

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6144192号
(P6144192)

(45) 発行日 平成29年6月7日(2017.6.7)

(24) 登録日 平成29年5月19日(2017.5.19)

(51) Int.Cl.

F 1

G06F	3/01	(2006.01)	G06F	3/01	570
G06T	7/20	(2017.01)	G06T	7/20	300A
G09B	21/00	(2006.01)	G09B	21/00	E

請求項の数 5 (全 13 頁)

(21) 出願番号	特願2013-272408 (P2013-272408)
(22) 出願日	平成25年12月27日 (2013.12.27)
(65) 公開番号	特開2015-125766 (P2015-125766A)
(43) 公開日	平成27年7月6日 (2015.7.6)
審査請求日	平成28年8月18日 (2016.8.18)

(73) 特許権者	392026693 株式会社 NTT ドコモ 東京都千代田区永田町二丁目11番1号
(74) 代理人	100088155 弁理士 長谷川 芳樹
(74) 代理人	100113435 弁理士 黒木 義樹
(74) 代理人	100121980 弁理士 沖山 隆
(74) 代理人	100128107 弁理士 深石 賢治
(72) 発明者	小林 茂子 東京都千代田区永田町二丁目11番1号 株式会社 NTT ドコモ内

最終頁に続く

(54) 【発明の名称】 画像認識装置、及び画像認識方法

(57) 【特許請求の範囲】

【請求項 1】

撮影対象者を画像データにして取得する画像取得手段と、

前記画像取得手段により取得された画像データに基づいて、撮影対象者の第1特定部分の特徴点を複数検出する第1特徴点検出手段と、

前記画像取得手段により取得された画像データに基づいて、撮影対象者の第1特定部分以外の部分である第2特定部分の特徴点を複数検出する第2特徴点検出手段と、

前記第1特徴点検出手段が検出した特徴点に基づいて、撮影対象者のジェスチャの区切りを特定する区切り特定手段と、

前記区切り特定手段が特定した区切りに基づいた期間における、前記第1特徴点検出手段が検出した複数の特徴点に基づいて、第1特定部分の動作を特定する動作特定手段と、

前記動作特定手段が特定した第1特定部分の動作から、撮影対象者のジェスチャの意味を特定する意味特定手段と、

前記意味特定手段が特定した意味を出力する出力手段と、
を備え、

前記意味特定手段が、第1特定部分の動作に基づいて撮影対象者のジェスチャの意味を特定できない場合、

前記動作特定手段は、前記区切り特定手段が特定した区切りに基づいた期間における、前記第2特徴点検出手段が検出した複数の特徴点に基づいて、第2特定部分の動作を特定し、

10

20

前記意味特定手段は、前記動作特定手段が特定した第1特定部分の動作及び第2特定部分の動作から撮影対象者のジェスチャの意味を特定する画像認識装置。

【請求項2】

前記意味特定手段が、第1特定部分の動作に基づいて撮影対象者のジェスチャの意味を特定できない場合、前記第2特徴点検出手段により第2特定部分の特徴点の検出処理を開始する、請求項1に記載の画像認識装置。

【請求項3】

撮影対象者の利き手を特定する利き手特定手段をさらに備え、

前記第1特徴点検出手段が、第1特定部分として指部分の特徴点を検出する場合、前記利き手特定手段が特定した利き手側の指部分の特徴点を検出し、

10

前記第2特徴点検出手段が、第2特定部分として指部分の特徴点を検出する場合、前記利き手特定手段が特定した利き手側の指部分の特徴点を検出する、請求項1又は2に記載の画像認識装置。

【請求項4】

前記第1特徴点検出手段又は前記第2特徴点検出手段が撮影対象者の腕部分の特徴点を検出する場合、

前記区切り特定手段は、腕部分の動きの有無に基づいて、撮影対象者のジェスチャの区切りを特定する、

請求項1～3の何れか一項に記載の画像認識装置。

【請求項5】

画像認識装置で実行する画像認識方法であって、

撮影対象者を画像データにして取得する画像取得ステップと、

前記画像取得ステップにより取得された画像データに基づいて、撮影対象者の第1特定部分の特徴点を複数検出する第1特徴点検出ステップと、

20

前記第1特徴点検出ステップで検出した特徴点に基づいて、撮影対象者のジェスチャの区切りを特定する区切り特定ステップと、

前記区切り特定ステップで特定した区切りに基づいた期間における、前記第1特徴点検出ステップで検出した特徴点に基づいて、第1特定部分の動作を特定する第1動作特定ステップと、

前記第1動作特定ステップで特定した第1特定部分の動作から、撮影対象者のジェスチャの意味を特定する第1意味特定ステップと、

30

前記第1意味特定ステップで、第1特定部分の動作に基づいて撮影対象者のジェスチャの意味を特定できない場合、前記画像取得ステップにより取得された画像データに基づいて、前記区切り特定ステップで特定した区切りに基づいた期間における撮影対象者の第1特定部分以外の部分である第2特定部分の特徴点を複数検出する第2特徴点検出ステップと、

前記区切り特定ステップで特定した区切りに基づいた期間における、前記第2特徴点検出ステップが検出した複数の特徴点に基づいて、第2特定部分の動作を特定する第2動作特定ステップと、

前記動作特定ステップで特定した第1特定部分の動作及び前記第2動作特定ステップで第2特定部分の動作から撮影対象者のジェスチャの意味を特定する第2意味特定ステップと、

40

前記第1意味特定ステップ又は前記第2意味特定ステップで特定した意味を出力する出力ステップと、

を備える、画像認識方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像認識装置、及び画像認識方法に関する。

【背景技術】

50

【0002】

従来における、手話などのジェスチャ認識の技術では、体全体、頭の位置、腕の動きだけでなく、指、口、顔の表情といった各々の動きを総合的に認識する。

【0003】

例えは、2つのカメラでユーザの画像を撮影し、それぞれの画像の画素値の差の相対的に大きい領域をユーザ領域として、ユーザ領域中の頂部を頭領域として、頂部よりも尖った領域を手先領域と特定して、手先の動きからジェスチャを判定する技術が知られている（特許文献1）。

【先行技術文献】**【特許文献】**

10

【0004】**【特許文献1】特開2009-211563号公報****【発明の概要】****【発明が解決しようとする課題】****【0005】**

しかし、特許文献1では、画像からユーザ領域を特定することができるが、全てのジェスチャに対して、体全体、頭の位置、腕、及び指の動きの検出や当該動きに対応する意味の特定した場合、処理負荷がかかるという問題点がある。

【0006】

本発明は、このような問題点を解決するためになされたものであり、ジェスチャ動作の意味を特定する処理負荷を軽減することができる画像認識装置及び画像認識方法を提供することを目的とする。

20

【課題を解決するための手段】**【0007】**

上述の課題を解決するために、本発明の画像認識装置は、撮影対象者を画像データにして取得する画像取得手段と、画像取得手段により取得された画像データに基づいて、撮影対象者の第1特定部分の特徴点を複数検出する第1特徴点検出手段と、画像取得手段により取得された画像データに基づいて、定期的に撮影対象者の第1特定部分以外の部分である第2特定部分の特徴点を複数検出する第2特徴点検出手段と、第1特徴点検出手段又は第2特徴点検出手段が検出した特徴点に基づいて、撮影対象者のジェスチャの区切りを特定する区切り特定手段と、区切り特定手段が特定した区切りに基づいた期間における、第1特徴点検出手段が検出した複数の特徴点に基づいて、第1特定部分の動作を特定する動作特定手段と、動作特定手段が特定した第1特定部分の動作から、撮影対象者のジェスチャの意味を特定する意味特定手段と、意味特定手段が特定した意味を出力する出力手段と、を備え、意味特定手段が、第1特定部分の動作に基づいて撮影対象者のジェスチャの意味を特定できない場合、動作特定手段は、区切り特定手段が特定した区切りに基づいた期間における、第2特徴点検出手段が検出した複数の特徴点に基づいて、第2特定部分の動作を特定し、意味特定手段は、動作特定手段が特定した第1特定部分の動作及び第2特定部分の動作から撮影対象者のジェスチャの意味を特定する。

30

【0008】

40

また、本発明に係る画像認識方法において、画像認識装置で実行する画像認識方法であつて、撮影対象者を画像データにして取得する画像取得ステップと、画像取得ステップにより取得された画像データに基づいて、撮影対象者の第1特定部分の特徴点を複数検出する第1特徴点検出ステップと、第1特徴点検出ステップ又は第2特徴点検出ステップで検出した特徴点に基づいて、撮影対象者のジェスチャの区切りを特定する区切り特定ステップと、区切り特定ステップで特定した区切りに基づいた期間における、第1特徴点検出ステップで検出した特徴点に基づいて、第1特定部分の動作を特定する第1動作特定ステップと、第1動作特定ステップで特定した第1特定部分の動作から、撮影対象者のジェスチャの意味を特定する第1意味特定ステップと、第1意味特定ステップで、第1特定部分の動作に基づいて撮影対象者のジェスチャの意味を特定できない場合、画像取得ステップに

50

より取得された画像データに基づいて、区切り特定ステップで特定した区切りに基づいた期間における撮影対象者の第1特定部分以外の部分である第2特定部分の特徴点を複数検出する第2特徴点検出ステップと、区切り特定ステップで特定した区切りに基づいた期間における、第2特徴点検出ステップで検出した複数の特徴点に基づいて、第2特定部分の動作を特定する第2動作特定ステップと、第1動作特定ステップで特定した第1特定部分の動作及び第2動作特定ステップで第2特定部分の動作から撮影対象者のジェスチャの意味を特定する第2意味特定ステップと、第1意味特定ステップ又は第2意味特定ステップで特定した意味を出力する出力ステップと、を備える。

【0009】

この発明によれば、第1特定部分の動作に基づいて、撮影対象者のジェスチャの意味を特定できない場合に限り、第2特定部分の動作を特定するので、撮影対象者の意味を特定する処理負荷を軽減させることができる。 10

【0010】

また、本発明の画像認識装置において、意味特定手段が、第1特定部分の動作に基づいて撮影対象者のジェスチャの意味を特定できない場合、第2特徴点検出手段により第2特定部分の特徴点の検出処理を開始するようにしてもよい。この場合、画像認識装置は、第2特徴点検出手段により検出された特徴点に基づいて動作を特定することを省略するだけでなく、第2特徴点の検出が必要になるタイミングまで第2特徴点の検出も停止するので、撮影対象者の意味を特定する処理負荷をより軽減させることができる。 20

【0011】

また、本発明の画像認識装置において、撮影対象者の利き手を特定する利き手特定手段をさらに備え、第1特徴点検出手段が、第1特定部分として指部分の特徴点を検出する場合、利き手特定手段が特定した利き手側の指部分の特徴点を検出し、第2特徴点検出手段が、第2特定部分として指部分の特徴点を検出する場合、利き手特定手段が特定した利き手側の指部分の特徴点を検出する、ようにしてもよい。この場合、画像認識装置は、指部分の特徴点を検出する場合に、利き手側の特徴点に基づいて動作を特定するので、検出する処理負荷を軽減させることができる。 20

【0012】

また、本発明の画像認識装置において、第1特徴点検出手段又は第2特徴点検出手段が撮影対象者の腕部分の特徴点を検出する場合、区切り特定手段は、腕部分の動きの有無に基づいて、撮影対象者のジェスチャの区切りを特定する、ようにしてもよい。この場合、画像認識装置は、腕部分の動きの有無に基づいて、ジェスチャの区切りを特定するので、ジェスチャの区切りを正確に特定することができる。 30

【発明の効果】

【0013】

本発明によれば、ジェスチャ動作の意味を判断する処理負荷を軽減することができる。

【図面の簡単な説明】

【0014】

【図1】本実施形態の画像認識装置10の機能構成を示すブロック図である。

【図2】画像認識装置10のハードウェア構成図である。 40

【図3】ジェスチャDB18及び文章解釈DB19のデータ例を示す図である。

【図4】動作とその意味の関係を示す図である。

【図5】本発明の実施形態に係る手話の意味特定処理を示すフローチャートである。

【図6】本発明の実施形態に係る手話の意味特定処理を示すフローチャートである。

【発明を実施するための形態】

【0015】

以下、図面を参照しながら、本実施形態を説明する。可能な場合には、同一の部分には同一の符号を付して、重複する説明を省略する。

【0016】

図1は、本実施形態の画像認識装置10の機能を示すブロック図である。この画像認識 50

装置 10 は、手話者などのジェスチャをテキストやイメージなどに翻訳するための装置であり、ジェスチャ入力部 11（画像取得手段）、画像記憶 DB（データベース）12、指特徴点検出部 13（第 1 特徴点検出手段、第 2 特徴点検出手段、利き手特定手段）、全体特徴点検出部 14（第 1 特徴点検出手段、第 2 特徴点検出手段）、区切り特定部 15（区切り特定手段）、動作特定部 16（動作特定手段）、意味特定部 17（意味特定手段）、ジェスチャ DB 18、文章解釈 DB 19 及び意味出力部 20（出力手段）を含んで構成されている。この画像認識装置 10 は、例えば、携帯端末やスマートフォンなどの携帯端末であることが好適である。

【 0 0 1 7 】

図 2 は、画像認識装置 10 のハードウェア構成図である。図 1 に示される画像認識装置 10 は、物理的には、図 2 に示すように、一または複数の CPU 101、主記憶装置である RAM 102 及び ROM 103、入力デバイスであるキーボード及びマウス等の入力装置 104、ディスプレイ等の出力装置 105、ネットワークカード等のデータ送受信デバイスである通信モジュール 106、半導体メモリ等の補助記憶装置 107などを含むコンピュータシステムとして構成されている。図 1 における各機能は、図 2 に示す CPU 101、RAM 102 等のハードウェア上に所定のコンピュータソフトウェアを読み込ませることにより、CPU 101 の制御のもとで入力装置 104、出力装置 105、通信モジュール 106 を動作させるとともに、RAM 102 や補助記憶装置 107 におけるデータの読み出し及び書き込みを行うことで実現される。以下、図 1 に示す機能ブロックに基づいて、各機能ブロックを説明する。

【 0 0 1 8 】

ジェスチャ入力部 11 は、手話者などの撮影対象者の動作を複数の画像データとして取得する部分であり、例えばカメラ（3 次元計測技術を備えたカメラ、簡易カメラ）である。なお、画像データには、撮影時刻が紐づけられている。ジェスチャ入力部 11 は、取得した画像データを画像記憶 DB 12 へ記憶する。

【 0 0 1 9 】

画像記憶 DB 12 は、ジェスチャ入力部 11 が取得した画像データを記憶するデータベースである。

【 0 0 2 0 】

指特徴点検出部 13 は、ジェスチャ入力部 11 により取得された画像データに基づいて、撮影対象者の指部分（第 1 特定部分又は第 2 特定部分）の特徴点を複数検出する部分である。具体的には、指特徴点検出部 13 は、各画像データ中における指部分の特徴点の位置を検出することにより、指部分の特徴点の位置を複数検出する。そして、指特徴点検出部 13 は、予め定められている利き手情報（例えば、画像認識装置 10 が保持している情報）に基づいて、利き手側の指部分の特徴点を検出する。なお、指特徴点検出部 13 は、画像記憶 DB 12 を介して画像データを取得してもよいし、ジェスチャ入力部 11 から直接画像データを取得してもよい。

【 0 0 2 1 】

全体特徴点検出部 14 は、ジェスチャ入力部 11 により取得された画像データに基づいて、撮影対象者の指以外の特徴点を複数検出する部分である。具体的には、全体特徴点検出部 14 は、各画像データ中における指以外の特徴点として、腕、顔、及び手等の部位（第 1 特定部分又は第 2 特定部分）の特徴点の位置を検出することにより、指以外の特徴点を複数検出する。本実施形態では、全体特徴点検出部 14 は、指以外の特徴点として、腕の位置を検出するものとする。なお、全体特徴点検出部 14 は、画像記憶 DB 12 を介して画像データを取得してもよいし、ジェスチャ入力部 11 から直接画像データを取得してもよい。

【 0 0 2 2 】

区切り特定部 15 は、指特徴点検出部 13 又は全体特徴点検出部 14 が検出した特徴点の位置に基づいて、撮影対象者のジェスチャの区切りを特定する部分である。具体的には、区切り特定部 15 は、指特徴点検出部 13 又は全体特徴点検出部 14 が複数の画像データ

10

20

30

40

50

タから検出した特徴点における、画像データの撮影時刻に基づいた特徴点（例えば、手首部分の特徴点）の位置の変化を検出する。区切り特定部15は、検出した位置の変化が無い又は、検出した位置の変化が予め定められている閾値より小さい箇所を撮影対象者のジェスチャの区切りであると特定する。

【0023】

動作特定部16は、区切り特定部15が特定した区切りに基づいた期間における、指特徴点検出部13又は全体特徴点検出部14が検出した特徴点に基づいて、検出部分（例えば、指、腕等）の動作を特定する部分である。具体的に、動作特定部16は、区切り特定部15が特定した期間に対応する複数の画像データ中から指特徴点検出部13又は全体特徴点検出部14が検出した特徴点の位置変化に基づいて、検出部分の動作を特定する。例えば、動作特定部16は、画像データの撮影時刻順に、区切り特定部15が特定した期間に対応する複数の画像データ中の指特徴点検出部13又は全体特徴点検出部14が検出した特徴点の位置変化として、変化度合いや変化方向を検出することにより、検出部分の動作を特定する。なお、上記区切り特定部15が特定した区切りに基づいた期間とは、ある区切りからその次の区切りまでの期間である。

【0024】

意味特定部17は、指特徴点検出部13又は全体特徴点検出部14の何れか1つが検出した特徴点に基づいて動作特定部16が特定した動作から、撮影対象者のジェスチャの意味を特定する。また、意味特定部17は、指特徴点検出部13又は全体特徴点検出部14の何れか1つが検出した特徴点に基づいた動作からジェスチャを特定した結果、ジェスチャの意味が複数である場合、指特徴点検出部13及び全体特徴点検出部14のそれぞれが検出した特徴点に基づいた動作（すなわち、指及び腕の動作）から、撮影対象者のジェスチャの意味を特定する。

【0025】

意味特定部17は、指特徴点検出部13又は全体特徴点検出部14の何れか1つが検出した特徴点に基づいて動作特定部16が特定した動作から、撮影対象者のジェスチャの意味を特定する際に、上記動作に対応する意味を、後述するジェスチャDB18及び文章解釈DB19を用いて検索することにより、撮影対象者のジェスチャの意味を特定する。

【0026】

また、意味特定部17は、指特徴点検出部13又は全体特徴点検出部14の何れか1つが検出した特徴点に基づいた動作からジェスチャを特定した結果、ジェスチャの意味が複数である場合、指特徴点検出部13及び全体特徴点検出部14のそれぞれが検出した特徴点に基づいた動作から、撮影対象者のジェスチャの意味を特定する際に、指特徴点検出部13及び全体特徴点検出部14のそれぞれが検出した特徴点に基づいた動作の意味を、ジェスチャDB18及び文章解釈DB19を用いて検索する。なお、意味特定部17は、最初に検索して取得した複数のジェスチャの意味をメモリ（例えば、RAM102）で保持しておき、当該複数のジェスチャの意味と、指特徴点検出部13及び全体特徴点検出部14のそれぞれが検出した特徴点に基づいた動作とに基づいて、撮影対象者のジェスチャの意味を特定するようにしてもよい。

【0027】

ジェスチャDB18は、手話者の動作のパターンを記述するデータベースである。これはジェスチャ認識のため的一般的なデータベースである。

【0028】

ジェスチャDB18のデータ例を図3(A)に示す。図3(A)は、指の動きに関するジェスチャのデータ例である。図3(A)に示すように、ジェスチャDB18は、「ジェスチャID」と、「手の向き」と、「手の角度」と、「指の突起有無」と、「手の動きの情報」とを少なくとも有する。図示していないが、ジェスチャDB18では、顔、腕の動きの情報もさらに有する。

【0029】

例えば、ジェスチャIDが「00000001」であるジェスチャは、手の向きが正面

10

20

30

40

50

であり、手の角度が 0 度（指先が上を向いた状態）であり、「指の突起有無」に基づき、親指のみ上がっていて、「動き」に基づき、手が止まっていることを示す。「指の突起有無」の情報では、「1」の値が割り当てられている指が上がっていることを示す。

【0030】

文章解釈 DB19 は、意味特定部 17 が、ジェスチャ DB18 を参照して、ある程度の動作を認識すると、その意味を解釈するためのデータベースである。ここには動作のパターンの識別子とその意味（意図）とが対応付けて記述されることになる。これも一般的なジェスチャ解析のためのデータベースである。

【0031】

文章解釈 DB19 が記憶しているデータ例を図 3 (B) に示す。図 3 (B) に示すように、文章解釈 DB19 は、「ジェスチャ ID」と「意味」とを有する。 10

【0032】

図 3 (B) の例では、ジェスチャ ID が「00000001」のジェスチャは、指文字で「あ」の意味であることを示す。また、ジェスチャ ID が「00002000」のジェスチャは、名詞の「スタイル」、「形」、「型」の意味があることを示す。すなわち、ジェスチャ ID が「00002000」のジェスチャは、複数の意味を有している。

【0033】

なお、文章解釈 DB19 で保持しているデータの「意味」は、指の動作のみに基づいて特定できる意味、「上体、腕、及び頭」の動作のみで特定できる意味、及び、指の動作と「上体、腕、及び頭」の動作との両方で特定できる意味がある。 20

【0034】

指の動作のみに基づいて特定できる意味、「上体、腕、及び頭」の動作のみに基づいて特定できる意味、及び指の動作と「上体、腕、及び頭」の動作との両方にに基づいて特定できる意味の例を図 4 に示す。

【0035】

図 4 に示すように、指の動作のみに基づいて特定できる意味として、五十音、アルファベット、数字等がある。また、「上体、腕、及び頭」の動作のみに基づいて特定できる意味として、体の特定の部位を示すものがある。そして、指の動作と「上体、腕、及び頭」の動作との両方にに基づいて特定できる意味として、約束（両手の小指を絡める）等がある。 30

【0036】

このように、指の動作だけで意味を特定できるものや、「上体、腕、及び頭」の動作だけで意味を特定できるものがあるので、指の動作と「上体、腕、及び頭」の動作との双方の動作を特定することなく、意味を特定することも可能である。この点に着目し、本実施形態では、最初に、指特徴点検出部 13 及び全体特徴点検出部 14 の一方から検出した特徴点に基づいて動作を特定し、その動作から 1 つの意味に特定できれば、その意味を出力し、上記動作に対応する複数の意味候補がある場合、指特徴点検出部 13 が検出した特徴点に基づいた動作、及び全体特徴点検出部 14 が検出した特徴点に基づいた動作から意味を特定する。

【0037】

意味出力部 20 は、意味特定部 17 が特定したジェスチャの意味を出力する部分である。具体的には、ディスプレイ等である。 40

【0038】

続いて、最初に全体特徴点検出部 14 が検出した特徴点に基づいた動作のみから意味を特定する場合の処理の説明を図 5 に示すフローチャートを用いて説明する。前提として、図 5 に示すフローチャートの処理が始まる前は、指特徴点検出部 13 は、停止しているものとする。図 5 に示すフローチャートの処理では、全体特徴点検出部 14 が第 1 特徴点検出手段に対応し、指特徴点検出部 13 が第 2 特徴点検出手段に対応する。

【0039】

撮影対象となる手話者の手話動作が、ジェスチャ入力部 11 により撮影され、撮影対象 50

の画像データが入力され、この画像データを画像記憶DB12へ記憶する(ステップS1)。全体特徴点検出部14は、画像データから腕に対応する特徴点の位置を検出する(ステップS2)。続いて、区切り特定部15は、直前に撮影された画像データ中の腕の特徴点の一又は複数の位置と比較して腕の特徴点位置の変化度合いを判定する(ステップS3)。

【0040】

区切り特定部15が特徴点位置の変化度合いを判定した結果、腕が下りた状態で、且つ腕の動きが一定時間停止していると判断した場合(ステップS4; YES)、区切り特定部15は、区切りを特定し、ステップS5へ移動する。

【0041】

区切り特定部15が特徴点位置の変化度合いを判定した結果、腕の動きが一定時間停止していない場合(ステップS4; NO)、ステップS1へ戻る。このように、区切り特定部15が、腕の動きが一定時間停止していると判断するまで、ステップS1～ステップS3を定期的に実行する。

【0042】

ステップS5において、動作特定部16は、区切り特定部15が特定した区切りに基づいた期間(区切り特定部15が特定した区切りとその直前の区切りの間)に対応する画像データに対して全体特徴点検出部14が検出した特徴点を用いて腕の動作を特定する(ステップS5)。そして、意味特定部17は、当該動作に基づいて意味を特定する(ステップS6)。

【0043】

上記期間における動作に基づいて意味特定部17が特定した意味が1つの場合(ステップS7; NO)、ステップS11へ移動する。意味特定部17が特定した意味が複数の場合(ステップS7; YES)、指特徴点検出部13による指部分の特徴点の検出処理を開始し、指特徴点検出部13は、上記期間に対応する画像データに対して、指部分の特徴点を検出する(ステップS8)。動作特定部16は、指特徴点検出部13が検出した特徴点を用いて指の動作を特定する(ステップS9)。そして、意味特定部17は、動作特定部16が特定した動作に基づいて動作の意味を特定し(ステップS10)、ステップS11へ移動する。

【0044】

ステップS11において、意味出力部20は、意味特定部17が特定した意味を出力し、処理を終了する。このように、画像認識装置10は、最初に全体特徴点検出部14が検出した特徴点を用いて腕の動作を特定し、当該腕の動作に基づいてジェスチャの意味を特定している。この結果、画像認識装置10は、腕の動作だけで意味を特定できる場合に、指特徴点検出部13が検出する特徴点に基づいて動作を特定する処理を省略しているので、ジェスチャの意味を特定する処理負荷を軽減することができる。

【0045】

続いて、最初に指特徴点検出部13が検出した特徴点に基づいた動作から意味を特定する場合の処理の説明を図6に示すフローチャートを用いて説明する。前提として、図6のフローチャートの処理前は、全体特徴点検出部14が停止しているものとする。図6に示すフローチャートの処理では、指特徴点検出部13が第1特徴点検出手段に対応し、全体特徴点検出部14が第2特徴点検出手段に対応する。

【0046】

撮影対象となる手話者の手話動作が、ジェスチャ入力部11により撮影され、撮影対象の画像データが入力され、この画像データを画像記憶DB12へ記憶する(ステップS21)。指特徴点検出部13は、画像データから腕に対応する特徴点の位置を検出する(ステップS22)。続いて、区切り特定部15は、直前に撮影された画像データ中の指の特徴位置と比較して指の特徴点位置の変化度合いを判定する(ステップS23)。

【0047】

区切り特定部15が特徴点の位置の変化度合いを判定した結果、指の動きが一定時間停

10

20

30

40

50

止していると判断した場合（ステップS24；YES）、区切り特定部15は、区切りを特定し、ステップS25へ移動する。

【0048】

区切り特定部15が特徴点の位置の変化度合いを判定した結果、指の動きが一定時間停止していない場合（ステップS24；NO）、ステップS21へ戻る。このように、区切り特定部15が、指の動きが一定時間停止していると判断するまで、ステップS21～ステップS23を定期的に実行する。

【0049】

ステップS25において、動作特定部16は、区切り特定部15が特定した区切りに基づいた期間（区切り特定部15が特定した区切りとその直前の区切りの間）に対応する画像データに対して指特徴点検出部13が検出した特徴点を用いて指の動作を特定する（ステップS25）。そして、意味特定部17は、当該動作に基づいて意味を特定する（ステップS26）。

10

【0050】

上記期間における動作に基づいて意味特定部17が特定した意味が1つの場合（ステップS27；NO）、ステップS31へ移動する。意味特定部17が特定した意味が複数の場合（ステップS27；YES）、全体特徴点検出部14による腕部分の特徴点の検出処理を開始し、全体特徴点検出部14は、上記期間に対応する画像データに対して、腕部分の特徴点を検出する（ステップS28）。動作特定部16は、全体特徴点検出部14が検出した特徴点を用いて腕の動作を特定する（ステップS29）。そして、意味特定部17は、動作特定部16が特定した動作に基づいて動作の意味を特定し（ステップS30）、ステップS31へ移動する。

20

【0051】

ステップS31において、意味出力部20は、意味特定部17が特定した意味を出力し、処理を終了する。このように、画像認識装置10は、最初に指特徴点検出部13が検出した特徴点を用いて指の動作を特定し、当該指の動作に基づいてジェスチャの意味を特定している。この結果、画像認識装置10は、指の動作だけで意味を特定できる場合に、全体特徴点検出部14が検出する特徴点に基づいて動作を特定する処理を省略しているので、ジェスチャの意味を特定する処理負荷を軽減することができる。

30

【0052】

つぎに、本実施形態における画像認識装置10の作用効果について説明する。

【0053】

本実施形態の画像認識装置10によれば、ジェスチャ入力部11は、撮影対象者を画像データにして取得する。指特徴点検出部13は、画像データに基づいて、撮影対象者の指部分の特徴点を複数検出し、全体特徴点検出部14は、画像データに基づいて、撮影対象者の腕部分の特徴点を複数検出し、区切り特定部15は、指特徴点検出部13検出した特徴点に基づいて、撮影対象者のジェスチャの区切りを特定する。動作特定部16は、区切り特定部15が特定した区切りに基づいた期間における、指特徴点検出部13が検出した特徴点に基づいて、指部分の動作を特定する。意味特定部17は、動作特定部16が特定した指部分の動作から、撮影対象者のジェスチャの意味を特定する。意味特定部17が、指部分の動作に基づいて撮影対象者のジェスチャの意味を特定できない場合（例えば、撮影対象者の意味が複数ある場合）、全体特徴点検出部14は、区切り特定部15が特定した区切りに基づいた期間に対応する画像データに基づいて特徴点を複数検出す。動作特定部16は、区切り特定部15が特定した区切りに基づいた期間における、全体特徴点検出部14が検出した複数の特徴点に基づいて、腕特定部分の動作を特定し、意味特定部17は、動作特定部16が特定した指部分の動作及び腕部分の動作から撮影対象者のジェスチャの意味を特定する。意味出力部20は、意味特定部17が特定した意味を出力する。

40

【0054】

このように、画像認識装置10は、指部分の動作に基づいて、撮影対象者のジェスチャの意味を特定できない場合に限り、腕部分の動作を特定するので、撮影対象者の意味を特

50

定する処理負荷を軽減させることができる。

【0055】

画像認識装置10では、意味特定部17が、指部分の動作に基づいて撮影対象者のジェスチャの意味を特定できない場合、全体特徴点検出部14による腕部分の特徴点の検出処理を開始する。この場合、画像認識装置10は、全体特徴点検出部14により検出された特徴点に基づいて動作を特定することを省略するだけでなく、腕部分の特徴点の検出が必要になるタイミングまで全体特徴点検出部14による特徴点の検出処理をしないので、撮影対象者の意味を特定する処理負荷をより軽減させることができる。

【0056】

画像認識装置10では、指特徴点検出部13は、撮影対象者の利き手を予め保持している情報に基づいて特定し、指特徴点検出部13が、利き手側の指部分の特徴点を検出する。この場合、指部分の特徴点を検出する場合に、利き手側の特徴点に基づいて指の動作を特定するので、両手の指の動作を特定する場合に比較して処理負荷を軽減させることができる。

10

【0057】

画像認識装置10では、区切り特定部15は、腕部分の動きの有無に基づいて、撮影対象者のジェスチャの区切りを特定する。ジェスチャの区切りの時点では、一般的に撮影対象者は、腕の動きを止めていることが多い（腕を下している場合は、より顕著である）。この傾向に基づいて、区切り特定部15は、腕部分の動きの有無に基づいて、ジェスチャの区切りを特定するので、正確にジェスチャの区切りを特定することができる。

20

【0058】

上述の実施形態では、最初に指特徴点検出部13が特徴点を検出する場合には、全体特徴点検出部14を停止させておく場合について述べたが、全体特徴点検出部14を予め起動させておいてもよい。この場合でも、意味特定部17が複数の意味を特定しない限り、全体特徴点検出部14が検出した特徴点に基づいて動作を特定しないので、処理負荷を軽減させることができる。

【0059】

上述の実施形態では、利き手に関する情報を予め保持し、指特徴点検出部13は、その保持している情報を用いて利き手を特定する場合について述べたが、過去のジェスチャの画像データを解析して、動作が大きい方の手を利き手と特定するようにしてもよい。これは、撮影対象者が利き手を優先して動かす傾向にあることに基づく。

30

【0060】

上述の実施形態では、ジェスチャ入力部11がカメラである場合について述べたが、温度センサに基づいた画像データを生成する装置によりジェスチャ入力部11を実現するようにしてもよい。

【0061】

上述の実施形態では、意味特定部17が、指部分の動作に基づいて撮影対象者のジェスチャの意味を特定した結果、撮影対象者のジェスチャの意味が複数である場合、動作特定部16が特定した指部分の動作及び腕部分の動作から撮影対象者のジェスチャの意味を特定する場合について述べたが、意味特定部17は、腕部分の動作から撮影対象者のジェスチャの意味を特定するようにしてもよい。すなわち、第1特定部分の動作に基づいて撮影対象者のジェスチャの意味を特定した結果、撮影対象者のジェスチャの意味が複数である場合、第2特定部分の動作から撮影対象者のジェスチャの意味を特定するようにしてもよい。

40

【0062】

上述の実施形態では、意味特定部17が、腕部分の動作に基づいて撮影対象者のジェスチャの意味を特定した結果、撮影対象者のジェスチャの意味が複数である場合、動作特定部16が指部分の動作を特定し、動作特定部16が特定した指部分の動作及び腕部分の動作から撮影対象者のジェスチャの意味を特定することについて述べたが、区切りに基づいた期間中の腕部分の動作が、腕を下した状態から変化が無い場合には、動作特定部16が

50

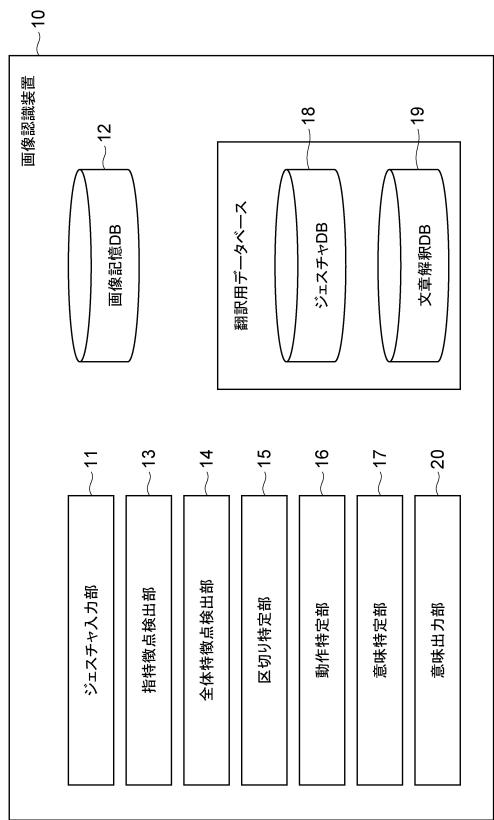
指部分の動作を特定する処理及び意味特定部 17 が指部分の動作及び腕部分の動作から撮影対象者のジェスチャの意味を特定する処理をせずに全体処理を終了してもよい。この場合、腕を下した状態では指を動作させている可能性が低いので不要な処理を省略することができる。

【符号の説明】

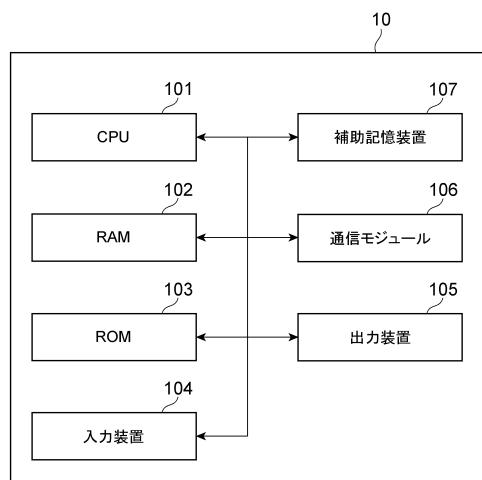
【0063】

10...画像認識装置、11...ジェスチャ入力部、12...画像記憶DB、13...指特徴点検出部、14...全体特徴点検出部、15...区切り特定部、16...動作特定部、17...意味特定部、18...ジェスチャDB、19...文章解釈DB、20...意味出力部。

【図1】



【図2】



【図3】

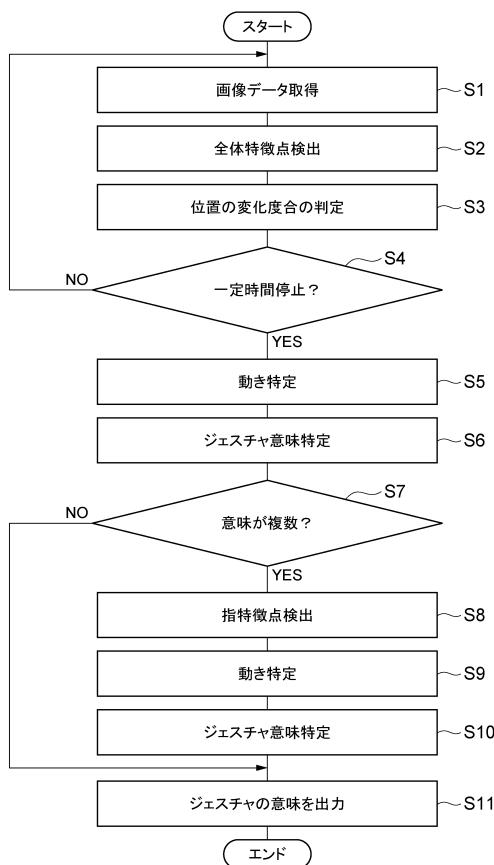
18			
ジェスチャID	手の向き	手の角度	指の突起有無(親指/…/小指)
000000001	正面	0度	1/0/0/0/0
000000002	正面	0度	0/0/0/0/1
...
000000014	正面	180度	1/1/1/0/0
...
00002000	正面	0度	1/1/0/0/0
			上下

19			
ジェスチャID	意味(カンマで区切られた場合は複数の意味が存在)		
00000001	(指文字)「あ」		
00000002	(指文字)「い」		
00000014	(指文字)「す」		
00002000	(名詞)「スタイル(すたいる)」「形(かたち)」「型(かた)」		

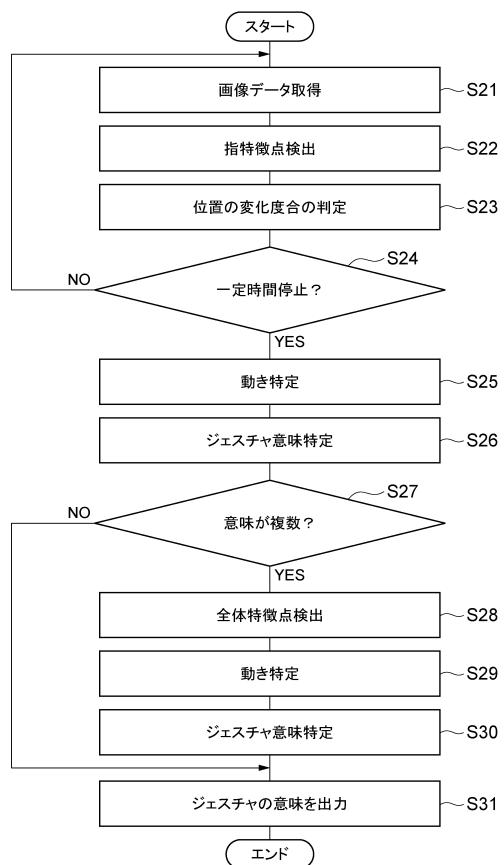
【図4】

指の動作のみで 特定できる意味(例)	上体、腕、頭の動作のみで 特定できる意味(例)	上体、腕、頭の動作と指の動作両方で 特定できる意味(例)
指文字 ・五十音 ・アルファベット ・数字 (名前、助詞は必ず指文字)	体の特定部位の前で腕(手)を回す動作 ・目の周り:目 ・口の周り:口 ・頭の周り:顔 ・腕の周り片手で回す:身体 ・両手を回す:轟らす ・親指を立てながら実施するが、 両手を回す動作が単独のため	回す動作に指の形もある場合 ・人差し指、中指を立てて(1と)の指文字 ・人差し指、小指を立てて(「き」)の指文字 ・人差し指、中指を立てて回す:人生 ・親指、小指を立てて回す:人生 ・人差し指、中指をカギ型に曲げて (1)の指文字(頭の当たりで回す)・ロマンティック
人間関係 親指:男性 人差し指:人(私、あなた) 中指:兄弟 薬指:姉妹 小指:女性	手、腕の動きで文字を書く動作 ・人:人 ・[m]:メートル(単位) ・[W]:ワイルズ	顔から両手が離れていく動作 ・両手の人差し指、中指を立てて ・両手の人差し指立てて: ・両手の起点に両側へ:榮誉 ・指を一本ずつ折りながら右へ移動:多い
3	・鼻の前の手を前に出す:良い	・人差し指、小指を立てて鼻の前から 前に出す:自慢
4	一般的な動作 ・ツバをかき混ぜる動作:コーヒー ・ラケットを振る動作:テニス ・靴に刀をおさめる動作:武士 ・ハンドルを握る動作:車	一般的な動作 ・両手の小指を絡める:結束

【図5】



【図6】



フロントページの続き

(72)発明者 渋谷 彰

東京都千代田区永田町二丁目11番1号 株式会社NTTドコモ内

(72)発明者 高橋 誠

東京都千代田区永田町二丁目11番1号 株式会社NTTドコモ内

(72)発明者 樋口 雄太

東京都千代田区永田町二丁目11番1号 株式会社NTTドコモ内

審査官 加内 慎也

(56)参考文献 特開平6-138815 (JP, A)

特開2003-50663 (JP, A)

特開平8-115408 (JP, A)

特開2001-286677 (JP, A)

特開2000-172163 (JP, A)

特開2001-56861 (JP, A)

特開2013-235582 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G 06 F 3 / 01

G 06 T 7 / 20

G 09 B 21 / 00