

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5846552号
(P5846552)

(45) 発行日 平成28年1月20日(2016.1.20)

(24) 登録日 平成27年12月4日(2015.12.4)

(51) Int.Cl.

G06T 7/20 (2006.01)
G09B 21/00 (2006.01)

F 1

G06T 7/20 300A
G09B 21/00 F

請求項の数 6 (全 11 頁)

(21) 出願番号 特願2011-125394 (P2011-125394)
 (22) 出願日 平成23年6月3日 (2011.6.3)
 (65) 公開番号 特開2012-252581 (P2012-252581A)
 (43) 公開日 平成24年12月20日 (2012.12.20)
 審査請求日 平成26年5月26日 (2014.5.26)

特許法第30条第1項適用 平成22年12月4日 公益社団法人計測自動制御学会九州支部発行の「第29回計測自動制御学会九州支部学術講演会予稿集」に発表

(73) 特許権者 504174135
 国立大学法人九州工業大学
 福岡県北九州市戸畠区仙水町1番1号
 (74) 代理人 100090697
 弁理士 中前 富士男
 (74) 代理人 100127155
 弁理士 来田 義弘
 (74) 代理人 100163267
 弁理士 今中 崇之
 (72) 発明者 タン ジューケイ
 福岡県北九州市戸畠区仙水町1-1 九州
 工業大学内
 (72) 発明者 石川 聖二
 福岡県北九州市戸畠区仙水町1-1 九州
 工業大学内

最終頁に続く

(54) 【発明の名称】手話認識方法及びその装置

(57) 【特許請求の範囲】

【請求項1】

手話者の手話を画像取得手段によって撮像して画像データに変換し、変換した前記画像データを認識手段によって文字として認識し、認識結果を前記手話者の意思として伝達手段によって外部に伝達する手話認識方法において、

前記画像取得手段を前記手話者の脳部前側に装着させると共に、前記認識手段及び前記伝達手段を前記手話者に装着させて、該画像取得手段、該認識手段、及び該伝達手段を該手話者と共に移動させることを特徴とする手話認識方法。

【請求項2】

請求項1記載の手話認識方法において、前記手話者の手話を、手話の特徴情報を基づいて複数のグループに分類し、該グループ毎に手話の基準画像データと該基準画像データに対応する文字を記憶してデータベースを予め作成し、前記画像データから前記特徴情報を抽出して該画像データが属する前記データベース内の前記グループを特定し、前記画像データと特定された前記グループ内の前記基準画像データとを比較して該画像データに類似する前記基準画像データを選定し、選定した該基準画像データに対応する文字を認識して、認識した文字を、文字画像、音声、又は音声付き文字画像として出力することを特徴とする手話認識方法。

【請求項3】

請求項2記載の手話認識方法において、前記手話者の手話は指文字であって、前記特徴情報は、手の向きと伸ばした指の本数から構成されることを特徴とする手話認識方法。

10

20

【請求項 4】

手話者の手話を画像取得手段によって撮像して画像データに変換し、変換した前記画像データを認識手段によって文字として認識し、認識結果を前記手話者の意思として伝達手段によって外部に伝達する手話認識装置において、

前記画像取得手段を前記手話者の胸部前側に装着させると共に、前記認識手段及び前記伝達手段を前記手話者に装着させて、該画像取得手段、該認識手段、及び該伝達手段を該手話者と共に移動させることを特徴とする手話認識装置。

【請求項 5】

請求項 4 記載の手話認識装置において、前記手話者の手話を、手話の特徴情報に基づいて複数のグループに予め分類し、該グループ毎に手話の基準画像データと該基準画像データに対応する文字を記憶して作成したデータベースを有し、前記画像データから前記特徴情報抽出して該画像データが属する前記データベース内の前記グループを特定し、前記画像データと特定された前記グループ内の前記基準画像データとを比較して該画像データに類似する前記基準画像データを選定し、選定した該基準画像データに対応する文字を認識して、認識した文字を、文字画像、音声、又は音声付き文字画像として出力することを特徴とする手話認識装置。

10

【請求項 6】

請求項 5 記載の手話認識装置において、前記手話者の手話は指文字であって、前記特徴情報は、手の向きと伸ばした指の本数から構成されることを特徴とする手話認識装置。

【発明の詳細な説明】

20

【技術分野】**【0001】**

本発明は、手話者（例えば、発話障害者、聴覚障害者）の手話を文字として認識し、認識結果を手話者の意思として外部に伝達する手話認識方法及びその装置に関する。

【背景技術】**【0002】**

発話障害者間のコミュニケーションは、例えば手話を介して行なわれているが、発話障害者が健常者に意思の伝達を行うには、健常者が手話を理解しなければ容易でない。そこで、図 6 に示すように、健常者 80 側に、例えばカメラ 81 を設置して手話行為者（手話者）82 が行う手話を撮影し、得られた画像データから手話者 82 の手話を、例えば文字として認識して、認識結果を健常者 80 が認識できる形式に変換して（図 6 では、「この方は、郵便局はどこかと尋ねています」という音声に変えて）伝える手話認識装置 83 や手法が種々提案されている（例えば、特許文献 1、非特許文献 1、2 参照）。

30

【先行技術文献】**【特許文献】****【0003】**

【特許文献 1】特許第 2779448 号公報

【非特許文献】**【0004】**

【非特許文献 1】三船智明、江刺宏恭、千種康民「手話 C A I のための手話記述言語と G U I の開発」、情報システム研究会技術報告、情報処理学会、1996 年 3 月、第 1996 卷、第 30 号、p. 7 - 14

40

【非特許文献 2】有賀光希、酒向慎司、北村正「日本手話の音韻構造を考慮した HMM に基づく手話認識」、電子情報通信学会技術研究報告、PRMU 研究会、2010 年 10 月、第 110 卷、第 219 号、p. 127 - 132

【発明の概要】**【発明が解決しようとする課題】****【0005】**

しかしながら、提案された手話認識装置 83 は、健常者 80 側に手話認識装置 83 を設置することを前提としているため、手話者 82 からの意思伝達は手話認識装置 83 を所持す

50

る健常者 80 に限定され、不特定多数の者に対して手話者 82 の意思を伝達することができないという問題がある。そして、手話を撮影するカメラ 81 が手話者 82 の正面に位置していないと、手話の正確な画像を取得することができず、手話の認識率が低下するという問題がある。また、カメラ 81 と手話者 82 との間が離れると、手話画像に加えて背景等の不要な画像も同時に取込まれるため、手話画像のみを抽出する前処理が必要になり、前処理で得られる手話画像の精度により手話の認識率が変動するという問題も生じる。更に、健常者 80 が、常時必要としない手話認識装置 83 を携行するということは期待できず、手話認識装置 83 の普及という面において現実的でないという問題がある。

【 0 0 0 6 】

本発明はかかる事情に鑑みてなされたもので、手話者が不特定多数の者に対して手話者の意思を正確に伝達することが可能な手話認識方法及びその装置を提供することを目的とする。 10

【課題を解決するための手段】

【 0 0 0 7 】

前記目的に沿う本発明に係る手話認識方法は、手話者の手話を画像取得手段によって撮像して画像データに変換し、変換した前記画像データを認識手段によって文字として認識し、認識結果を前記手話者の意思として伝達手段によって外部に伝達する手話認識方法において、

前記画像取得手段を前記手話者の脛部前側に装着させると共に、前記認識手段及び前記伝達手段を前記手話者に装着させて、該画像取得手段、該認識手段、及び該伝達手段を該手話者と共に移動させる。 20

【 0 0 0 8 】

本発明に係る手話認識方法において、前記手話者の手話を、手話の特徴情報に基づいて複数のグループに分類し、該グループ毎に手話の基準画像データと該基準画像データに対応する文字を記憶してデータベースを予め作成し、前記画像データから前記特徴情報を抽出して該画像データが属する前記データベース内の前記グループを特定し、前記画像データと特定された前記グループ内の前記基準画像データとを比較して該画像データに類似する前記基準画像データを選定し、選定した該基準画像データに対応する文字を認識して、認識した文字を、文字画像、音声、又は音声付き文字画像として出力することが好ましい。 30

【 0 0 0 9 】

本発明に係る手話認識方法において、前記手話者の手話は指文字であって、前記特徴情報は、手の向きと伸ばした指の本数から構成することができる。

【 0 0 1 0 】

前記目的に沿う本発明に係る手話認識装置は、手話者の手話を画像取得手段によって撮像して画像データに変換し、変換した前記画像データを認識手段によって文字として認識し、認識結果を前記手話者の意思として伝達手段によって外部に伝達する手話認識装置において、

前記画像取得手段を前記手話者の脛部前側に装着させると共に、前記認識手段及び前記伝達手段を前記手話者に装着させて、該画像取得手段、該認識手段、及び該伝達手段を該手話者と共に移動させている。 40

【 0 0 1 1 】

本発明に係る手話認識装置において、前記手話者の手話を、手話の特徴情報に基づいて複数のグループに予め分類し、該グループ毎に手話の基準画像データと該基準画像データに対応する文字を記憶して作成したデータベースを有し、前記画像データから前記特徴情報を抽出して該画像データが属する前記データベース内の前記グループを特定し、前記画像データと特定された前記グループ内の前記基準画像データとを比較して該画像データに類似する前記基準画像データを選定し、選定した該基準画像データに対応する文字を認識して、認識した文字を、文字画像、音声、又は音声付き文字画像として出力することが好ましい。

【 0 0 1 2 】

10

20

30

40

50

本発明に係る手話認識装置において、前記手話者の手話は指文字であって、前記特徴情報は、手の向きと伸ばした指の本数から構成することができる。

【発明の効果】

【0013】

本発明に係る手話認識方法及びその装置においては、手話者の手話を撮像して画像データに変換する画像取得手段を、手話をっている手話者の脇部前側に配置するので、画像取得手段を手話者と共に移動させることができる。これにより、手話者は、手話を認識できない不特定多数の者に対して、随時意思の伝達を容易に図ることができる。そして、手話者は、使い慣れた手話を使用するので、筆談、キーボード入力による文字表示や音声出力等の他の伝達方式と比較して、自分の意思を迅速、臨機応変、かつ容易に伝達することができる。

10

また、画像取得手段が、画像取得手段を常時必要とする手話者側に所持されるため、手話認識方法及びその装置の普及を促進することができる。更に、画像取得手段が手話者の脇部前側に配置されるので、手話を近距離から撮影することができる。これにより、手話の正確な画像データを取得することができ、手話の認識が容易にできると共に、背景等の不要画像の取込が制限されるため、手話の認識を安定して行うことができる。

【0014】

本発明に係る手話認識方法及びその装置において、手話者の手話を、手話の特徴情報に基づいて複数のグループに分類し、グループ毎に手話の基準画像データと基準画像データに対応する文字を記憶してデータベースを予め作成する場合、手話者の特徴（癖）を考慮した手話認識方法及びその装置を構築することができる。これによって、手話の認識率を向上することができる。そして、手話者が自作した指文字（例えば、熟語や慣用語句等で示される使用頻度の高い表現や複雑な意思を、簡潔に表示する手話者専用の指文字）の基準画像データをデータベースに登録すれば、手話者の自作した指文字の認識が可能になる。これにより、手話者は、通常の指文字に加えて、自作した指文字を自由に使用することができ、通常の指文字と自作の指文字を組合せて使用することで、自分の意思を適格、迅速、かつ容易に伝達することができる。

20

また、画像データから特徴情報を抽出して画像データが属するデータベース内のグループを特定し、画像データと特定されたグループ内の基準画像データとを比較して画像データに類似する基準画像データを選定する場合、基準画像データの選定精度を向上することができる。これによって、手話の認識率を向上することができる。

30

更に、選定した基準画像データに対応する文字を認識して、認識した文字を、文字画像、音声、又は音声付き文字画像として出力する場合、手話を認識できない不特定多数の者に対して確実に手話者の意思を伝達することができる。

【0015】

本発明に係る手話認識方法及びその装置において、手話者の手話が指文字である場合、手話を一文字ずつ認識することができ、正確な意思の伝達が可能になる。

また、特徴情報が、手の向きと伸ばした指の本数から構成される場合、指文字の分類を容易に行うことができる。

【図面の簡単な説明】

40

【0016】

【図1】本発明の一実施の形態に係る手話認識装置の説明図である。

【図2】同手話認識装置による手話の認識手順を示す説明図である。

【図3】手の特徴情報の抽出手順の説明図である。

【図4】(A)はシルエット画像、(B)は細線化画像、(C)は端点及び端点円の画像の説明図である。

【図5】指文字の大まかな分類の説明図である。

【図6】従来の手話認識装置の説明図である。

【発明を実施するための形態】

【0017】

50

続いて、添付した図面を参照しつつ、本発明を具体化した実施の形態につき説明し、本発明の理解に供する。

図1に示すように、本発明の一実施の形態に係る手話認識装置10は、手話者11の手話の一例である指文字を画像取得手段の一例であるビデオカメラ12によって画像データに変換し、変換した画像データを文字（例えば、図1では、「郵便局はどこですか」という文字列）として認識して、認識した文字を文字画像、音声、又は音声付き文字画像として手話を認識できない健常者13に伝達するものである。そして、手話認識装置10は、ビデオカメラ12を手話をっている手話者11の胸部前側（例えば胸部）に配置し、ビデオカメラ12によって、ビデオカメラ12の前で行なわれる手話者11の指文字を撮像し、しかも、手話者11に装着されて手話者11と共に移動する。ここで、指文字は、平仮名一文字ずつと対応しているので、手話者11の意思を一文字ずつ認識することができ、正確な意思の伝達が可能になる。なお、手話認識装置10は、手話者11の手の動きを伴う「の、も、り、を、ん」を除く平仮名に対応する指文字を対象としている。以下、詳細に説明する。

【0018】

手話認識装置10は、手話者11の複数の指文字をビデオカメラ12で予め撮影して指文字毎に基準画像データを取得し、得られた基準画像データにそれぞれ文字を対応させてデータベースとして記憶するデータベース作成機能を備えた学習手段を有している。また、手話認識装置10は、手話者11が意思の伝達のために表示した指文字の画像データと記憶している基準画像データとを比較して、画像データに類似する基準画像データを選定し、選定した基準画像データに対応する文字を認識し、認識した文字の文字信号を出力する判別機能を備えた認識手段を有している。更に、手話認識装置10は、認識した文字の文字信号を、例えば音声として出力する伝達手段の一例であるスピーカ14を有している。ここで、学習手段及び認識手段は、データベース作成機能及び判別機能を発現するプログラムを、例えば、可搬式パーソナルコンピュータ（ウェアブルパーソナルコンピュータ）15に搭載することにより構成できる。

【0019】

手話者11の指文字を撮像して基準画像データを作成するので、手話者12の特徴（癖）を考慮してデータベースを構築することができる。また、指文字は、「手の向き（上方向、横（水平）方向、下方向の3区分）」及び「伸ばした指の本数」から構成される手の特徴情報に基づいて複数のグループに分類することができる。このため、データベースは、指文字の手の特徴情報に基づいて分類されたグループ毎に、基準画像データと基準画像データに対応する文字を記憶している。ここで、指文字を表示している指は、重なることがあるため、「伸ばした指の本数」は、0又は1本、1±1本、2±1本、3±1本、4±1本、5又は4本の6区分に分類する。

【0020】

認識手段では、入力された画像データから、「手の向き」及び「伸ばした指の本数」で構成される手の特徴情報を抽出し、画像データが、指文字の特徴情報に基づいて予め分類されたデータベース内のどのグループに属するかを決定する。画像データが属するグループが決定されると、決定されたグループ内の基準画像データと画像データとを比較して、画像データに類似する基準画像データを選定し、選定した基準画像データに対応する文字を認識する。

【0021】

画像データから手の特徴情報を抽出する場合、先ず、画像データから細線化画像を作成して連結数が1の端点を全て探す。次いで、画像データのシルエット画像の輪郭部分と探した各端点との最小距離を求めて端点を中心とする円の半径とし、最大半径の円の中心となる端点位置を手首位置、それ以外の端点位置を指先位置とする。そして、画像データから手の重心位置を求めて、手首位置と重心位置の位置関係、指先位置同士の相対的位置関係から、手の向き（上方向、横（水平）方向、又は下方向）を決める。また、指先位置に対応する端点の個数を「伸ばした指の本数」とする。

10

20

30

40

50

【 0 0 2 2 】

ここで、認識手段は、画像データに対して文字を認識するたびに認識結果を認識順に予め設定された個数だけ記憶する認識記憶部を有し、画像データに対して文字が新たに認識される度に、新たに認識された文字と認識記憶部に記憶されている複数の文字から文字群を形成し、文字群を構成している文字の中で、予め設定された個数以上で、かつ最大個数となる文字を、画像データに対応する文字とする補間機能が設けられている。

【 0 0 2 3 】

手話認識装置 10 (即ち、ビデオカメラ 12、可搬式パーソナルコンピュータ 15、スピーカ 14) は、手話者 11 に装着されて手話者 11 と共に移動するので、手話者 11 は、指文字を認識できない不特定多数の健常者 13 に対して、隨時意思の伝達を容易に図ることができる。また、手話認識装置 10 が、手話認識装置 10 を常時必要とする手話者 11 に所持されるため、手話認識装置 10 の普及を促進することもできる。

また、ビデオカメラ 12 が手話者 11 の腕部前側に配置されるので、指文字を近距離から撮影することができ、指文字の正確な画像データの取得が可能になって、指文字の認識が容易にできる。そして、指文字の画像データを撮影する際、背景や通行人等の不要画像の取込が制限されるため、指文字の認識を安定して行うことができる。

更に、画像データが属するグループを特定して認識候補となる文字を絞り込むので、指文字を文字として認識する際の認識精度 (認識率) が向上する。そして、認識された文字は、スピーカ 14 から音声として出力されるので、指文字を認識できない健常者 13 に対して手話者 11 の意思を容易に伝達することができる。

【 0 0 2 4 】

続いて、手話認識装置 10 を使用して、手話者 11 が表示した指文字を認識する手話認識方法について説明する。

先ず、手話認識装置 10 に設けられたビデオカメラ 12 を手話者 11 の腕部前側に配置し、ビデオカメラ 12 の前で手話者 11 が表示する複数の指文字をビデオカメラ 12 で予め撮影し、学習手段を用いて指文字毎に基準画像データを取得し、得られた基準画像データにそれぞれ文字を対応させてデータベースとして記憶しておく。なお、データベースでは、指文字を、手の形状を示す手の向き及び伸ばした指の本数に基づいて予め複数のグループに分類し、グループ毎にグループに属する指文字の基準画像データと指文字に対応する文字とを対として記憶している。

【 0 0 2 5 】

手話者 11 は、腕部前側にビデオカメラ 12 を配置すると共に、データベースが構築された学習手段及び認識手段が搭載された可搬式パーソナルコンピュータ 15 を背負い、スピーカ 14 を肩に取付けることにより、手話認識装置 10 を装着する。そして、手話者 11 が自分の意思を、指文字を認識できない健常者 13 に伝達しようとする場合、図示しないスイッチを操作して手話認識装置 10 を起動させ、意思を伝える指文字をビデオカメラ 12 の前に表示する。

【 0 0 2 6 】

図 2 に示すように、手話者 11 がビデオカメラ 12 の前で表示した指文字は、ビデオカメラ 12 で撮像されて、入力画像が形成され、入力画像は画像データに変換されて、認識手段に入力される。認識手段に入力された画像データから手の特徴情報を抽出し、画像データが、指文字の特徴情報に基づいて予め分類されたデータベース内のどのグループに属するかを決定する。入力画像が属するグループが決定されると、決定されたグループ内の基準画像データと画像データとを比較して、画像データに類似する基準画像データを選定し、選定した基準画像データと対となった文字を特定する。そして、画像データに対応する文字が特定されると、特定された文字を画像データの認識結果として出力する(以上、認識処理)。

【 0 0 2 7 】

ここで、最新の画像データに対して文字が新たに認識されると、新たに認識された文字と認識記憶部に認識順に記憶されている複数の文字から文字群を形成し、文字群を構成して

10

20

30

40

50

いる文字の中で、予め設定された個数以上で、かつ最大個数となる文字を、画像データに対応する文字として決定する補間処理を行う。そして、認識手段からは、補間処理で決定された文字の文字信号が、スピーカ14に出力される。以下、手の特徴情報の抽出方法及び認識処理の方法についてそれぞれ説明する。

【0028】

手の特徴情報を抽出する場合、図3に示すように、入力画像中のノイズ除去を行った後、手の領域を抽出するために、肌色に合致する領域を求め、指文字の手の形状情報を取得する。肌色領域の抽出には、HSV色空間を用いる。変数である色相(H)、彩度(S)、及び明度(V)をそれぞれ、例えば、 $5 < H < 30$ 、 $S < 30$ 、 $V > 0.4$ と設定する。抽出結果は手のシルエット画像として得られる(図4(A)参照)。そして、シルエット画像から、肌色領域部分の重心位置を求める。10

【0029】

次いで、Hilditch(ヒルディッチ)の細線化アルゴリズムを用いてシルエット画像に細線化を施す。得られた細線化画像を図4(B)に示す。なお、図4(B)は、細線化画像をシルエット画像に重ねて示している。細線化画像において、各画素の8近傍を調べて連結数を求める。連結数が1である端点を探し、その端点を指先又は手首の候補とする。そして、求めた端点において、端点を中心とする円の半径を徐々に大きくして、円がシルエット画像の輪郭部分と最初に接触する際の半径を求める(図4(C)参照)。

【0030】

一般に、手首位置の端点を中心とする円の半径は、指先の端点を中心とする円の半径よりも大きいので、半径が最大となる円の中心(端点)を手首位置とすることことができ、手首以外の端点を指先位置とすることができます。そして、手首位置と重心位置との位置関係、端点同士の相対的位置関係から、画像データの手の向きが、上方向、横(水平)方向、下方向の3区分のいずれに該当するかを決める。また、手首以外の端点の個数から、伸ばした指の本数を求める。その結果、画像データが、データベース内で予め分類されたどのグループに属するかを決める(即ち、表示された指文字(画像データ)をデータベース内のグループに基づいて大まかに分類する)。20

【0031】

なお、画像データは、手の向きに関しては、上方向、横(水平)方向、下方向のいずれかに分類できるが、伸ばした指の本数に関しては、指の本数が0本又は5本の場合は2つのグループにそれぞれ含まれるとし、指の本数が1~4本の場合は3つのグループにそれぞれ含まれるとする。このため、図5に示すように、画像データから文字を認識するには、特定されたグループ毎に含まれる基準画像データと画像データとの比較を行う。30

【0032】

画像データからの文字の認識は、画像データが属すると特定されたグループの中で、画像データに類似する基準画像データを選定し、選定した基準画像に対応する文字を求ることになる。ここで、画像データに類似する基準画像データの選定には、例えば、グレースケールにおける固有空間法を使用する。なお、固有空間法を使用する場合、グループに含まれる基準画像データ毎に、基準画像データのグレースケールにおける輝度分布を構成する基準画像濃度値ベクトル(画素の濃度値を画素番号順に並べたもの)を求め、求めた基準画像濃度値ベクトルを固有ベクトルの線形結合で記述する際に固有ベクトル及び基準画像線形結合係数を決定して(学習して)、データベースに予め記憶しておく。40

【0033】

そして、画像データに類似する基準画像データを選定するには、先ず、画像データの輝度分布を構成する画像濃度値ベクトルを求め、画像データが属すると特定されたグループの有する固有ベクトルの線形結合で画像濃度値ベクトルを近似した際に使用する画像線形結合係数を求める。次いで、得られた画像線形結合係数に最も近い基準画像線形結合係数を探し出し、探し出した基準画像線形結合係数を有する基準画像濃度値ベクトルに対応する基準画像データを、画像データに最も類似するとして選定する。

【0034】

10

20

30

40

50

なお、固有ベクトル及び基準画像線形結合係数の決定、画像線形結合係数の算出に要する時間を短縮するため、基準画像濃度値ベクトルの主成分分析を行い、固有ベクトルの次元を圧縮することが好ましい。固有ベクトルの次元の圧縮は以下の手順で行う。先ず、基準画像濃度値ベクトルから構成される濃度値行列から分散共分散行列を求めて、分散共分散行列の固有値を求める。次いで、固有値を大きい順に並べ、最大の固有値から最小の固有値に向けて固有値を順次足し合せた固有値部分和を求め、固有値部分和を固有値総和（最大の固有値から最小の固有値までを全て足し合せたもの）で除した累積寄与率が、例えば0.8未満の範囲で、最大となる固有値部分和を求める。そして、求めた固有値部分和を構成する固有値にそれぞれ対応する固有ベクトルを対象とすることで、固有ベクトルの次元を圧縮することができる。

10

【0035】

ここで、認識手段には認識記憶部が設けられており、画像データに対して文字が認識される毎に（例えば1秒毎に）認識結果が順次設定された個数（例えば30個、時間にして1秒間分）だけ記憶されている。そして、取得した画像データに対して文字が新たに認識されると、新たに認識された文字と認識記憶部に記憶されている複数（30個）の文字から文字群を形成し、文字群を構成している文字の中で、例えば度数（個数）が5以上で、かつ最大度数（最大個数）の文字を、画像データに対応する文字として認識する補間機能が設けられている。これによって、順次表示される指文字と指文字の間で誤認識が行なわれるのを低減できる。

20

そして、認識手段からは、画像データに対応する文字の文字信号が、スピーカ14に出力され、指文字を認識できない健常者13に対して、音声として手話者11の意思を伝達する。

【実施例】

【0036】

手話者の胸部にビデオカメラを取付け、ビデオカメラの前で表示される41種類の指文字を撮影し、画像データを取得した。画像データは、41種類の指文字の静止画像データからなるデータ群を1回分として6回分からなる（即ち、246枚の静止画像データから構成される）。そして、5回分の静止画像データを学習（基準画像データの作成）に用い、残りの1回分の静止画像データに対して図2に示す手順で文字の認識を行うLeave-one-out（リープ・ワン・アウト）法による評価を6回繰り返した。手の特徴情報を抽出して指文字の画像データを予め分類してから文字の認識を行う方法、即ち、固有ベクトルで張られる一つの固有空間の有する基準画像データ数（クラス数）を減らして認識処理を行うことで、246個の指文字の画像データ中、196個の指文字の画像データに対して正しい文字が認識され、認識率は79.7%であった。

30

【0037】

なお、同一の画像データを使用してLeave-one-out法による評価を行う際に、手の特徴情報を考慮せず、41種類の指文字の基準画像データを学習して固有ベクトル及び基準画像線形結合係数を決定し、文字の認識を行うと、246個の指文字の画像データ中、78個の指文字の画像データに対して正しい文字が認識され、認識率は31.7%であった。このことから、手の特徴情報を抽出して指文字の画像データを予め分類してから文字の認識を行うと、認識率を向上できることが確認できた。

40

【0038】

以上、本発明を、実施の形態を参照して説明してきたが、本発明は何ら上記した実施の形態に記載した構成に限定されるものではなく、特許請求の範囲に記載されている事項の範囲内で考えられるその他の実施の形態や変形例も含むものである。

更に、本実施の形態とその他の実施の形態や変形例にそれぞれ含まれる構成要素を組合せたものも、本発明に含まれる。

例えば、スピーカの代わりにディスプレイを装着することで、認識した文字を文字画像として表示できる。また、スピーカと共にディスプレイを装着することで、音声付き文字画像として出力することができる。

50

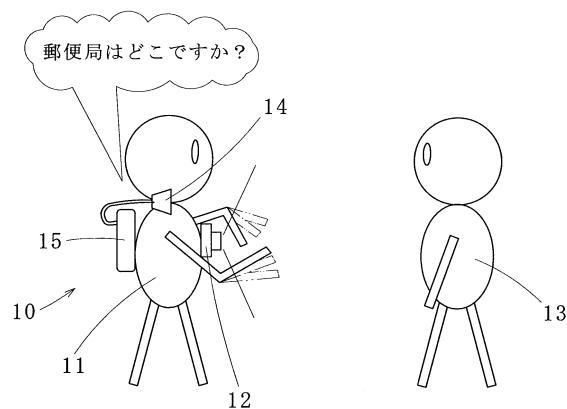
更に、認識した文字の文字画像及び／又は音声の伝達手段に、健常者（手話者が意思を伝達しようとする相手）の有する携帯電話のディスプレイ及び／又はスピーカを利用するようすることもできる。

【符号の説明】

【0039】

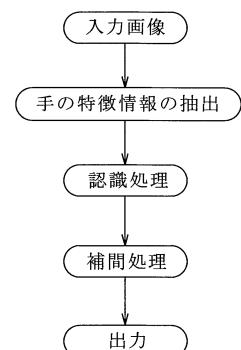
10：手話認識装置、11：手話者、12：ビデオカメラ、13：健常者、14：スピーカ、15：可搬式パーソナルコンピュータ

【図1】

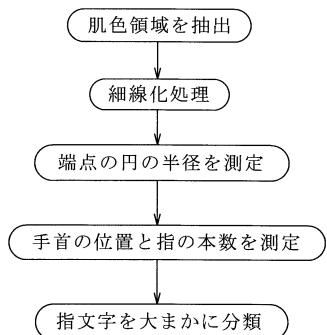


【図2】

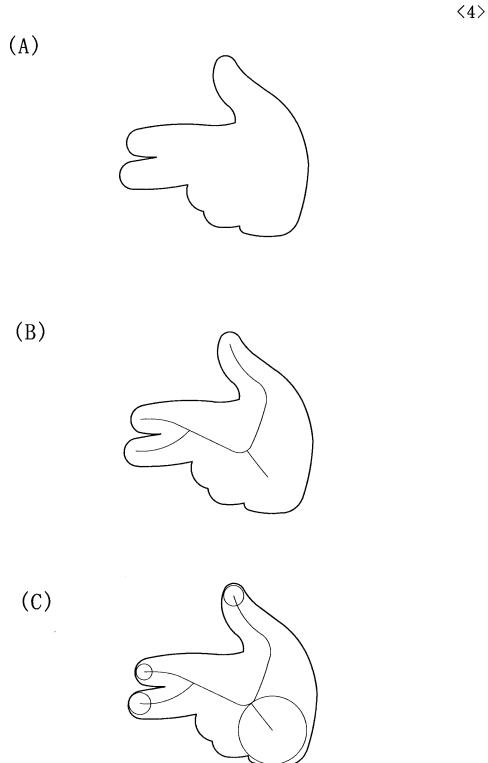
<1> <2>



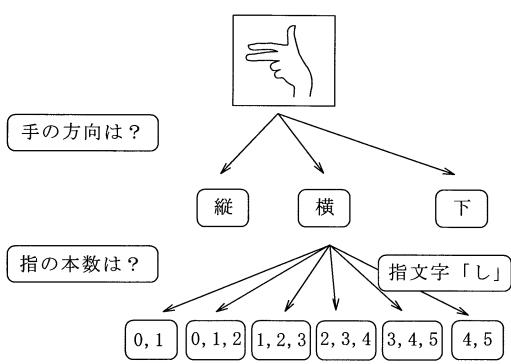
【図3】



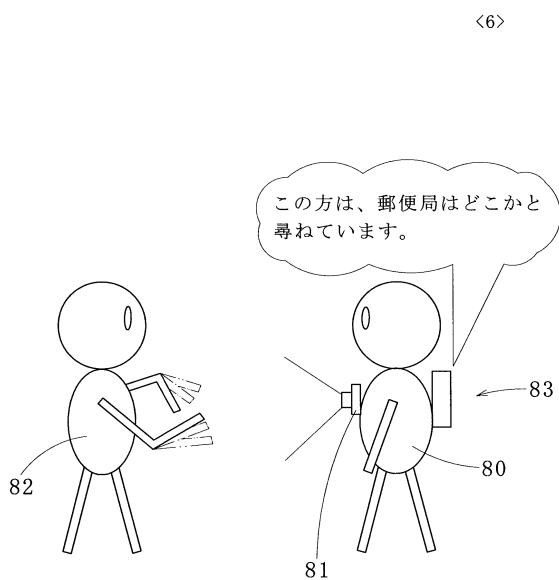
【図4】



【図5】



【図6】



フロントページの続き

審査官 佐藤 実

(56)参考文献 特開2007-241833(JP, A)

特開平02-144675(JP, A)

特開平10-26999(JP, A)

特開平9-319297(JP, A)

保利栄作 外1名, 距離画像を用いたウェアラブルな3次元デスクトップ環境の構築, 電子情報
通信学会技術研究報告 PRMU2001-222~232 パターン認識・メディア理解, 社
団法人電子情報通信学会, 2002年 2月14日, 第101巻 第652号, 第1-8頁

(58)調査した分野(Int.Cl., DB名)

G06T 7/20

G09B 21/00