽出し、逐次要素と静的動作要素の対応リストを作成する。

対応リスト中の逐次要素のフォーマットを、図２６に示す。

図２６において、２６０１は手話形態素辞書中における逐次要素の通し番号，２６０２は逐次要素を構成する静的動作要素の数，２６０３，２６０４は逐次要素を構成する静的動作要素の手話形態素辞書中における通し番号である。

図２５のステップ２５０２では、静的動作要素認識部１０３から静的動作要素の認識結果を１時刻分読み込む。次にステップ２５０３では、認識結果が最後であれば処理を終了する。そうでなければ、ステップ２５０４に移る。ステップ２５０４では、各逐次要素について、必要な動作要素を読み込んだ静的動作要素の認識結果から選択し、その時刻における逐次要素の評価値を求める。時刻ｔにおける評価値Ｅ７（ｔ）は、逐次要素を構成する

静的動作要素の数をｎ，ｉ番目の静的動作要素の時刻ｔにおける評価値をＥ６（ｔ，ｉ）として、下記（数６）の式によって求めることができる。なお、Ｅ６（ｔ，ｉ）は、前記（数１）の式により容易に求めることができる。

【数６】

$$E7(t) = \left(\prod_{j=1}^{n} E6(t, j) \right)^{\frac{1}{n}}$$

ステップ２５０５では、各逐次要素について、求めた逐次要素の評価値と過去の逐次要素の評価値の履歴を記憶してあるバッファの内容から、評価値が極大になる時間範囲を検索し、その時間範囲と評価値を逐次要素の認識結果として出力する。なお、逐次要素の認識結果のフォーマットは、図２１に示すフォーマットと同じである。

ステップ２５０６では、次の時刻での処理で使用するために、求めた逐次要素の評価値をバッファに格納する。

【００２１】

図２７は、図１に示す静的動作要素統合部における統合処理のフローチャートである。

静的動作要素統合部１０６では、動的な動作要素のみによって構成された逐次要素の評価結果に静的な動作要素の評価値を統合して、逐次要素全体の評価値を求める。図２７では、まずステップ２７０１において、手話形態素辞書中から動的な動作要素を含む逐次要素および動作要素状態数が２以上の静的動作要素から構成される同時要素を含む逐次要素を抽出し、逐次要素と静的な動作要素の対応リストを作成する。対応リスト中の各逐次要素のフォーマットは、図２６に示すフォーマットと同じである。

ステップ２７０２では、動的逐次要素認識部１０４から動的逐次要素の認識結果を一つ読み込む。ステップ２７０３において、動的逐次要素の認識結果が最後であれば処理を終了する。そうでなければ、ステップ２７０４に移る。ステップ２７０４では、読み込んだ動的逐次要素の時間範囲において、動的逐次要素に対応する静的動作要素の認識結果（評価値）を静的動作要素認識部１０３から読み込む。ステップ２７０５では、読み込んだ静的動作要素の評価値の平均を動的逐次要素に対応する静的動作要素の評価値として求める。ステップ２７０６では、動的逐次要素の評価値と静的動作要素の評価値とから、逐次要素全体の評価値を求める。この評価値の計算は、二種類の評価値の平均や相乗平均等を計算することにより求めることができる。また逐次要素を構成する動作要素の種類毎に評価値を保存しておき、それらの平均や相乗平均等を計算することにより求めることもできる。

ステップ２７０７では、動的逐次要素の時間範囲と求めた評価値を統合結果として出力する。統合結果のフォーマットは、図２１に示すフォーマットと同じである。

図２７に示すフローチャートでは、静的動作要素の評価値は全ての時刻について求め、その中から必要な時間範囲の評価値のみを受け取ることになるが、動的な逐次要素認識部の認識結果から得られる時間範囲についてのみ、必要な静的動作要素の評価値を求めるように静的動作要素認識部１０３に指示を送り、その評価結果を受け取って静的動作要素の統合処理を行うようにすることもできる。

【００２２】

図２８は、図１に示す手話形態素認識部における認識処理のフローチャートである。

手話形態素認識部１０７は、逐次要素の認識結果から手話形態素を認識する。図２８では、まずステップ２８０１において、手話形態素辞書から手話形態素とそれを構成する逐次要素の対応リストを作成する。対応リスト中の各手話形態素のフォーマットを図２９に示す。図２９において、２９０１は手話形態素名，２９０２は手話形態素を構成する逐次要素間の重なり度，２９０３は手話形態素を構成する逐次要素の数，２９０４，２９０５は手話形態素を構成する逐次要素の手話形態素辞書中における通し番号である。

ステップ２８０２において、逐次要素の認識結果を一つ読み込む。ステップ２８０３にお

いて、逐次要素の認識結果が最後であれば処理を終了する。そうでなければ、ステップ２８０４に移る。ステップ２８０４では、読み込んだ逐次要素と過去の逐次要素の履歴を格納してあるバッファ中の逐次要素から、構成される逐次要素列に対応する手話形態素を対応リストから検索する。ステップ２８０５では、検索された手話形態素について、逐次要素の評価値および逐次要素間の重なりあるいはギャップに基づいて評価値を求める。評価値Ｅ９は、手話形態素を構成する逐次要素の数をｎ，ｉ番目の逐次要素の評価値をＥ８（ｉ），ｉ番目とｉ＋１番目の逐次要素間の重なりあるいはギャップをＧ（ｉ）（重なりの場合正，ギャップの場合負），逐次要素間の重なりあるいはギャップの平均をＡ，分散をσとして、下記（数７）の式によって求めることができる。なお、Ｅ８（ｉ）は、前式（数１）により容易に求められる。

【数７】

$$E9 = \left(\prod_{i=1}^{n} E8(i) \times \prod_{i=1}^{n-1} e^{-\frac{(G(i)-A)^2}{2\times\sigma}} \right)^{\frac{1}{2\times n-1}}$$

【００２３】

ステップ２８０６において、求めた評価値，手話形態素名および時間範囲を認識結果として出力する。手話形態素の時間範囲は、認識に使用された逐次要素の時間範囲に基づいて求める。例えば、最初の逐次要素の開始時刻から最後の逐次要素の終了時刻を、手話形態素の時間範囲とすることができる。

図３０に、手話形態素の認識結果のフォーマットを示す。図３０において、３００１は手話形態素の時間範囲の開始時刻，３００２は手話形態素の時間範囲の終了時刻，３００３は手話形態素名，３００４は手話形態素の評価値である。

なお、前式（数１），（数２），（数３），（数４），（数６），（数７）では、評価値の計算はそれぞれを構成要素の評価値の相乗平均を用いているが、単純な平均（構成要素の評価値を全て加算し、その数で割算する）を用いても良い。また、構成要素の評価値とギャップや重なりに対する評価値に対して重み付けをしても良い。

【００２４】

【発明の効果】

以上説明したように、本発明によれば、動的な動作要素の認識結果から決定される時間範囲に基づいて静的な動作要素を評価するため、静的な動作要素の認識範囲による認識精度の低下がなくなる。また静的な動作要素のみによって構成される逐次要素についても、各時刻毎に逐次要素全体の評価値を求め、評価値の極大値となる時刻を検出することにより認識を行うため、それぞれの動作要素の検出時間のずれによる認識低下がなくなる。さらに動作要素の特徴を表す属性値を、実際の動作データから求めた連続量によって表現し、それに基づいて認識された動作要素を評価するため、動作要素の適切な評価値を求めることが可能となり、手話の認識精度を向上することができる。

【図面の簡単な説明】

【図１】本発明の一実施例を示す手話認識装置の概念ブロック図である。

【図２】本発明において、動作要素に基づく手話動作モデルを示す図である。

【図３】図１に示す動的動作要素認識部の構造を示すブロック図である。

【図４】図１に示す静的動作要素認識部の構造を示すブロック図である。

【図５】本発明の一実施例を実現するための手話認識装置のハードウェア構成図である。

【図６】図１の手話入力装置から入力されるデータのフォーマット図である。

【図７】動的動作要素認識のためのパラメータのフォーマット図である。

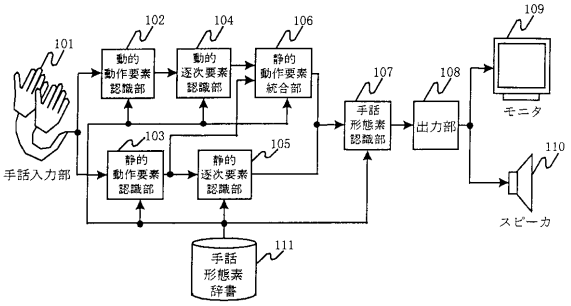
【図８】手話形態素辞書の記述フォーマットである。

【図９】動作要素の記述フォーマットである。

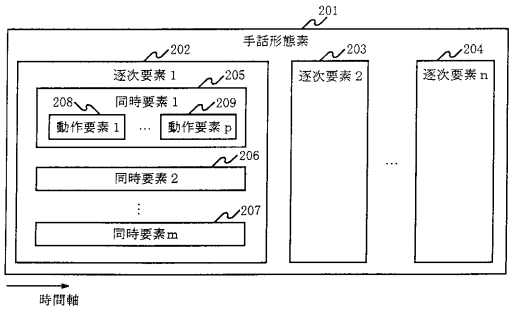
【図１０】動作要素の種類およびそれぞれの属性値の種類を示す図である。

【図１１】動作要素の属性値のフォーマットである。
【図１２】動作要素状態数が２以上の同時要素に含まれる静的動作要素の認識処理を説明するためのフローチャートである。
【図１３】静的動作要素の検出処理を説明するための図である。
【図１４】図１３において認識された動的動作要素のフォーマット図である。
【図１５】図１に示す動的逐次要素認識部の認識処理を説明するためのフローチャートである。
【図１６】同時要素と動作要素の対応リストにおける同時要素のフォーマット図である。
【図１７】逐次要素と同時要素の対応リストにおける逐次要素のフォーマット図である。
【図１８】図１に示す動的逐次要素認識部における同時要素の認識処理を説明するためのフローチャートである。
【図１９】認識された同時要素のフォーマット図である。
【図２０】図１に示す動的逐次要素認識部における逐次要素の認識処理を説明するためのフローチャートである。
【図２１】図２０で認識された逐次要素のフォーマット図である。
【図２２】図１に示す静的動作要素認識部の認識処理を説明するためのフローチャートである。
【図２３】静的動作要素リスト中の動作要素のフォーマット図である。
【図２４】図２２で認識された静的動作要素のフォーマット図である。
【図２５】図１に示す静的逐次要素認識部の認識処理を説明するためのフローチャートである。
【図２６】静的逐次要素リスト中の逐次要素のフォーマット図である。
【図２７】図１に示す静的動作要素統合部の統合処理を説明するためのフローチャートである。
【図２８】図１に示す手話形態素認識部の認識処理を説明するためのフローチャートである。
【図２９】手話形態素と逐次要素の対応リスト中の手話形態素のフォーマット図である。
【図３０】図２８で認識された手話形態素のフォーマット図である。
【符号の説明】
１０１…手話入力部、１０２…動的動作要素認識部、１０８…出力部、
１０３…静的動作要素認識部、１０４…動的逐次要素認識部、
１０５…静的逐次要素認識部、１０６…動的動作要素統合部、
１０７…手話形態素認識部、１０９…モニタ、１１０…スピーカ、
１１１…手話形態素辞書、２０１…手話形態素、２０８，２０９…動作要素、
２０２，２０３，２０４…逐次要素、２０５，２０６，２０７…同時要素、
３０１，３０２，３０３…動的動作要素認識部、３０４～３０６…パラメータ、
４０１，４０２，４０３…静的動作要素認識部、５０４～５１１…メモリ、
６０１…手の位置、６０５…手の方向、６０９…指の曲げ、
７０１…動作要素名、７０５，７０７…パラメータ種類、
７０６，７０８…パラメータ、９０１…動作要素の種類、９０２…手の部位、
９０３～９０５…属性値、１１０１～１１０３…属性値の平均値、
１１０４～１１０６…属性値の分散、１３０１…評価値の時系列、
１３０２…時間範囲の開始時刻、１３０３…時間範囲の終了時刻。

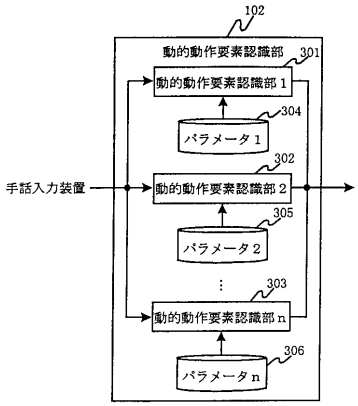
【図１】



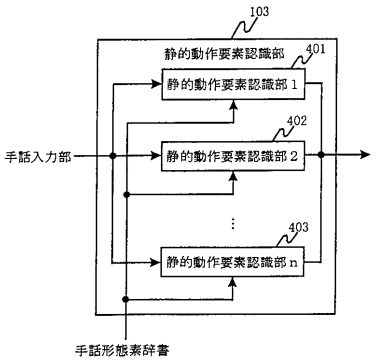
【図２】



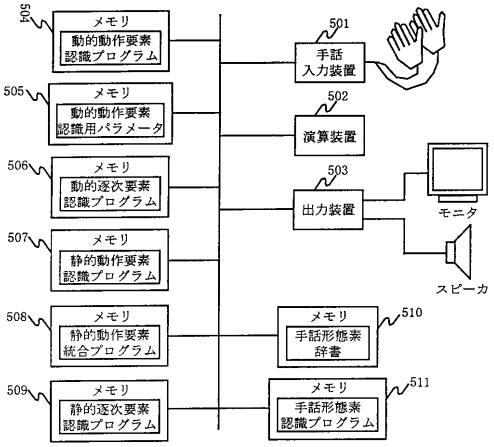
【図３】



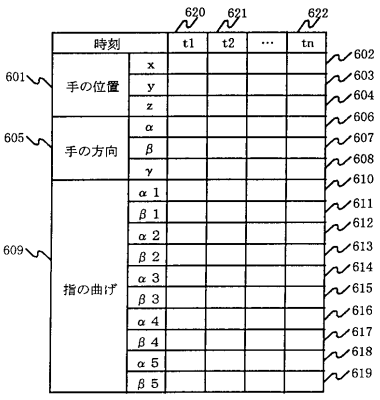
【図４】



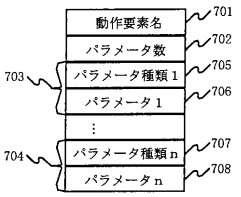
【図５】



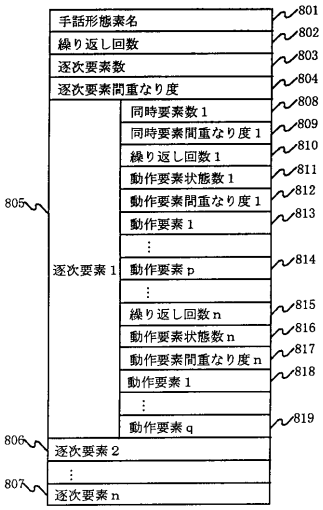
【図６】



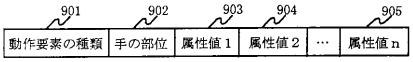
【図７】



【図８】



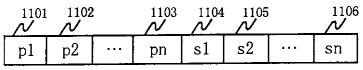
【図９】



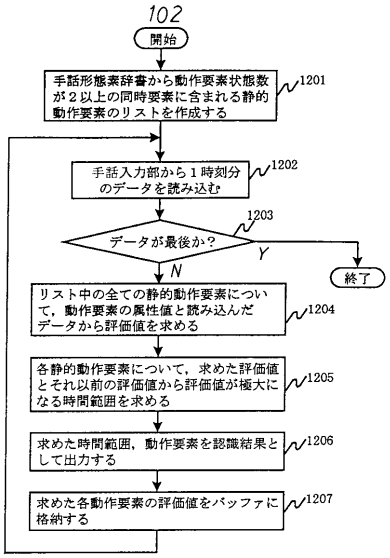
【図１０】

動作要素の種類	属性値の種類
形状	各指の間接の曲げ角度
形状振動	各指の間接の曲げ角度，振幅，周期
方向	方向ベクトル
方向振動	振動中心方向，振動平面，振幅，周期
位置	位置ベクトル
位置振動	振動中心位置，振動方向，振幅，周期
直線円弧	動作方向，動作平面，始点，終点，直線度，長さ，速さ
円	中心位置，動作平面，長径，短径，速さ
波線	動作方向，動作平面，始点，終点，移動距離，振幅，周期
螺旋	動作方向，動作平面，始点，終点，移動距離，長径，短径，速さ
関係形状	右手各指の間接の曲げ角度，左手各指の間接の曲げ角度
関係方向	右手の方向ベクトルと左手の方向ベクトルの角度
関係距離	右手と左手の距離
関係位置	--方の手を基準とした他方の手の相対位置ベクトル

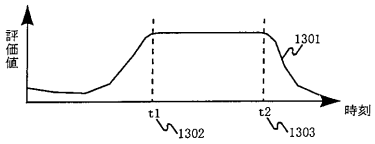
【図１１】



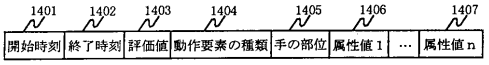
【図１２】



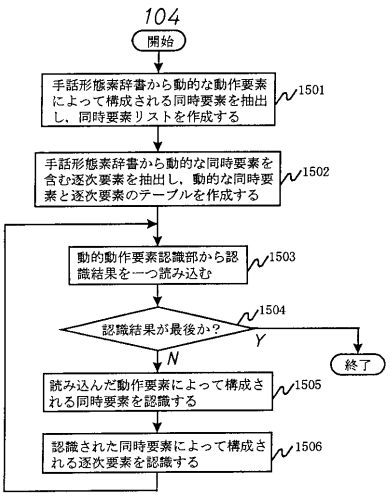
【図１３】



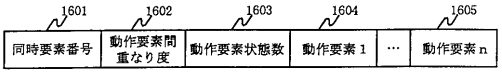
【図１４】



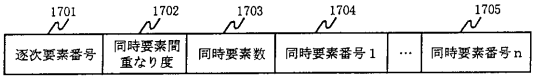
【図１５】



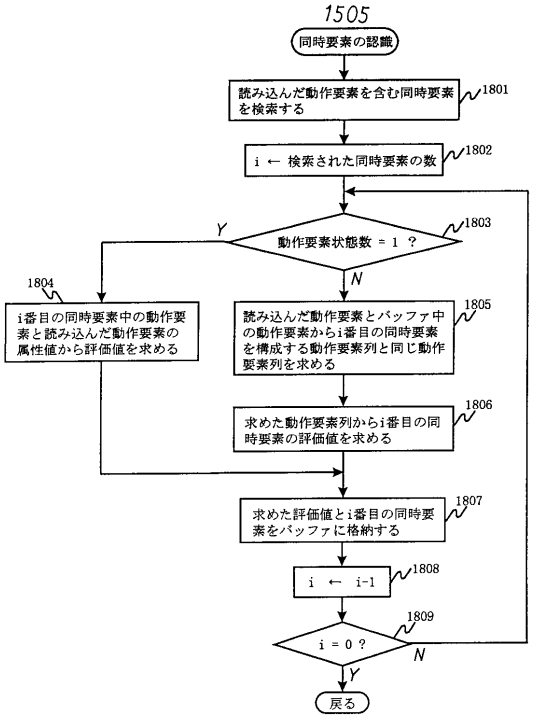
【図１６】



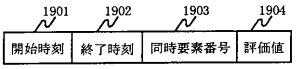
【図１７】



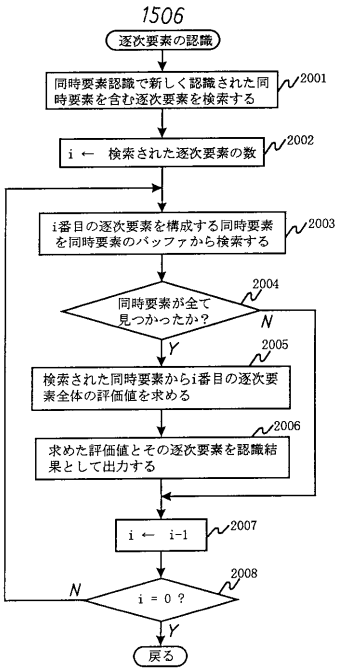
【図１８】



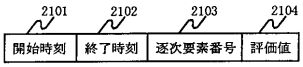
【図１９】



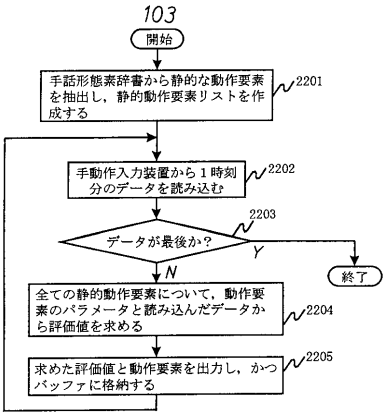
【図２０】



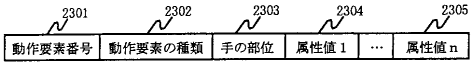
【図２１】



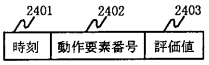
【図２２】



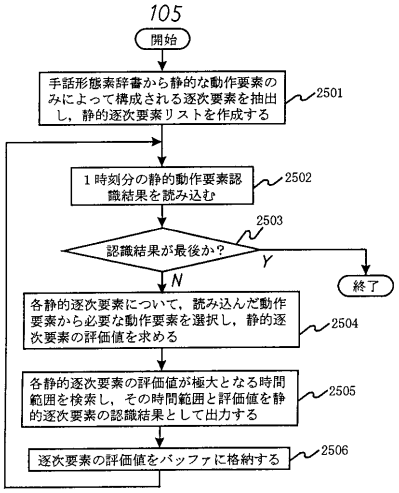
【図２３】



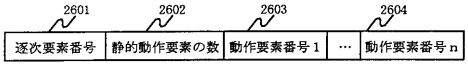
【図２４】



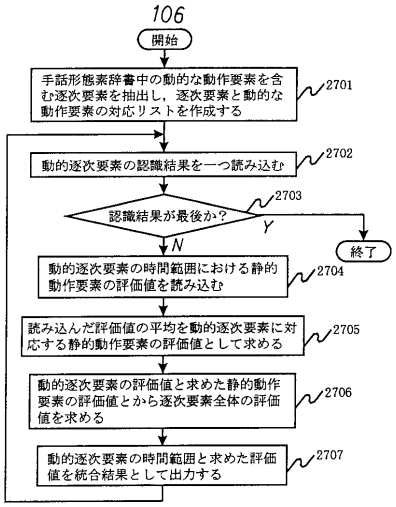
【図２５】



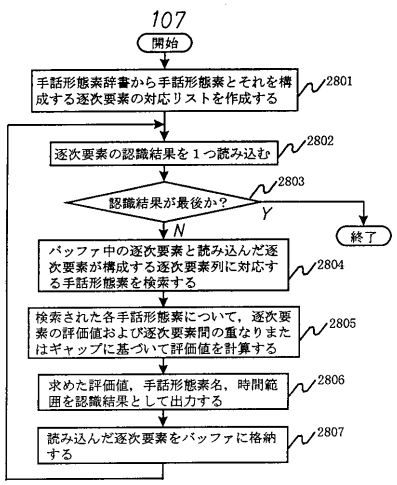
【図２６】



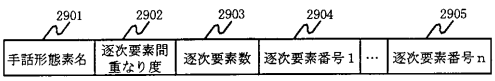
【図２７】



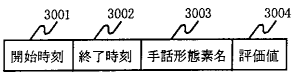
【図２８】



【図２９】



【図３０】



フロントページの続き

(72)発明者　大木　優
東京都国分寺市東恋ケ窪一丁目２８０番地　株式会社日立製作所中央研究所内

審査官　脇岡　剛

(56)参考文献　特開平０８－２７２９７４（ＪＰ，Ａ）
特開平０７－２８２２３５（ＪＰ，Ａ）
特開平０８－１１５４０８（ＪＰ，Ａ）
特開平０６－０６７６０１（ＪＰ，Ａ）
特開平０６－３３３０２２（ＪＰ，Ａ）
特開昭６２－０３４２８５（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.[7]，ＤＢ名)
G06T　1/00　340
G06F　3/033　310
G06T　7/20　300
G09B 21/00
JICSTファイル(JOIS)