

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6098498号  
(P6098498)

(45) 発行日 平成29年3月22日(2017.3.22)

(24) 登録日 平成29年3月3日(2017.3.3)

(51) Int.Cl.

F I

G O 6 T 1/00 (2006.01)

G O 6 T 1/00 3 1 5

G O 6 T 7/60 (2017.01)

G O 6 T 7/60 1 5 O B

G O 6 F 3/0346 (2013.01)

G O 6 F 3/0346 4 2 2

請求項の数 11 (全 34 頁)

(21) 出願番号 特願2013-262010 (P2013-262010)  
 (22) 出願日 平成25年12月19日(2013.12.19)  
 (65) 公開番号 特開2015-118582 (P2015-118582A)  
 (43) 公開日 平成27年6月25日(2015.6.25)  
 審査請求日 平成27年12月14日(2015.12.14)

(73) 特許権者 000002185  
 ソニー株式会社  
 東京都港区港南1丁目7番1号  
 (74) 代理人 100093241  
 弁理士 宮田 正昭  
 (74) 代理人 100101801  
 弁理士 山田 英治  
 (74) 代理人 100086531  
 弁理士 澤田 俊夫  
 (74) 代理人 100095496  
 弁理士 佐々木 榮二  
 (74) 代理人 110000763  
 特許業務法人大同特許事務所

最終頁に続く

(54) 【発明の名称】 情報処理装置、情報処理方法、並びにプログラム

(57) 【特許請求の範囲】

【請求項 1】

異なる視点から撮影した画像を連結して生成した連結画像から特徴量を抽出する特徴量抽出部と、

前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの位置判定処理を実行する特定オブジェクト認識部を有し、

前記特徴量抽出部は、フィルタ構成領域を離間して設定した離間型フィルタを適用した特徴量抽出処理を実行し、

前記離間型フィルタを構成する離間した各フィルタ領域を、前記連結画像の2つの画像の各々に設定して特徴量抽出処理を実行する情報処理装置。

【請求項 2】

前記特徴量抽出部は、

予め実行した学習処理において前記特定オブジェクトの第1状態と第2状態との識別力が高いと判定したフィルタを選択的に適用して特徴量抽出を実行する請求項1に記載の情報処理装置。

【請求項 3】

前記情報処理装置は、

前記特定オブジェクトの第1状態を示す連結画像と、第2状態を示す連結画像を多数入力して、異なる形式の離間型フィルタを適用した特徴量抽出により、識別力の高いフィルタを判別する処理を実行する学習処理部を有する請求項1または2に記載の情報処理装置

。

【請求項 4】

前記特定オブジェクト認識部は、

前記特徴量抽出部の抽出した異なる離間型フィルタ対応の特徴量を加算したスコアを算出し、加算スコアに基づいて前記特定オブジェクトの位置判定処理を実行する請求項 1 ~ 3いずれかに記載の情報処理装置。

【請求項 5】

前記特定オブジェクト認識部は、

前記特徴量抽出部の抽出した異なる離間型フィルタ対応の画素対応の特徴量を加算した画素対応のスコアを算出し、算出した画素対応のスコアをさらに加算した総合スコアに基づいて前記特定オブジェクトの位置判定処理を実行する請求項 1 ~ 4いずれかに記載の情報処理装置。

10

【請求項 6】

前記特徴量抽出部は、

前記特定オブジェクトである指がボードに接触している第 1 状態と、指がボードに接触していない第 2 状態を判別する指標値としての特徴量を算出し、

前記特定オブジェクト認識部は、

前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトである指がボードに接触しているか否かの判定処理を実行する請求項 1 ~ 5いずれかに記載の情報処理装置。

【請求項 7】

前記特徴量抽出部は、さらに、

画像内に特定オブジェクトが存在するか否かを判別する指標値としての特徴量を抽出し、

20

前記特定オブジェクト認識部は、

前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの有無を判定する特定オブジェクト検出処理を実行する請求項 1 ~ 6いずれかに記載の情報処理装置。

【請求項 8】

前記離間型フィルタは、異なる輝度のフィルタ領域を離間して設定したフィルタである請求項 1 ~ 7いずれかに記載の情報処理装置。

【請求項 9】

前記特徴量抽出部は、

前記連結画像の一方の画像の特徴点と、他方の画像上の前記特徴点に対応するエッジライン上に前記離間型フィルタを設定して特徴量抽出を実行する請求項 1 ~ 8いずれかに記載の情報処理装置。

30

【請求項 10】

情報処理装置において実行する情報処理方法であり、

特徴量抽出部が、異なる視点から撮影した画像を連結して生成した連結画像から特徴量を抽出する特徴量抽出処理、

特定オブジェクト認識部が、前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの位置判定処理を実行する特定オブジェクト認識処理を実行し、

40

前記特徴量抽出部は、フィルタ構成領域を離間して設定した離間型フィルタを適用した特徴量抽出処理を実行し、

前記離間型フィルタを構成する離間した各フィルタ領域を、前記連結画像の 2 つの画像の各々に設定して特徴量抽出処理を実行する情報処理方法。

【請求項 11】

情報処理装置において情報処理を実行させるプログラムであり、

特徴量抽出部に、異なる視点から撮影した画像を連結して生成した連結画像から特徴量を抽出させる特徴量抽出処理、

特定オブジェクト認識部に、前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの位置判定処理を実行する特定オブジェクト認識処理を実行させ、

50

前記特徴量抽出部には、フィルタ構成領域を離間して設定した離間型フィルタを適用した特徴量抽出処理を実行させ、

前記離間型フィルタを構成する離間した各フィルタ領域を、前記連結画像の２つの画像の各々に設定して特徴量抽出処理を実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【０００１】

本開示は、情報処理装置、情報処理方法、並びにプログラムに関する。さらに詳細にはカメラの撮影画像に基づいて特定オブジェクトの有無や位置を解析する情報処理装置、情報処理方法、並びにプログラムに関する。

【背景技術】

【０００２】

画像に撮影された被写体の位置を解析するシステムとしてステレオカメラを用いた距離解析処理が知られている。

これは、２つの異なる位置から撮影した２枚の画像の対応点を解析し、この対応点情報に基づいてカメラから被写体までの距離（デプス）を算出するものである。

しかし、２枚の画像の対応点に基づく距離算出処理は、計算コストや処理コストが大きくなり、データ処理能力の十分な装置を用いないと処理遅延が発生する場合がある。

【０００３】

また、画像から特定のオブジェクト、例えば人の顔や手を検出する処理として特徴量を用いた処理が知られている。これは、検出対象の特徴量を示す画像データを予め取得し、撮影画像の中から、この特徴量と類似する画像領域を検出する処理である。例えば人の顔を検出する場合は、人の顔の特徴量を示す画像データを予め取得し、撮影画像の中から、この特徴量と類似する特徴量を持つ画像領域を顔領域として判定する。

【０００４】

特徴量としては、例えば画像内のエッジ抽出によって得られる特徴量が用いられる。予め実行する学習処理によって、複数の特徴量データを集積し、集積した学習データと撮影画像に含まれる画像領域との比較により目的のオブジェクトの検出が実行される。（特許文献１～３参照）

【０００５】

しかし、従来技術において開示された特徴量を用いたオブジェクト検出処理は、例えば１枚の画像の中から、特定の特徴量と類似する領域を検出して特定のオブジェクトの画像領域を特定することは可能であるが、そのオブジェクトの空間的な位置を判定するものとはなっていない。

【先行技術文献】

【特許文献】

【０００６】

【特許文献１】ＵＳ公開２００４－００１３３０４

【特許文献２】ＵＳ公開２００２－０１０２０３４

【特許文献３】ＵＳ特許７０９９５１０

【発明の概要】

【発明が解決しようとする課題】

【０００７】

本開示は、カメラの撮影画像に基づいて例えば人の指など特定オブジェクトが画像に含まれるか否かの判定処理や、特定オブジェクトの位置判定を行う情報処理装置、情報処理方法、並びにプログラムを提供することを目的とする。

【課題を解決するための手段】

【０００８】

本開示の第１の側面は、

異なる視点から撮影した画像を連結して生成した連結画像から特徴量を抽出する特徴量

10

20

30

40

50

抽出部と、

前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの位置判定処理を実行する特定オブジェクト認識部を有し、

前記特徴量抽出部は、フィルタ構成領域を離間して設定した離間型フィルタを適用した特徴量抽出処理を実行する情報処理装置にある。

【0009】

さらに、本開示の情報処理装置の一実施態様において、前記特徴量抽出部は、前記離間型フィルタを構成する離間した各フィルタ領域を、前記連結画像の2つの画像の各々に設定して特徴量抽出処理を実行する。

【0010】

さらに、本開示の情報処理装置の一実施態様において、前記特徴量抽出部は、予め実行した学習処理において前記特定オブジェクトの第1状態と第2状態との識別力が高いと判定したフィルタを選択的に適用して特徴量抽出を実行する。

【0011】

さらに、本開示の情報処理装置の一実施態様において、前記情報処理装置は、前記特定オブジェクトの第1状態を示す連結画像と、第2状態を示す連結画像を多数入力して、異なる形式の離間型フィルタを適用した特徴量抽出により、識別力の高いフィルタを判別する処理を実行する学習処理部を有する。

【0012】

さらに、本開示の情報処理装置の一実施態様において、前記特定オブジェクト認識部は、前記特徴量抽出部の抽出した異なる離間型フィルタ対応の特徴量を加算したスコアを算出し、加算スコアに基づいて前記特定オブジェクトの位置判定処理を実行する。

【0013】

さらに、本開示の情報処理装置の一実施態様において、前記特定オブジェクト認識部は、前記特徴量抽出部の抽出した異なる離間型フィルタ対応の画素対応の特徴量を加算した画素対応のスコアを算出し、算出した画素対応のスコアをさらに加算した総合スコアに基づいて前記特定オブジェクトの位置判定処理を実行する。

【0014】

さらに、本開示の情報処理装置の一実施態様において、前記特徴量抽出部は、前記特定オブジェクトである指がボードに接触している第1状態と、指がボードに接触していない第2状態を判別する指標値としての特徴量を算出し、前記特定オブジェクト認識部は、前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトである指がボードに接触しているか否かの判定処理を実行する。

【0015】

さらに、本開示の情報処理装置の一実施態様において、前記特徴量抽出部は、さらに、画像内に特定オブジェクトが存在するか否かを判別する指標値としての特徴量を抽出し、前記特定オブジェクト認識部は、前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの有無を判定する特定オブジェクト検出処理を実行する。

【0016】

さらに、本開示の情報処理装置の一実施態様において、前記離間型フィルタは、異なる輝度のフィルタ領域を離間して設定したフィルタである。

【0017】

さらに、本開示の情報処理装置の一実施態様において、前記特徴量抽出部は、前記連結画像の一方の画像の特徴点と、他方の画像上の前記特徴点に対応するエピポーラライン上に前記離間型フィルタを設定して特徴量抽出を実行する。

【0018】

さらに、本開示の第2の側面は、

情報処理装置において実行する情報処理方法であり、

特徴量抽出部が、異なる視点から撮影した画像を連結して生成した連結画像から特徴量を抽出する特徴量抽出処理、

10

20

30

40

50

特定オブジェクト認識部が、前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの位置判定処理を実行する特定オブジェクト認識処理を実行し、

前記特徴量抽出部は、フィルタ構成領域を離間して設定した離間型フィルタを適用した特徴量抽出処理を実行する情報処理方法にある。

【0019】

さらに、本開示の第3の側面は、

情報処理装置において情報処理を実行させるプログラムであり、

特徴量抽出部に、異なる視点から撮影した画像を連結して生成した連結画像から特徴量を抽出させる特徴量抽出処理、

特定オブジェクト認識部に、前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの位置判定処理を実行する特定オブジェクト認識処理を実行させ、

前記特徴量抽出部には、フィルタ構成領域を離間して設定した離間型フィルタを適用した特徴量抽出処理を実行させるプログラムにある。

【0020】

なお、本開示のプログラムは、例えば、様々なプログラム・コードを実行可能な情報処理装置やコンピュータ・システムに対して、コンピュータ可読な形式で提供する記憶媒体、通信媒体によって提供可能なプログラムである。このようなプログラムをコンピュータ可読な形式で提供することにより、情報処理装置やコンピュータ・システム上でプログラムに応じた処理が実現される。

【0021】

本開示のさらに他の目的、特徴や利点は、後述する本開示の実施例や添付する図面に基づくより詳細な説明によって明らかになるであろう。なお、本明細書においてシステムとは、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

【発明の効果】

【0022】

本開示の一実施例の構成によれば、特定オブジェクトの位置判定を異なる視点からの撮影画像の連結画像から抽出した特徴量に基づいて実行する装置、方法が実現される。

具体的には、異なる視点からの撮影画像を連結した連結画像から特徴量を抽出し、抽出した特徴量に基づいて特定オブジェクトの位置判定処理を実行する。特徴量抽出部は、フィルタ構成領域を離間して設定した離間型フィルタを適用した特徴量抽出処理を実行し、離間型フィルタの各フィルタ領域を、連結画像の2つの画像の各々に設定して特徴量抽出処理を実行する。特定オブジェクト認識部は、特徴量抽出部の抽出した異なる離間型フィルタ対応の特徴量を加算したスコアに基づいて特定オブジェクトの位置判定処理を実行する。

本構成により、特定オブジェクトの位置判定を異なる視点からの撮影画像の連結画像から抽出した特徴量に基づいて実行する装置、方法が実現される。

なお、本明細書に記載された効果はあくまで例示であって限定されるものではなく、また付加的な効果があってもよい。

【図面の簡単な説明】

【0023】

【図1】本開示の情報処理装置が実行する処理の概要について説明する図である。

【図2】本開示の情報処理装置が実行する処理のシーケンスについて説明するフローチャートを示す図である。

【図3】本開示の情報処理装置の構成と処理について説明する図である。

【図4】特徴量抽出に適用するフィルタについて説明する図である。

【図5】特徴量抽出に適用するフィルタについて説明する図である。

【図6】特徴量抽出処理について説明する図である。

【図7】スコア算出処理について説明する図である。

【図8】オブジェクト検出のための学習処理について説明する図である。

【図 9】オブジェクト検出のための学習処理について説明する図である。

【図 10】オブジェクト検出のための学習処理について説明する図である。

【図 11】特徴量抽出処理について説明する図である。

【図 12】特徴量抽出処理について説明する図である。

【図 13】総合スコア算出処理について説明する図である。

【図 14】特定オブジェクトの位置判定処理に適用する画像とフィルタについて説明する図である。

【図 15】画像の切り出し処理と L R 連結画像の生成処理について説明する図である。

【図 16】L R 連結画像の例について説明する図である。

【図 17】オブジェクト位置判定のための学習処理について説明する図である。

10

【図 18】オブジェクト位置判定のための学習処理について説明する図である。

【図 19】オブジェクト位置判定のための学習処理について説明する図である。

【図 20】ボードに対する指が接触した画像と比較画像の差異について説明する図である。

【図 21】エポポーララインを適用したフィルタ適用領域の限定について説明する図である。

【図 22】離間型フィルタの種類と特徴について説明する図である。

【図 23】情報処理装置の構成例について説明する図である。

【発明を実施するための形態】

【0024】

20

以下、図面を参照しながら本開示の情報処理装置、および情報処理方法、並びにプログラムの詳細について説明する。なお、説明は以下の項目に従って行なう。

1. 本開示の情報処理装置の実行する処理の概要について
2. 情報処理装置の構成と処理について
3. 情報処理装置の各構成部の実行する処理について
  - 3-1. 学習処理部の特徴量抽出部が実行する処理について
  - 3-2. 機械学習部の実行する処理について
  - 3-3. 認識処理部の実行する処理について
4. 特定オブジェクトの位置判定処理について
  - 4-1. 学習処理部の実行する処理について
  - 4-2. 認識処理部の実行する処理について
5. フィルタ適用位置を限定した実施例について
6. 離間型フィルタのタイプと特性について
7. 情報処理装置の構成例について
8. 本開示の構成のまとめ

30

【0025】

[ 1. 本開示の情報処理装置の実行する処理の概要について ]

まず、図 1 以下を参照して本開示の情報処理装置が実行する処理の概要について説明する。

本開示の情報処理装置は、複数の異なる位置に設置されたカメラの撮影画像を入力し、カメラ撮影画像の解析に基づいて、以下の各処理を実行する。

40

( a ) 撮影画像内に特定のオブジェクトが含まれるか否かを判定する「特定オブジェクト検出処理」

( b ) 上記処理 ( a ) で検出された特定オブジェクトについての位置を判定する「特定オブジェクトの位置判定処理」

これら ( a ) , ( b ) の各処理を実行する。

【0026】

なお、以下の実施例では、検出処理、および位置判定の対象とする特定オブジェクトを図 1 ( a ) に示す人の指 30 とした実施例を説明する。

また、特定オブジェクトの位置判定処理は、指 30 がボード 21 に接触 ( タッチ ) して

50

いるか否かの２つの状態の判定処理を行なう実施例について説明する。

【００２７】

図１（ａ）は、画像の撮影環境の一例を示す図である。

図１（ａ）は、ボード２１上に人の指３０があり、指３０がボード２１に接触（タッチ）している状態を示している。

なお、指３０は、随時、移動し、ボード２１に接触していない離れた状態になる場合や、ボード２１から離れて、ボード２１上には存在しなくなる場合もある。

すなわち、指３０の状態としては、以下の３種類の状態がある。

（状態１）指３０がボード２１に接触した状態

（状態２）指３０がボード２１の上にあるが接触していない状態

10

（状態３）指３０がボード２１に接触せず、ボード２１の上部にもない状態

【００２８】

本開示の情報処理装置は、図１に示す２つのカメラ、すなわちカメラＬ１１とカメラＲ１２の撮影画像に基づいて、上記の３状態を判別する。

【００２９】

カメラＬ１１はボード２１の中心部から見て左上部の位置に固定され、ボード２１の平面を左上部から撮影する。

一方、カメラＲ１２はボード２１の中心部から見て右上部の位置に固定され、ボード２１の平面を右上部から撮影する。

ボード２１上に指３０が存在する場合は、指３０を含む画像をそれぞれ左上部、および右上部から撮影する。

20

【００３０】

各カメラの撮影画像の例を図１（ｂ）に示す。

Ｌ画像５１は、カメラＬ１１によって撮影された画像である。

Ｒ画像５２は、カメラＲ１２によって撮影された画像である。

この画像例は、指３０がボード２１に接触している状態における撮影画像である。

指３０が、ボード２１に接触していない場合や、指３０がボード２１の上に無い場合には、図１（ｂ）に示す画像とは異なる画像が撮影されることになる。

【００３１】

なお、カメラＬ１１、カメラ１２の撮影画像は静止画でも、動画でもよい。静止画である場合は、撮影された２枚の静止画を利用して上記の（状態１）～（状態３）のいずれの状態にあるかを判定する。

30

また、撮影画像が動画である場合は、例えば各撮影フレーム単位、あるいは、予め規定した所定フレーム間隔で上記の（状態１）～（状態３）のいずれの状態にあるかを判定する。

【００３２】

次に、図２に示すフローチャートを参照して本開示の情報処理装置の実行する処理シーケンスについて説明する。

本開示の情報処理装置は、例えば図１に示す２つの異なる位置に設定されたカメラであるカメラＬ１１、カメラＲ１２の撮影画像を入力して、これらの画像に基づいて、上述した３つの状態、すなわち、

40

（状態１）指３０がボード２１に接触した状態

（状態２）指３０がボード２１の上にあるが接触していない状態

（状態３）指３０がボード２１に接触も上部にもない状態

上記３状態のいずれの状態であるかを判定する処理を実行する。

【００３３】

図２に示すフローの各ステップの処理について説明する。

（ステップＳ１０１）

まず、情報処理装置は、撮影画像に指が含まれるか否かの指検出処理を実行する。この処理は、図１（ｂ）に示すＬ画像５１、またはＲ画像５２のいずれか一方の画像のみを利

50

用して実行する。

この指検出処理には、予め蓄積した特徴量を示す学習データを利用する。この処理の具体的処理については、後段で説明する。

【 0 0 3 4 】

(ステップ S 1 0 2 ~ S 1 0 3 )

情報処理装置は、ステップ S 1 0 2 において、撮影画像から指が検出されたか否かを判定する。

指が検出されなかった場合は、ステップ S 1 0 3 に進み、指が検出されなかったことを示す解析結果を出力し、ステップ S 1 0 8 に進む。

なお、解析結果の出力先は、解析結果に応じた処理を実行するデータ処理部であり、情報処理装置内のデータ処理部でもよいし、情報処理装置に接続された外部装置であってもよい。

【 0 0 3 5 】

(ステップ S 1 0 4 )

ステップ S 1 0 2 において、撮影画像から指が検出されたと判定した場合は、ステップ S 1 0 4 に進む。

ステップ S 1 0 4 では、指の位置検出処理として、指がボードに接触 (タッチ) しているか否かの接触判定処理を実行する。

この接触判定処理には、図 1 ( b ) に示す L 画像 5 1 と、R 画像 5 2 の異なる視点から撮影した複数枚の画像を利用する。

この接触判定処理においても、予め蓄積した特徴量を示す学習データを利用する。この処理の具体的処理については、後段で説明する。

【 0 0 3 6 】

(ステップ S 1 0 5 ~ S 1 0 7 )

情報処理装置は、ステップ S 1 0 5 において、撮影画像から検出した指がボードに接触した状態であるか否かを判定し、接触していないと判定した場合は、ステップ S 1 0 6 に進み、指がボードに接触していない状態であることを示す解析結果を出力し、ステップ S 1 0 8 に進む。

一方、接触していると判定した場合は、ステップ S 1 0 7 に進み、指がボードに接触している状態であることを示す解析結果を出力し、ステップ S 1 0 8 に進む。

なお、解析結果の出力先は、解析結果に応じた処理を実行するデータ処理部であり、情報処理装置内のデータ処理部でもよいし、情報処理装置に接続された外部装置であってもよい。

【 0 0 3 7 】

(ステップ S 1 0 8 )

次に、情報処理装置は、ステップ S 1 0 8 において、解析処理が終了したか否かを判定する。解析対象となる画像の入力がない場合には、処理終了と判定する。例えば、動画を撮影しながら予め規定したフレーム間隔ごとに解析を継続的に実行する場合には、図 2 に示すフローに従った処理を所定フレーム間隔ごとに繰り返し実行する。この場合には、ステップ S 1 0 8 の判定は N o となり、ステップ S 1 0 1 に戻り、新たな入力画像に対して、ステップ S 1 0 1 以下の解析処理を繰り返し実行する。

【 0 0 3 8 】

このように、本開示の情報処理装置は、1枚の撮影画像に基づくオブジェクト検出処理を実行し、続いて異なる視点から撮影した画像を適用して、オブジェクト位置判定処理を実行する。

なお、ここで説明する実施例では、オブジェクト検出処理は指の検出処理であり、オブジェクト位置判定処理は、指がボードに接触しているか否かの2つの状態の判別処理として実行する。

【 0 0 3 9 】

[ 2 . 情報処理装置の構成と処理について ]

10

20

30

40

50



次に本開示の情報処理装置の構成例について説明する。

本開示の情報処理装置の構成例を図3に示す。情報処理装置100は、学習処理部110と認識処理部120を有する。

学習処理部110は、画像入力部111、特徴量抽出部112、機械学習部113、認識辞書記憶部114を有する。

学習処理部110は、多数の画像を学習画像として入力し、学習画像に基づいて、認識処理部120が画像から特定オブジェクトの有無や位置の識別処理に用いる認識辞書を生成する。生成した認識辞書は認識辞書記憶部114に格納される。

#### 【0040】

一方、認識処理部120は、画像入力部121、特徴量抽出部122、特定オブジェクト認識部123を有する。

認識処理部120は、解析対象画像を入力して、特定オブジェクトの検出と位置判定を行なう。認識処理部120は、学習処理部110が学習データに基づいて生成し、認識辞書記憶部114に格納した認識辞書を用いて、特定オブジェクトの検出と位置判定を行なう。

図2を参照して説明したフローチャートに従った処理は、この認識処理部120の実行する処理に相当する。

#### 【0041】

なお、図3に示す情報処理装置100は、先に図2に示すフローを参照して説明したように、以下の2つの処理を行なう。

(処理1) 画像に指が含まれるか否かのオブジェクト検出処理、

(処理2) 検出された指がボードに接触しているか否かを判定するオブジェクト位置判定処理(接触判定処理)

#### 【0042】

情報処理装置100の学習処理部110は、上記の(処理1)、(処理2)に適用する学習データとしての認識辞書を生成する。

また、認識処理部120は、学習処理部の生成した上記の(処理1)、(処理2)に対応する認識辞書を個別に適用して、以下の各処理、すなわち、

(処理1) 画像に指が含まれるか否かのオブジェクト検出処理、

(処理2) 検出された指がボードに接触しているか否かを判定するオブジェクト位置判定処理(接触判定処理)

これらの各処理を実行する。

#### 【0043】

[3. 情報処理装置の各構成部の実行する処理について]

以下、図3に示す情報処理装置100の各処理部の実行する処理について、順次、説明する。

#### 【0044】

[3-1. 学習処理部の特徴量抽出部が実行する処理について]

まず、図3に示す情報処理装置100の学習処理部110内の特徴量抽出部112が実行する処理について説明する。

学習処理部110は、画像入力部111から多数の撮影画像を学習画像として入力する。

学習画像は、例えば、図1(a)を参照して説明した撮影環境において撮影された多数の画像である。この学習画像には、例えば、以下の各状態の画像が各々多数含まれる。

(状態1) 指30がボード21に接触した状態

(状態2) 指30がボード21の上にあるが接触していない状態

(状態3) 指30がボード21に接触も上部にもない状態

なお、学習処理に際して入力する学習画像の各々には、各画像が上記のどの状態の画像であるかを示す属性情報(ラベル、タグ等)が設定されている。

#### 【0045】

10

20

30

40

50

特徴量抽出部 1 1 2 は、画像入力部 1 1 1 から入力する学習画像に含まれる特徴量を抽出する。特徴量抽出には、様々なフィルタが用いられる。図 4、図 5 に特徴量抽出に適用する矩形フィルタの例を示す。

【 0 0 4 6 】

図 4 ( a ) に示す 1 次微分フィルタは、画像内において白から黒、または、黒から白へ変化する特徴を持つ画素領域を入力画像から抽出する処理に適したフィルタである。

垂直・水平方向フィルタは白から黒、または、黒から白へ変化する方向が垂直または水平方向の場合、斜め方向フィルタは白から黒、または、黒から白へ変化する方向が斜め方向の画素領域を効率的に抽出することができる。

【 0 0 4 7 】

図 4 ( b ) に示す 2 次微分フィルタは、画像内において白 / 黒 / 白、または、黒 / 白 / 黒と変化する特徴を持つ画素領域を入力画像から抽出する処理に適したフィルタである。

垂直・水平方向フィルタは変化方向が垂直または水平方向の場合、斜め方向フィルタは、変化方向が斜め方向である画素領域を効率的に抽出することができる。

【 0 0 4 8 】

図 5 ( c ) に示す 3 次微分フィルタは、画像内において白 / 黒 / 白 / 黒、または、黒 / 白 / 黒 / 白と変化する特徴を持つ画素領域を入力画像から抽出する処理に適したフィルタである。

垂直・水平方向フィルタは変化方向が垂直または水平方向の場合、斜め方向フィルタは、変化方向が斜め方向である画素領域を効率的に抽出することができる。

なお、図 5 ( a ) ~ ( c ) に示す矩形フィルタは従来から特徴量抽出に用いられる既存のフィルタである。

【 0 0 4 9 】

図 5 ( d ) に示す離間型フィルタは、既存のフィルタではなく、本開示の情報処理装置の実行するオブジェクト位置判定処理に利用するフィルタとして考案したフィルタである。

図 1 に示す指 3 0 の位置、具体的には、指 3 0 がボード 2 1 に接触しているか否かを判定するために利用するフィルタである。

【 0 0 5 0 】

図 5 ( d ) に示す離間型フィルタは、白と黒のパターンが離間して設定されていることが特徴であり、この点が、図 4 ( a ) , ( b ) , 図 5 ( c ) に示す 1 次 ~ 3 次微分フィルタとは異なる。

この図 5 ( d ) に示す離間型フィルタを適用した指の位置判定処理については後段で詳細に説明する。

【 0 0 5 1 】

図 3 に示す情報処理装置 1 0 0 の学習処理部 1 1 0 内の特徴量抽出部 1 1 2 の実行するフィルタを適用した特徴量抽出の一例について、図 6 を参照して説明する。

図 6 は、図 4 ( A ) に示す 1 次微分フィルタの左上端に示す黒から白へ横方向に変化するフィルタ 1 3 0 を適用して、画像から同様の变化領域を持つ画素領域を抽出する処理例を説明する図である。

【 0 0 5 2 】

図 6 ( A ) は入力画像である。ここで、入力画像は学習画像である。

図 6 ( B ) は、図 4 ( A ) に示す 1 次微分フィルタの左上端に示す、黒から白へ横方向に変化するフィルタ 1 3 0 を適用したフィルタ演算を実行して得られる画像である。

フィルタ演算は、図に示すように以下の式に従って実行する。

【 0 0 5 3 】

10

20

30

40

【数 1】

$$G_{d,\theta} \otimes I(x_i, y_i, s_i)$$

・・・ (式 1)

【0054】

上記式 (式 1) に示す  $G_d$  は、図 4、図 5 を参照して説明した各フィルタに対応するガウス関数  $G$  の  $d$  次微分関数である。 $d$  は、図 4、図 5 を参照して説明した次数に相当し、 $\theta$  がフィルタの角度、すなわち図 4、図 5 を参照して説明した垂直、水平 ( $0^\circ$ ,  $90^\circ$ )、斜め ( $45^\circ$ ,  $135^\circ$ ) 等のフィルタの設定角度に相当する。

10

また、 $I(x_i, y_i, s_i)$  の  $(x_i, y_i)$  は画素位置を示し、 $s_i$  は画像のスケールを意味する。 $i$  は、適用する画像のスケール識別子である。

$I(x_i, y_i, s_i)$  はスケール  $s_i$  の画像の画素位置  $(x_i, y_i)$  の画素値、例えば輝度を示す。

上記 (式 1) は、 $G_d$  によって定義されるフィルタと、画像の各画素値との畳み込み演算を実行する式である。

図 6 に示す例は、図 4 (A) に示す 1 次微分フィルタの左上端に示す黒から白へ横方向に変化するフィルタ 130 を適用したフィルタ演算の実行例である。

20

このフィルタ演算によって、図 6 (A) 入力画像の画素値は図 6 (B) に示す画像のように変換される。

【0055】

図 6 (C) は、図 4 (A) に示すフィルタ 130 に対応する右方向に黒から白へ変化するパターンと、逆に白から黒に変化するパターンを等化させる演算を実行して得られる画像である。この処理により、例えば背景輝度に対して不変な特徴パターンの抽出が実現される。

この画像の生成に適用するフィルタ演算は以下の式に示す演算である。

【0056】

【数 2】

30

$$\left| G_{d,\theta} \otimes I(x_i, y_i, s_i) \right|$$

・・・ (式 2)

【0057】

上記式 (式 2) は、前述の (式 1) の絶対値を算出する処理に相当する。

【0058】

40

図 6 (D) は、画素位置のずれ等に対する耐性を高めるために、上記 (式 2) で得られる結果を、図 6 (C) に示す画像の所定範囲の画素領域で平滑化する処理を行なったものである。

この画像の生成に適用するフィルタ演算は以下の式に示す演算である。

【0059】

【数 3】

$$\sum_{x,y} |G_{d,\theta} \otimes I(x_i, y_i, s_i)|$$

・・・(式 3)

【0060】

例えば、上記(式 3)によって得られる値が、 $G_{d,\theta}$ によって定義されるフィルタに基づいて得られる特徴量(x)となる。

10

【0061】

図 6 に示す例は、図 4 (A) に示す 1 次微分フィルタの左上端に示す黒から白へ横方向に変化するフィルタ 130 を適用した特徴量算出処理例であり、フィルタ 130 のパターンと類似する画素領域ほど特徴量(x)の値が大きくなる。

なお、特徴量(x)は各画素対応の値として算出される。

その他のフィルタについても同様の処理を実行して、各フィルタ対応の特徴量(x)を抽出し、抽出結果を図 3 に示す機械学習部 113 に出力する。

【0062】

[ 3-2 . 機械学習部の実行する処理について ]

20

次に、図 3 に示す情報処理装置 100 の学習処理部 110 内の機械学習部 113 の実行する処理について説明する。

【0063】

図 4、図 5 に示したフィルタは、それぞれ 1 つの弱識別器(WL: Weak Learner)となり得る。弱識別器は、フィルタを探索領域に重ね合わせて得られる矩形特徴、例えば黒矩形に対応する領域内の輝度値の和と白矩形に対応する領域内の輝度値の和との差(特徴量 x)が閾値(th)よりも大きいかで、物体を検出したか否かを判定する。

【0064】

なお、本開示の処理では、特定オブジェクト(指)の検出処理、さらに、

特定オブジェクト(指)の位置判定処理として、指がボードに接触しているか否かの位置判定(接触判定)も行う。

30

オブジェクト検出処理では、フィルタ適用によって得られた特徴量(x)が閾値より大きいか否かで、特定のオブジェクトが検出されたか否かを判定する。

またオブジェクト位置判定処理(接触判定処理)では、フィルタ適用によって得られた特徴量(x)が閾値より大きいか否かで、オブジェクト(指)がボードに接触しているか否かを判定する。

【0065】

図 4、図 5 に示す白黒パターンの矩形型フィルタの適用においては、例えば、指の領域が背景より輝度値が低いという学習結果を利用して、矩形特徴に基づいて入力画像から指領域をある程度の確率で判別することができる。

40

個々の弱識別器のみでは十分な識別能力を発揮することはできないが、複数の弱識別器の識別結果を線形結合することで、より強い識別力を持つ識別器を構築することができる。

機械学習部 113 の実行する処理の 1 つは、各フィルタ対応の弱識別器によって得られた特徴量抽出結果を統合して、より強い識別器を生成することであり、認識処理部 120 において特定オブジェクト(指)の検出や位置判定に最適な特徴量の選択情報(スコア)を生成することである。

【0066】

この処理について、図 7 を参照して説明する。

なお、図 7 以下では、画像内に指が含まれるか否かのオブジェクト検出処理を行なう場

50

合の処理例について説明する。指がボードに接触しているか否かの位置判定処理については後段で説明する。

【 0 0 6 7 】

図 7 は、各フィルタを弱識別器 1 ~ n として設定し、各フィルタ対応の弱識別器 1 ~ n によって得られる特徴量 ( x ) に基づいて、各画像領域に検出対象 ( 例えば指 ) が含まれるか否かの判別結果を出力する関数  $f_1(x) \sim f_n(x)$  を設定し、これらの複数の弱識別器によって得られる判別結果に基づくスコアの算出処理例を説明する図である。

【 0 0 6 8 】

各フィルタ対応の弱識別器 1 ~ n によって得られる特徴量 ( x ) に基づいて、各画像領域に検出対象 ( 例えば指 ) が含まれるか否かの判別結果を出力する関数  $f_n(x)$  は以下の式 ( 式 4 ) で示される。

$$f_n(x) = a_n \times g(x > t_{hn}) + b_n \quad \cdots (式 4)$$

【 0 0 6 9 】

なお、式 ( 4 ) において  $a_n$  および  $b_n$  は係数 ( パラメータ ) である。また、 $g(x > t_{hn})$  は、特徴量  $x$  が閾値  $t_{hn}$  より大きい場合、所定領域が検出対象らしいことを示す数値「1」を出力し、特徴量  $x$  が閾値  $t_{hn}$  以下である場合、所定領域が検出対象らしくないことを示す数値「0」を出力する関数である。

$n$  はフィルタ識別子である。

【 0 0 7 0 】

上記 ( 式 4 ) によって算出される値が、 $f_n(x) = a_n + b_n$  である場合、所定領域は検出対象らしく、 $f_n(x) = b_n$  である場合、所定領域は検出対象らしくないとの判別がなされる。

【 0 0 7 1 】

このような弱識別器の判別結果  $f_n(x)$  の加算値をスコア  $F(x)$  とする。

スコア  $F(x)$  は、例えば次式 ( 式 5 ) で示される。

$$F(x) = f_n(x) \quad \cdots (式 5)$$

【 0 0 7 2 】

すなわち、スコア  $F(x)$  は  $N$  個の弱識別器の判別結果  $f_1(x) \sim f_n(x)$  の和となる。また、スコア  $F(x)$  の値が所定の閾値  $t_h$  より大きい場合、画像中に検出対象があるとされ、スコア  $F(x)$  の出力値が閾値  $t_h$  以下である場合、画像中に検出対象がないとされる。

【 0 0 7 3 】

機械学習部 113 は、例えば、多数の学習画像データとして、

指が含まれる画像、

指が含まれない画像、

これらの 2 種類のカテゴリに分類される画像を入力して、図 4、図 5 を参照して説明したフィルタ 1 ~ n を適用した判別結果を出力し、その出力結果に基づく統計学習を実行して、識別力の高い弱識別器を選択する処理を実行する。

この処理について、図 8 を参照して説明する。

【 0 0 7 4 】

図 8 には、学習画像の例として、

( A ) 撮影画像中に識別対象オブジェクトである指が含まれる「指あり画像」

( B ) 撮影画像中に識別対象オブジェクトである指が含まれない「指なし画像」

これらの各画像を示している。

これらの学習画像は、上記 ( A ) , ( B ) のいずれであることを示す属性情報 ( ラベル , タグ等 ) が設定された画像である。

【 0 0 7 5 】

これらの各画像に対して図 4、図 5 を参照して説明したフィルタ  $f_1 \sim f_n$  ( = 弱識別器 1 ~ n ) を適用して得られる特徴量 ( x ) を算出する。図 8 に示す 1 つの矩形の  $x$  各々が、1 つのフィルタを適用した 1 つの画像のある画素対応の特徴量  $x$  を示している。

10

20

30

40

50

なお、オブジェクト検出処理には、図4(a), (b) ~ 図5(c)に示す1次 ~ 3次微分フィルタを用いる。

オブジェクト位置判定処理には、図5(d)に記す離間型フィルタを用いる。

#### 【0076】

フィルタ適用処理によって算出される特徴量 $x$ の値は、フィルタ $f_1 \sim f_n$ (=弱識別器 $1 \sim n$ )の識別力に応じて異なる値となるが、識別力が高ければ指あり画像と、指なし画像とでは、その差が大きくなる。

図9に、各学習画像に対してフィルタ $f_1$ を適用した場合の特徴量( $x$ )と度数との対応データを示す。

この度数分布データに示すように、ある程度の識別力がある場合は、指あり画像データの度数分布と、指なし画像データの度数分布は、異なる山を形成する。

#### 【0077】

しかし、この分布は、適用するフィルタによって異なるものとなる。

図10(1)は、識別力の低い弱識別器(フィルタ)に基づく度数分布データの例を示している。

また、図10(2)は、識別力の高い弱識別器(フィルタ)に基づく度数分布データの例を示している。

このように、識別力の低い弱識別器(フィルタ)を適用すると、指あり画像と指なし画像を区別することが困難になり、識別力の高い弱識別器(フィルタ)を適用すると、指あり画像と指なし画像を区別することが容易となる。

#### 【0078】

この度数分布データに基づいて、識別力の高い弱識別器(フィルタ)のみを選択して、この選択されたフィルタのみから図7を参照して説明したスコアを算出する処理を行なうことで、より信頼度の高い識別処理が実現される。

図10(3)は、識別力の高い選択された複数の弱識別器(フィルタ)のみから算出したスコアの度数分布データの例を示している。

スコアの閾値は、例えばブースティングやサポート・ベクター・マシン(Support Vector Machine: SVM)などの統計学習器を用いた学習により決定する。

#### 【0079】

機械学習部113は、このような統計学習を実行して識別力の高いフィルタの選択を実行する。

また、先に図7を参照して説明した各フィルタ(弱識別器)により判別結果を出力する関数 $f_{n1}(x)$ 、すなわち、

$$f_n(x) = a_n \times g(x > th_n) + b_n$$

上記式において利用されるパラメータ $a_n$ ,  $b_n$ ,  $th_n$ についても、この統計学習によって、識別レベルを高く設定するための最適値を算出する。

#### 【0080】

この学習処理によって算出したパラメータと、識別力の高低を示すデータ(タグ)を各フィルタ(弱識別器)に設定して認識辞書を生成し、認識辞書記憶部114に格納する。

#### 【0081】

### [3-3. 認識処理部の実行する処理について]

次に、図3に示す情報処理装置100の認識処理部120の実行する処理について説明する。

まず、画像内に指が含まれるか否かの特定オブジェクト検出処理について説明する。

指がボードに接触しているか否かの位置判定処理については後段で説明する。

#### 【0082】

認識処理部120は、画像入力部121、特徴量抽出部122、特定オブジェクト認識部123を有する。

認識処理部120は、解析対象画像を入力して、特定オブジェクトの検出と位置判定を

10

20

30

40

50

行なう。認識処理部 120 は、学習処理部 110 が学習データに基づいて生成し、認識辞書記憶部 114 に格納した認識辞書を用いて、特定オブジェクトの検出と位置判定を行なう。

図 2 を参照して説明したフローチャートに従った処理は、この認識処理部 120 の実行する処理に相当する。

#### 【0083】

認識処理部 120 の画像入力部 121 は、解析対象とする画像を入力する。すなわち、図 1 (a) に示す撮影環境のカメラ L11 の撮影画像とカメラ R12 の撮影画像である。

この解析対象画像は、学習画像とは異なり、撮影画像中に指が含まれるか否かが不明な画像である。

なお、先に図 2 を参照して説明したように、

指が存在するか否かの指検出処理には、L 画像または R 画像の一方のみを使用し、指がボードに接触しているか否かの判定処理には L 画像と R 画像の 2 枚の画像を使用する。

#### 【0084】

特徴量抽出部 122 は、先に説明した学習処理部 110 の特徴量抽出部 112 の実行する特徴量抽出と同様、図 4、図 5 を参照して説明したフィルタを適用した特徴量抽出処理を実行する。

ただし、認識処理部 120 の特徴量抽出部 122 で適用するフィルタ（弱識別器）は、学習処理部 110 の学習結果によって、識別力が高いと判定されたフィルタのみを選択して実行する。

これらは、認識辞書記憶部 114 に格納されたフィルタ情報（タグ）に基づいて選択される。

また、各フィルタ（弱識別器）を用いて算出する特徴量（ $x$ ）の算出式、すなわち、

$$f_n(x) = a_n \times g(x > th_n) + b_n$$

上記式のパラメータ  $a_n$ 、 $b_n$ 、 $th_n$  についても、前述した学習処理部 110 の学習処理によって決定され、認識辞書記憶部 114 に格納された値を利用する。

#### 【0085】

認識処理部 120 は、このように識別力の高い選択されたフィルタを用いて特徴量を算出し、スコアの算出を実行し、スコアの値に基づいて、指検出処理と、指がボードに接触しているか否かの判定処理を行なう。

以下、図を参照して、認識処理部 120 の実行する処理について説明する。

図 11 は、認識処理部 120 の画像探索処理手順を模式的に示す図である。

#### 【0086】

まず、入力画像から、中間画像として積分画像を作成する。

入力画像に対して、直接、図 6 を参照して説明したフィルタ適用処理を実行してもよいが、中間画像として入力画像に基づく積分画像を生成し、積分画像に対してフィルタ適用を実行することで高速処理が可能となる。

なお、積分画像の生成処理と、積分画像を用いた特徴量算出の高速化処理については、本出願人の先の出願である特開 2011 - 180792 に記載されている。本開示の処理においても、この特許文献に記載されたと同様の処理手順に従って積分画像の生成と、積分画像を適用した特徴量算出を実行する。

#### 【0087】

積分画像は、垂直・水平矩形フィルタ用と斜め方向矩形フィルタ用とで作成方法が異なる。認識処理部 120 の特徴量抽出部 122 は、図 4、図 5 を参照して説明したフィルタに対応する垂直・水平矩形フィルタ用と、斜め方向矩形フィルタ用の 2 種類の積分画像（Integral image）を作成する。

#### 【0088】

認識処理部 120 の特徴量抽出部 122 は、入力画像から生成した積分画像を用い、積分画像上で各フィルタを適用して特徴量を算出する。すなわち、先に図 6 を参照して説明した特徴量（ $x$ ）の算出処理を実行する。具体的には、積分画像をスキャンして、各スキ

10

20

30

40

50

ヤン位置 (  $x$  ,  $y$  ) において、前述の式 ( 式 1 ) ~ ( 式 3 ) を適用して特徴量 (  $x$  ) を算出する。

ここで、適用するフィルタは識別レベルの高い選択されたフィルタ ( 弱識別器 ) である。

#### 【 0 0 8 9 】

特定オブジェクト認識部 1 2 3 は、複数の選択されたフィルタ ( 弱識別器 ) を適用して得られた特徴量を加算してスコアを算出する。この処理は、先に図 7 を参照して説明した処理である。

積分画像を用いることで、スキャン位置毎の矩形特徴を高速に計算することができる。

検出スコアが、予め設定した閾値以上に到達したときに、そのスキャン位置 (  $x$  ,  $y$  ) で目的オブジェクト、本実施例では指が検出されたと判定する。

#### 【 0 0 9 0 】

画像をスキャンして得られた最大の検出スコア次第で、検出結果が否定的 ( *rejection* )、すなわち、目的オブジェクト ( 指 ) が検出されなかったという結果が返されることもある。また、スケール変換すなわち入力画像の大きさを変えながら、積分画像の生成と検出スコアの計算を繰り返し行なう構成としてもよい。

#### 【 0 0 9 1 】

なお、最初に計算した積分画像をスケール変換すれば任意のサイズのウィンドウの探索が可能になるが、積分画像をスケール変換すると演算量が増大し、積分画像を使用して処理を高速化する効果を相殺することになる。そこで、図 1 1 に示した例では、入力画像をスケール変換する度に積分画像を再計算するようにしている。

#### 【 0 0 9 2 】

画像のスキャン処理例を図 1 2 に示す。

画像の右上から水平方向に順次スキャン処理を実行し、各画素点で個々の選択フィルタに基づく特徴量 (  $x$  ) を算出し、その後、選択フィルタ各々に基づいて算出された特徴量 (  $x$  ) を加算して各画素点に対応するスコアを算出する。

算出したスコアが、あらかじめ規定した閾値を超えたら目的オブジェクト ( 指 ) が検出されたと判定する。

#### 【 0 0 9 3 】

図 1 2 に示すように、スコアは各画素位置対応のスコアとして算出される。しかし、例えば 1 つの画素位置のみスコアが高くても、画像内に目的オブジェクト ( 指 ) が検出されたとはいえない。すなわち、例えば指の形状に応じた複数の画素位置に対応する複数の高スコアが検出されなければ、画像内に目的オブジェクト ( 指 ) が検出されたとはいえない。

従って、最終的に指が検出されたか否かを判定するためには、複数の画素位置のスコアを総合して評価する必要がある。

#### 【 0 0 9 4 】

特定オブジェクト認識部 1 2 3 は、このスコアの総合評価処理を実行する。

図 1 3 を参照してスコアの総合評価処理例について説明する。

図 1 3 には、積分画像に対するスキャン処理によって設定した各画素位置 ( 画素位置 1 ~  $k$  ) でのフィルタ適用処理を示す画像と、これら各画素位置で算出したスコア 1 ~ スコア  $k$  を示している。

指が検出されると、これらの各画素位置でのスコアは高くなる。

#### 【 0 0 9 5 】

画像に目的オブジェクト ( 指 ) が含まれるか否かを判定する最終的な指標値としての総合評価値である総合スコアは、これら各画素位置のスコアの加算値として算出する。すなわち、

総合スコア = スコア 1 + スコア 2 + . . . + スコア  $k$

このように画像の各画素対応のスコアを加算した値を総合スコアとし、総合スコアが予め設定した閾値以上である場合は、目的オブジェクト ( 指 ) が検出されたと判定する。



## 【 0 0 9 6 】

## 〔 4 . 特定オブジェクトの位置判定処理について 〕

次に、認識処理部 1 2 0 の実行するもう 1 つの処理である特定オブジェクトの位置判定処理、具体的には、図 1 に示す指 3 0 がボード 2 1 に接触しているか否かを判定する処理について説明する。

## 【 0 0 9 7 】

## 〔 4 - 1 . 学習処理部の実行する処理について 〕

指 3 0 がボード 2 1 に接触しているか否かを判定する処理を行なう場合も、図 3 に示す情報処理装置 1 0 0 の学習処理部 1 1 0 が多数の学習画像を利用して認識辞書を生成する。この処理の流れは、上述した説明に従った処理となる。

ただし、この学習処理で学習対象とする画像は、図 1 に示すカメラ L 1 1 の撮影画像である L 画像と、カメラ R 1 2 の撮影画像である R 画像を連結した連結画像とする。

また、特徴量算出に用いるフィルタ（弱識別器）は、図 5（d）に示す離間型フィルタとする。

## 【 0 0 9 8 】

図 1 4 に処理対象とする画像の例とフィルタ例を示す。

図 1 4（A）は、処理対象画像であり、図 1 に示すカメラ L 1 1 の撮影画像である L 画像と、カメラ R 1 2 の撮影画像である R 画像を連結した画像となる。

なお、学習処理部 1 1 0 における学習画像を適用した学習処理においても、認識処理部 1 2 0 における解析対象画像を入力した認識処理においても、図 1 4（A）に示すような L R 連結画像を利用する。

## 【 0 0 9 9 】

図 1 4（B）は特徴量（x）の算出に適用するフィルタの例を示している。これらのフィルタは、先に図 5（d）を参照して説明した離間型フィルタであり、白領域と黒領域とが離間して設定されたフィルタである。

なお、図には、白領域と黒領域を各々 1 つの領域とした 1 次元微分形式相当のフィルタのみを示しているが、白領域を 2 つ、黒領域を 1 つとした 2 次元微分形式や、白領域を 2 つ、黒領域も 2 つ設定した 3 次元形式のフィルタを利用してもよい。

## 【 0 1 0 0 】

学習画像を適用した学習処理と、解析対象となる画像を入力した認識処理の双方においてこの離間型フィルタを適用する。

## 【 0 1 0 1 】

学習処理においては、

（a）指がボードに接触している画像、

（b）指がボードに接触していない画像、

上記 2 つのカテゴリの画像として多数の L R 連結画像を生成して学習処理を行なう。

## 【 0 1 0 2 】

なお、学習処理に際しては、学習処理部 1 1 0 の特徴量抽出部 1 1 2 が図 1 4（A）に示す L R 連結画像の生成を行うが、この際、指の含まれる画像領域のみを切り出して連結して連結画像を生成する処理を行なってもよい。

学習処理部 1 1 0 の実行する学習処理の目的は、指がボードに接触しているか否かを識別するために適したフィルタを選択することと、フィルタの特徴量算出式に用いるパラメータを決定することである。これらの処理を効率的に実行するために画像領域の切り出しを行って連結する。

なお、学習処理において決定するパラメータとは、先に説明した各フィルタ（弱識別器）の特徴量算出式、すなわち、

$$f_n(x) = a_n \times g(x > t_{hn}) + b_n$$

上記式のパラメータ  $a_n$  ,  $b_n$  ,  $t_{hn}$

である。

## 【 0 1 0 3 】

10

20

30

40

50

L R 連結画像の生成と画像切り出し処理例について、図 15 を参照して説明する。

図 15 には、

- (1) 図 1 に示すカメラ L の撮影画像である撮影 L 画像
- (2) 図 1 に示すカメラ R の撮影画像である撮影 R 画像、
- (3) 特徴量抽出処理対象とする L R 連結画像

これらの画像の例を示している。

#### 【0104】

図 15 (3) に示す L R 連結画像は、撮影 L 画像と撮影 R 画像の一部の画像領域を切り出して連結した画像である。

切り出し処理の方法としては様々な手法が適用できるが、例えば、撮影 R 画像に含まれる指の指先と指の根本位置を切り出し画像の中心位置に設定するように R 画像を切り出す。撮影 L 画像については、撮影 R 画像の切り出し位置に対応する領域を切り出す。このような処理を実行してそれぞれの切り出し画像を連結して L R 連結画像を生成する。

この処理を行なうことで、多数の L R 連結画像の処理が必要となる学習処理において、効率的にフィルタの評価やパラメータの決定が可能となる。

#### 【0105】

図 16 は、画像切り出し処理と連結処理によって生成した特徴量抽出のための L R 連結画像の例を示す図である。

なお、図 16 に示す L R 連結画像例 1 ~ 2 はいずれも模式的な積分画像の例である。学習処理においても解析処理と同様、先に図 11 を参照して説明した積分画像を生成して、積分画像に対するスキャン処理を実行して特徴量の算出を行うことが可能である。

#### 【0106】

図 3 の学習処理部 110 の特徴量抽出部 112 は、このような L R 画像を利用したスキャン処理を実行して特徴量抽出を行う。なお、適用するフィルタは、図 5 (d) や図 14 (B) に示す離間型フィルタである。

#### 【0107】

図 3 の学習処理部 110 の機械学習部 113 は、特徴量抽出部の算出した特徴量を入力してフィルタの評価やパラメータの決定を行う。

学習処理においては、

- (a) 指がボードに接触している画像、
- (b) 指がボードに接触していない画像、

上記 2 つのカテゴリの画像として多数の L R 連結画像を生成して学習処理を行なう。

#### 【0108】

これらの 2 種類のカテゴリに分類される画像を入力して、図 5 (d) や図 14 (B) に示す離間型フィルタを適用した判別結果を出力し、その出力結果に基づく統計学習を実行して、識別力の高い弱識別器を選択する処理を実行する。

この処理について、図 17 を参照して説明する。

図 17 には、学習画像の例として、

- (A) 指がボードに接触している画像「指接触 L R 連結画像」
- (B) 指がボードに接触していない画像「指非接触 L R 連結画像」

これらの各画像を示している。

これらの学習画像は、上記 (A) , (B) のいずれであるかを示す属性情報 (ラベルやタグ) が設定された画像である。

#### 【0109】

これらの各画像に対して図 5 (d) や図 14 (B) に示す離間型フィルタ  $f_1 \sim f_n$  (= 弱識別器 1 ~ n) を適用して得られる特徴量 ( $x$ ) を算出する。図 17 に示す 1 つの矩形の各々が算出特徴量  $x$  を示している。

#### 【0110】

なお、学習画像として利用する L R 連結画像は、例えば動画画像の撮影処理によって得られる複数の連続撮影画像フレームを平均化して作成した平均画像としてもよい。

例えば、

第 1 フレームから第 20 フレームの平均画像に基づいて生成した L R 連結画像を第 1 の学習画像、

第 21 フレームから第 40 フレームの平均画像に基づいて生成した L R 連結画像を第 2 の学習画像、

このような処理を実行して、複数の学習画像を生成して処理を行なってもよい。

#### 【0111】

この特徴量  $x$  の値は、フィルタ  $f_1 \sim f_n$  (= 弱識別器  $1 \sim n$ ) の識別力に応じて異なる値となるが、識別力が高ければ「指接触 L R 連結画像」と、「指非接触 L R 連結画像」とでは、その差が大きくなる。

図 18 に、各学習画像に対してフィルタ  $f_1$  を適用した場合の特徴量 ( $x$ ) と度数との対応データを示す。

この度数分布データに示すように、ある程度の識別力がある場合は、「指接触 L R 連結画像」対応の度数分布と、「指非接触 L R 連結画像」対応の度数分布は、異なる山を形成する。

#### 【0112】

しかし、この分布は、適用するフィルタによって異なるものとなる。

図 19 (1) は、識別力の低い弱識別器 (フィルタ) に基づく度数分布データの例を示している。

また、図 19 (2) は、識別力の高い弱識別器 (フィルタ) に基づく度数分布データの例を示している。

このように、識別力の低い弱識別器 (フィルタ) を適用すると、「指接触 L R 連結画像」と「指非接触 L R 連結画像」を区別することが困難になり、識別力の高い弱識別器 (フィルタ) を適用すると、「指接触 L R 連結画像」と「指非接触 L R 連結画像」を区別することが容易となる。

#### 【0113】

この度数分布データに基づいて、識別力の高い弱識別器 (フィルタ) のみを選択して、この選択されたフィルタのみから図 7 を参照して説明したスコアを算出する処理を行なうことで、より信頼度の高い識別処理が実現される。

図 19 (3) は、識別力の高い選択された複数の弱識別器 (フィルタ) のみから算出したスコアの度数分布データの例を示している。

スコアの閾値は、例えばブースティングやサポート・ベクター・マシン (Support Vector Machine: SVM) などの統計学習器を用いた学習により決定する。

#### 【0114】

機械学習部 113 は、このような統計学習を実行して識別力の高いフィルタの選択を実行する。さらにパラメータの決定を行う。

すなわち、先に図 7 を参照して説明した各フィルタ (弱識別器) により判別結果を出力する関数、

$$f_n(x) = a_n \times g(x > th_n) + b_n$$

上記式において利用されるパラメータ  $a_n$ ,  $b_n$ ,  $th_n$  についても、この統計学習によって、識別レベルを高く設定するための最適値を算出する。

#### 【0115】

この学習処理によって算出したパラメータと、識別力の高低を示すデータをフィルタ情報として各フィルタ (弱識別器) に対応付けた認識辞書を生成し、認識辞書記憶部 114 に格納する。

#### 【0116】

[4-2. 認識処理部の実行する処理について]

次に、図 3 に示す情報処理装置 100 の認識処理部 120 がオブジェクト位置判定処理 (接触判定処理) において実行する処理について説明する。

認識処理部 1 2 0 の画像入力部 1 2 1 は、解析対象画像となる L 画像と R 画像を入力する。

特徴量抽出部 1 2 2 は、これらの入力 L 画像と R 画像を用いて、L R 連結画像を生成して特徴量算出を行う。

#### 【 0 1 1 7 】

特徴量抽出部 1 2 2 は、異なる視点から撮影した画像を連結して生成した連結画像から特徴量を抽出する際に、フィルタ構成領域を離間して設定した離間型フィルタを適用し、離間型フィルタを構成する離間した各フィルタ領域を、連結画像の 2 つの画像の各々に設定して特徴量抽出処理を実行する。

また、特定オブジェクト認識部 1 2 3 は、特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの位置判定処理を実行する。

10

#### 【 0 1 1 8 】

特徴量抽出部 1 2 2 は、先に説明した学習処理部 1 1 0 の特徴量抽出部 1 1 2 の実行する特徴量抽出と同様、図 5 ( d ) や図 1 4 ( B ) に示す離間型フィルタを適用した特徴量抽出処理を実行する。

ただし、認識処理部 1 2 0 の特徴量抽出部 1 2 2 で適用するフィルタ ( 弱識別器 ) は、学習処理部 1 1 0 の学習結果によって、識別力が高いと判定されたフィルタのみを利用して実行する。

すなわち、指がボードに接触した状態 ( 第 1 状態 ) と、接触していない状態 ( 第 2 状態 ) の識別力が高いと判定された離間型フィルタのみを利用して実行する。

20

#### 【 0 1 1 9 】

これらは、認識辞書記憶部 1 1 4 に格納されたフィルタ情報 ( タグ ) に基づいて選択される。

また、各フィルタ ( 弱識別器 ) を用いて算出する特徴量 (  $x$  ) の算出式、すなわち、

$$f_n(x) = a_n \times g(x > t_{hn}) + b_n$$

上記式のパラメータ  $a_n$  ,  $b_n$  ,  $t_{hn}$  も、前述した学習処理部 1 1 0 の学習処理で決定し認識辞書記憶部 1 1 4 に格納した値を利用する。

#### 【 0 1 2 0 】

認識処理部 1 2 0 は、識別力の高いと判定されたフィルタを選択的に用いて特徴量を算出し、スコアの算出を実行し、スコアの値に基づいて、指検出処理と、指がボードに接触しているか否かの判定処理を行なう。

30

#### 【 0 1 2 1 】

以下、図を参照して、認識処理部 1 2 0 の実行する処理について説明する。

認識処理部 1 2 0 の実行する処理は、先に特定オブジェクト検出処理の説明において、図 1 1 以下を参照して説明した処理とほぼ同様の処理である。

ただし、処理対象画像が L R 連結画像であり、適用するフィルタが図 5 ( d ) や図 1 4 ( B ) に示す離間型フィルタとなる点が異なる。

#### 【 0 1 2 2 】

なお、認識処理において処理対象とする L R 連結画像は、L R カメラ各々で撮影した 1 枚の撮影画像フレームを連結した L R 画像でもよいし、先に説明した学習画像と同様、動画像の撮影処理によって得られる複数の連続撮影画像フレームを平均化して作成した平均画像としてもよい。

40

例えば、

第 1 フレームから第 2 0 フレームの平均画像に基づいて生成した L R 連結画像を生成して処理を行なってもよい。

#### 【 0 1 2 3 】

認識処理部 1 2 0 の実行する処理について説明する。

認識処理部 1 2 0 は、先に図 1 1 を参照して説明した手順に従って、L R 連結画像から中間画像として積分画像を作成する。

#### 【 0 1 2 4 】

50

認識処理部 120 の特徴量抽出部 122 は、積分画像を用いて、積分画像上で各フィルタを適用して特徴量を算出する。すなわち、例えば先に図 6 を参照して説明した特徴量算出処理を実行する。具体的には、積分画像をスキャンして、各スキャン位置 (x, y) において、前述の式 (式 1) ~ (式 3) を適用して特徴量 (x) を算出する。

ここで、適用するフィルタは、学習処理において選択された識別力の高い離間型フィルタ (弱識別器) である。

複数の選択されたフィルタ (弱識別器) を適用して得られた特徴量を加算してスコアが算出される。この処理は、先に図 7 を参照して説明した処理と同様の処理である。

積分画像を用いることで、スキャン位置毎の矩形特徴を高速に計算することができる。

検出スコアが、予め設定した閾値以上に到達したときに、そのスキャン位置 (x, y) で「指がボードに接触している」と判定する。

#### 【0125】

画像をスキャンして得られた最大の検出スコア次第で、検出結果が否定的 (rejection)、すなわち、「指がボードに接触していない」という結果が返されることもある。また、スケール変換すなわち入力画像の大きさを変えながら、積分画像の生成と検出スコアの計算を繰り返し行なう構成としてもよい。

#### 【0126】

なお、最初に計算した積分画像をスケール変換すれば任意のサイズのウィンドウの探索が可能になるが、積分画像をスケール変換すると演算量が増大し、積分画像を使用して処理を高速化する効果を相殺することになる。そこで、図 11 に示した例では、入力画像をスケール変換する度に積分画像を再計算するようにしている。

#### 【0127】

なお、先に図 12 を参照して説明した処理と同様、スコアは各画素位置対応のスコアとして算出される。

前述したように、例えば 1 つの画素位置のみスコアが高くても、指がボードに接触していると結論付けることはできない。すなわち、複数の画素位置に対応する複数の高スコアが検出されなければ、画像にある指がボードに接触している状態にあると判定することはできない。

従って、最終的に指がボードに接触しているか否かを判定するためには、複数の画素位置のスコアを総合して評価する必要がある。

#### 【0128】

特定オブジェクト認識部 123 は、このスコアの総合評価処理を実行する。

この処理は、先に図 13 を参照して説明したスコアの総合評価処理と同様の処理となる。

先に図 13 を参照して説明した処理と同様の総合スコア算出を行う。すなわち、画像の指がボードに接触しているか否かを判定する最終的な指標値としての総合評価値である総合スコアは各画素位置のスコアの加算値として算出する。すなわち、

総合スコア = スコア 1 + スコア 2 + ... + スコア k

このように画像の各画素対応のスコアを加算した値を総合スコアとし、総合スコアが予め設定した閾値以上である場合は、指がボードに接触していると判定する。

#### 【0129】

図 20 は、指がボードに接触している場合の LR 連結画像と、指がボードに接触していない場合の LR 連結画像の差異を説明する図である。

図 20 には以下の 2 つの平均画像を示している。

(a) 指がボードに接触している LR 連結画像の平均画像

(b) 指がボードに接触していない LR 連結画像の平均画像

#### 【0130】

R 画像の指の位置を一定とした LR 連結画像を処理対象画像 (学習画像、解析画像) に設定することで、(a) に示す指がボードに接触している LR 連結画像は、以下の特徴を有することになる。すなわち、異なるタイミングで撮影された画像でも、LR カメラの位

10

20

30

40

50

置が固定されており、視差が一定であるので、L R 連結画像上では、ほぼ同じ位置に指の位置が設定される。従って、L 画像の平均画像において、指の輪郭が比較的、明確に示される。

#### 【 0 1 3 1 】

一方、( b ) に示す指がボードに接触していない画像では、指がボードに接触していないので、異なるタイミングで撮影された画像の指の位置が画像ごとに異なる。この結果、L 画像の平均画像では指の輪郭がぼやける。

#### 【 0 1 3 2 】

##### [ 5 . フィルタ適用位置を限定した実施例について ]

上述したように、指がボードに接触しているか否かを判定するための処理として、複数の異なる視点から撮影した画像の連結画像である L R 連結画像を利用し、離間型フィルタを適用した特徴量抽出を行う実施例について説明した。

#### 【 0 1 3 3 】

しかし、特徴量抽出に際して、L R 連結画像のすべての画像の画素位置に対するフィルタ適用処理は、時間を要し、結果を出力するまでの時間が長くなる可能性がある。

このような遅延を防止し、高速処理を可能とする実施例として、フィルタ適用位置を限定した実施例について説明する。

図 2 1 を参照してこの実施例について説明する。

本実施例は、L R 連結画像に設定されるエピポーラライン上にフィルタ適用位置を限定した実施例である。

#### 【 0 1 3 4 】

エピポーララインは、異なる視点からの 2 つの画像において、同一オブジェクト位置の探索ラインとして設定されるラインであり、被写体距離の計測などにおいて利用されるラインである。

#### 【 0 1 3 5 】

例えば図 2 1 ( A ) に示す R 画像を基準画像とした場合、R 画像上のある点 ( 特徴点 1 5 0 ) に対応する L 画像上の点は、L 画像に設定したエピポーラライン 1 5 1 上に限定される。これはエピポーラ拘束 ( e p i p o l a r c o n s t r a i n t ) と呼ばれる。

#### 【 0 1 3 6 】

本実施例では、L R 連結画像における R 画像のある特徴点に離間型フィルタの一方のフィルタ領域を設定した場合、分離フィルタのもう一方のフィルタ領域を L 画像に設定した特徴点对応のエピポーラライン上に設定する。

このように、フィルタ適用位置をエピポーラライン上にのみ限定してフィルタ適用処理を実行する。

#### 【 0 1 3 7 】

具体的には、例えば、L R 連結画像の R 画像部分からエッジ領域等の特徴点と判断される点を選択し、その特徴点に対応するエピポーララインを L 画像上に設定する。次に離間型フィルタの一方のフィルタ領域を R 画像の特徴点に設定し、もう一方のフィルタ領域を L 画像上のエピポーラライン上に設定する。

このような設定を行ってフィルタ適用処理を実行して特徴量を算出する。さらに、L 画像上のエピポーラライン上で、フィルタをスライドさせて各画素位置での特徴量を算出してもよい。

このような処理を行なうことで、フィルタ適用処理を実行する画素領域を限定することが可能となり、処理量を削減し、高速化処理を実現する。

#### 【 0 1 3 8 】

図 2 1 ( B 1 ) は、このようなエピポーラ拘束を適用したフィルタ適用処理を実行した例である。

一方、図 2 1 ( B 2 ) は、なエピポーラ拘束を適用せずにフィルタ適用処理を実行した例である。

図 2 1 ( B 2 ) の例では、L R 全連結画像の全画素領域に対するフィルタ適用処理を行

10

20

30

40

50

なって全画素対応の特徴量算出が実行されるが、そのほとんどは、有効な特徴量が得られないと推定される。

一方、図 2 1 ( B 1 ) の例では、L 画像と R 画像の対応点が検出されるエッジラインによって限定された範囲でフィルタ適用処理を行なうので、処理時間が短縮されるとともに、有効な特徴量のみを選択的に抽出することが可能となる。

【 0 1 3 9 】

[ 6 . 離間型フィルタのタイプと特性について ]

先の実施例において説明したように、指がボードに接触しているか否かを判定するための処理には、複数の異なる視点から撮影した画像の連結画像である L R 連結画像を利用し、先に図 5 ( d ) や図 1 4 ( B ) において説明した離間型フィルタを適用した特徴量抽出を行う。

【 0 1 4 0 】

指がボードに接触しているか否かの判定処理に適用する離間型フィルタとしては様々なフィルタが利用可能である。

フィルタの形式によって、画像から抽出される特徴量の持つ意味合い、すなわち抽出される画像特性が異なるものとなる。

【 0 1 4 1 】

図 2 2 に、離間型フィルタの構成例と、各離間型フィルタによって抽出される特徴量データに反映される主な画像特性についてまとめた表を示す。

【 0 1 4 2 】

( 1 ) フィルタ 1 は、ほぼ正方形型の傾き 4 5 ° の白黒パターンのフィルタ領域が離間した離間型フィルタである。

このフィルタの適用処理によって得られる特徴量データは、主として視差情報を反映した特徴量となる。

( 2 ) フィルタ 2 は、長方形型の傾き 4 5 ° の白黒パターンのフィルタ領域が離間した離間型フィルタである。

このフィルタの適用処理によって得られる特徴量データは、主としてオブジェクト形状を反映した特徴量となる。

( 3 ) フィルタ 3 は、長方形型の傾き 0 ° の白黒パターンのフィルタ領域が離間した離間型フィルタである。

このフィルタの適用処理によって得られる特徴量データは、主として視差情報を反映した特徴量となる。

( 4 ) フィルタ 4 は、幅の狭い長方形型の傾き 0 ° の白黒パターンのフィルタ領域が離間した離間型フィルタである。

このフィルタの適用処理によって得られる特徴量データは、主としてオブジェクト形状を反映した特徴量となる。

( 5 ) フィルタ 5 は、小型の正方形型の傾き 4 5 ° の白黒パターンのフィルタ領域が離間した離間型フィルタである。

このフィルタの適用処理によって得られる特徴量データは、主としてオブジェクト形状を反映した特徴量となる。

【 0 1 4 3 】

なお、図 2 2 に示す離間型フィルタは、適用可能なフィルタの一例であり、この他にもさまざまなパターンのフィルタが適用可能である。

【 0 1 4 4 】

[ 7 . 情報処理装置の構成例について ]

次に、図 2 3 を参照して上述した処理を実行する情報処理装置 2 0 0 のハードウェア構成例について説明する。

CPU (Central Processing Unit) 201は、ROM (Read Only Memory) 202、または記憶部208に記憶されているプログラムに従って各種の処理を実行するデータ処理部として機能する。例えば、先に説明した実施例に従った処理等を実行する。

【0145】

RAM (Random Access Memory) 203には、CPU 201が実行するプログラムやデータなどが適宜記憶される。これらのCPU 201、ROM 202、およびRAM 203は、バス204により相互に接続されている。

【0146】

CPU 201はバス204を介して入出力インタフェース205に接続され、入出力インタフェース205には、各種スイッチ、キーボード、マウス、マイクロホンなどよりなる入力部206、ディスプレイ、スピーカなどよりなる出力部207が接続されている。CPU 201は、入力部206から入力される指令に対応して各種の処理を実行し、処理結果を例えば出力部207に出力する。

【0147】

入出力インタフェース205に接続されている記憶部208は、例えばハードディスク等からなり、CPU 201が実行するプログラムや各種のデータを記憶する。通信部209は、インターネットやローカルエリアネットワークなどのネットワークを介して外部の装置と通信する。

【0148】

入出力インタフェース205に接続されているドライブ210は、データの記録再生対象となるリムーバブルメディア211を駆動し、データの記録、再生を実行する。

CPU 201が実行するプログラムは、例えば、パッケージメディア等としてのリムーバブルメディア211に記録して提供することができる。

【0149】

[ 8 . 本開示の構成のまとめ ]

以上、特定の実施例を参照しながら、本開示の実施例について詳解してきた。しかしながら、本開示の要旨を逸脱しない範囲で当業者が実施例の修正や代用を成し得ることは自明である。すなわち、例示という形態で本発明を開示してきたのであり、限定的に解釈されるべきではない。本開示の要旨を判断するためには、特許請求の範囲の欄を参酌すべきである。

【0150】

なお、本明細書において開示した技術は、以下のような構成をとることができる。

( 1 ) 異なる視点から撮影した画像を連結して生成した連結画像から特徴量を抽出する特徴量抽出部と、

前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの位置判定処理を実行する特定オブジェクト認識部を有し、

前記特徴量抽出部は、フィルタ構成領域を離間して設定した離間型フィルタを適用した特徴量抽出処理を実行する情報処理装置。

【0151】

( 2 ) 前記特徴量抽出部は、前記離間型フィルタを構成する離間した各フィルタ領域を、前記連結画像の2つの画像の各々に設定して特徴量抽出処理を実行する前記( 1 )に記載の情報処理装置。

【0152】

( 3 ) 前記特徴量抽出部は、予め実行した学習処理において前記特定オブジェクトの第1状態と第2状態との識別力が高いと判定したフィルタを選択的に適用して特徴量抽出を実行する前記( 1 )または( 2 )に記載の情報処理装置。

【0153】

( 4 ) 前記情報処理装置は、前記特定オブジェクトの第1状態を示す連結画像と、第2状態を示す連結画像を多数入力して、異なる形式の離間型フィルタを適用した特徴量抽出

10

20

30

40

50



により、識別力の高いフィルタを判別する処理を実行する学習処理部を有する前記(1)~(3)いずれかに記載の情報処理装置。

【0154】

(5)前記特定オブジェクト認識部は、前記特徴量抽出部の抽出した異なる離間型フィルタ対応の特徴量を加算したスコアを算出し、加算スコアに基づいて前記特定オブジェクトの位置判定処理を実行する前記(1)~(4)いずれかに記載の情報処理装置。

【0155】

(6)前記特定オブジェクト認識部は、前記特徴量抽出部の抽出した異なる離間型フィルタ対応の画素対応の特徴量を加算した画素対応のスコアを算出し、算出した画素対応のスコアをさらに加算した総合スコアに基づいて前記特定オブジェクトの位置判定処理を実行する前記(1)~(5)いずれかに記載の情報処理装置。

10

【0156】

(7)前記特徴量抽出部は、前記特定オブジェクトである指がボードに接触している第1状態と、指がボードに接触していない第2状態を判別する指標値としての特徴量を算出し、前記特定オブジェクト認識部は、前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトである指がボードに接触しているか否かの判定処理を実行する前記(1)~(6)いずれかに記載の情報処理装置。

【0157】

(8)前記特徴量抽出部は、さらに、画像内に特定オブジェクトが存在するか否かを判別する指標値としての特徴量を抽出し、前記特定オブジェクト認識部は、前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの有無を判定する特定オブジェクト検出処理を実行する前記(1)~(7)いずれかに記載の情報処理装置。

20

【0158】

(9)前記離間型フィルタは、異なる輝度のフィルタ領域を離間して設定したフィルタである前記(1)~(8)いずれかに記載の情報処理装置。

【0159】

(10)前記特徴量抽出部は、前記連結画像の一方の画像の特徴点と、他方の画像上の前記特徴点に対応するエッジライン上に前記離間型フィルタを設定して特徴量抽出を実行する前記(1)~(9)いずれかに記載の情報処理装置。

【0160】

30

(11) 情報処理装置において実行する情報処理方法であり、

特徴量抽出部が、異なる視点から撮影した画像を連結して生成した連結画像から特徴量を抽出する特徴量抽出処理、

特定オブジェクト認識部が、前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの位置判定処理を実行する特定オブジェクト認識処理を実行し、

前記特徴量抽出部は、フィルタ構成領域を離間して設定した離間型フィルタを適用した特徴量抽出処理を実行する情報処理方法。

【0161】

(12) 情報処理装置において情報処理を実行させるプログラムであり、

特徴量抽出部に、異なる視点から撮影した画像を連結して生成した連結画像から特徴量を抽出させる特徴量抽出処理、

40

特定オブジェクト認識部に、前記特徴量抽出部の抽出した特徴量に基づいて、特定オブジェクトの位置判定処理を実行する特定オブジェクト認識処理を実行させ、

前記特徴量抽出部には、フィルタ構成領域を離間して設定した離間型フィルタを適用した特徴量抽出処理を実行させるプログラム。

【0162】

また、明細書中において説明した一連の処理はハードウェア、またはソフトウェア、あるいは両者の複合構成によって実行することが可能である。ソフトウェアによる処理を実行する場合は、処理シーケンスを記録したプログラムを、専用のハードウェアに組み込まれたコンピュータ内のメモリにインストールして実行させるか、あるいは、各種処理が実

50

行可能な汎用コンピュータにプログラムをインストールして実行させることが可能である。例えば、プログラムは記録媒体に予め記録しておくことができる。記録媒体からコンピュータにインストールする他、LAN (Local Area Network)、インターネットといったネットワークを介してプログラムを受信し、内蔵するハードディスク等の記録媒体にインストールすることができる。

#### 【0163】

なお、明細書に記載された各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的にあるいは個別に実行されてもよい。また、本明細書においてシステムとは、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

10

#### 【産業上の利用可能性】

#### 【0164】

以上、説明したように、本開示の一実施例の構成によれば、特定オブジェクトの位置判定を異なる視点からの撮影画像の連結画像から抽出した特徴量に基づいて実行する装置、方法が実現される。

具体的には、異なる視点からの撮影画像を連結した連結画像から特徴量を抽出し、抽出した特徴量に基づいて特定オブジェクトの位置判定処理を実行する。特徴量抽出部は、フィルタ構成領域を離間して設定した離間型フィルタを適用した特徴量抽出処理を実行し、離間型フィルタの各フィルタ領域を、連結画像の2つの画像の各々に設定して特徴量抽出処理を実行する。特定オブジェクト認識部は、特徴量抽出部の抽出した異なる離間型フィルタ対応の特徴量を加算したスコアに基づいて特定オブジェクトの位置判定処理を実行する。

20

本構成により、特定オブジェクトの位置判定を異なる視点からの撮影画像の連結画像から抽出した特徴量に基づいて実行する装置、方法が実現される。

#### 【符号の説明】

#### 【0165】

- 11 カメラL
- 12 カメラR
- 21 ボード
- 30 指
- 51 L画像
- 52 R画像
- 100 情報処理装置
- 110 学習処理部
- 111 画像入力部
- 112 特徴量抽出部
- 113 機械学習部
- 114 認識辞書記憶部
- 120 認識処理部
- 121 画像入力部
- 122 特徴量抽出部
- 123 特定オブジェクト認識部
- 200 情報処理装置
- 201 CPU
- 202 ROM
- 203 RAM
- 204 バス
- 205 入出力インタフェース
- 206 入力部
- 207 出力部

30

40

50

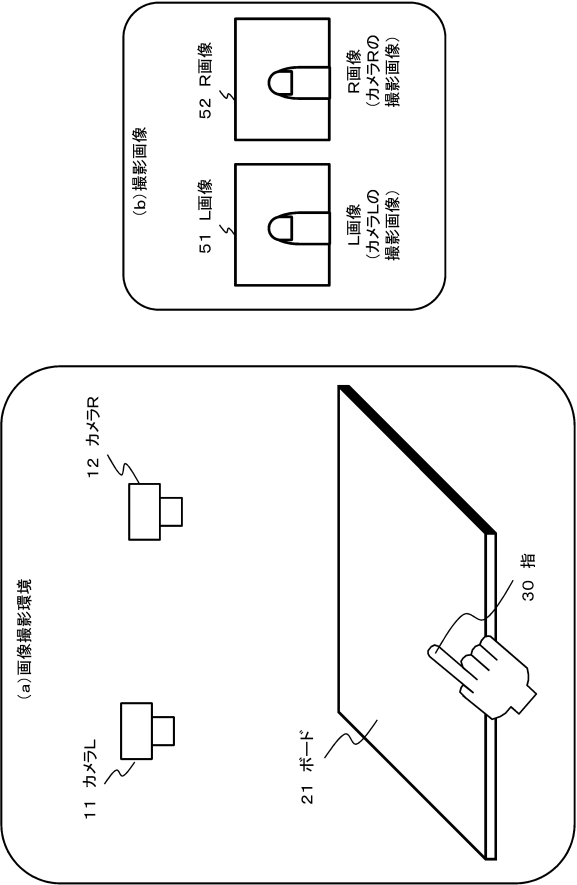
2 0 8    記憶部

2 0 9    通信部

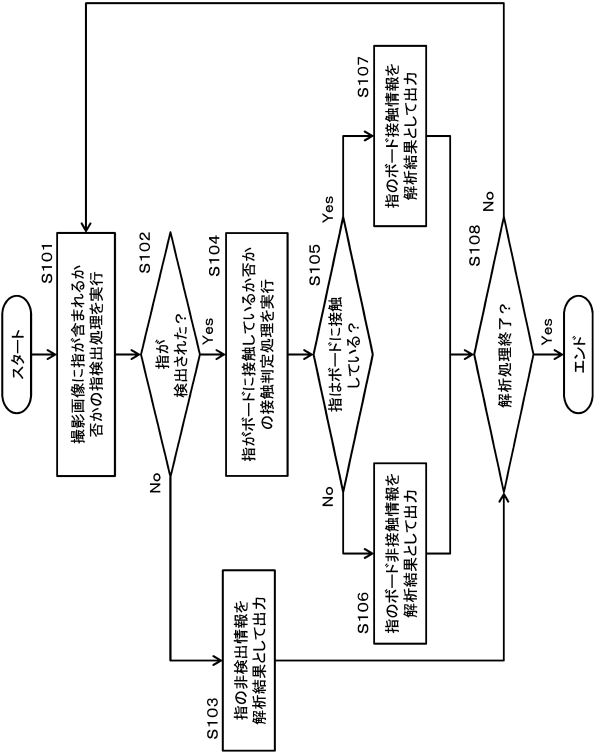
2 1 0    ドライブ

2 1 1    リムーバブルメディア

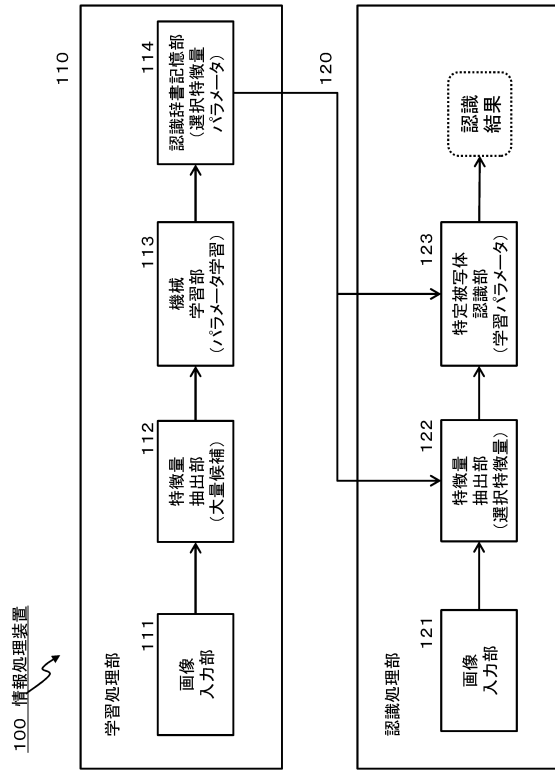
【 図 1 】



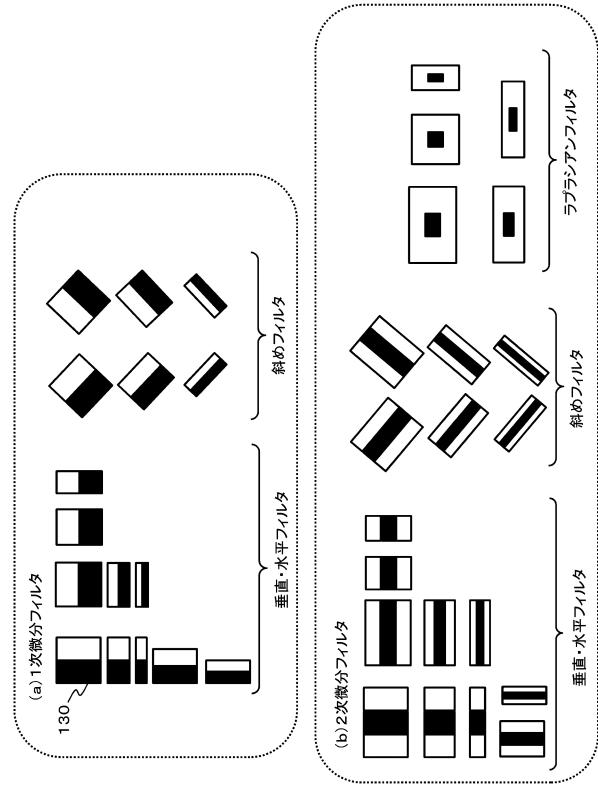
【 図 2 】



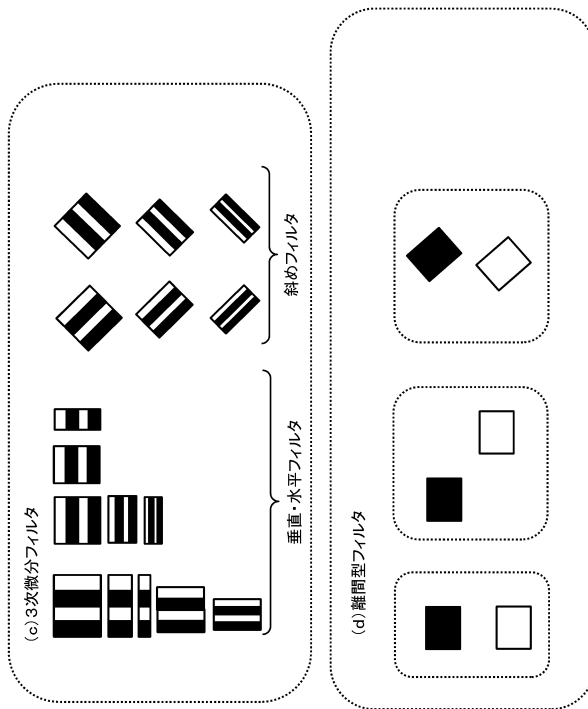
【図 3】



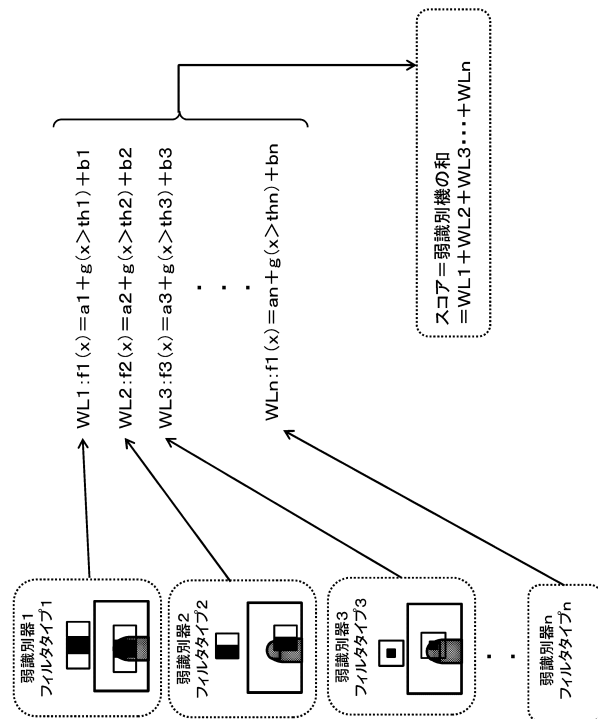
【図 4】



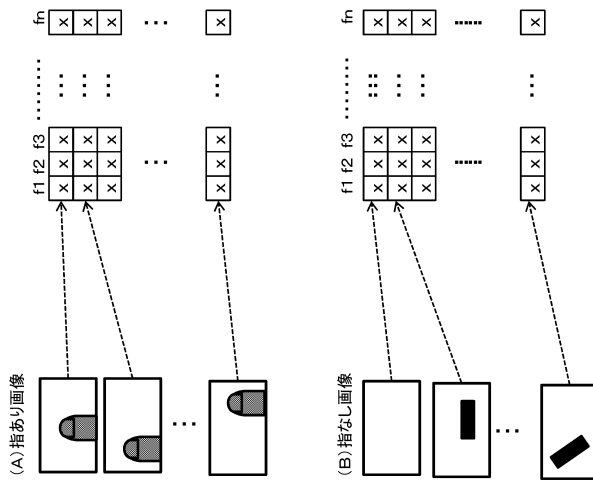
【図 5】



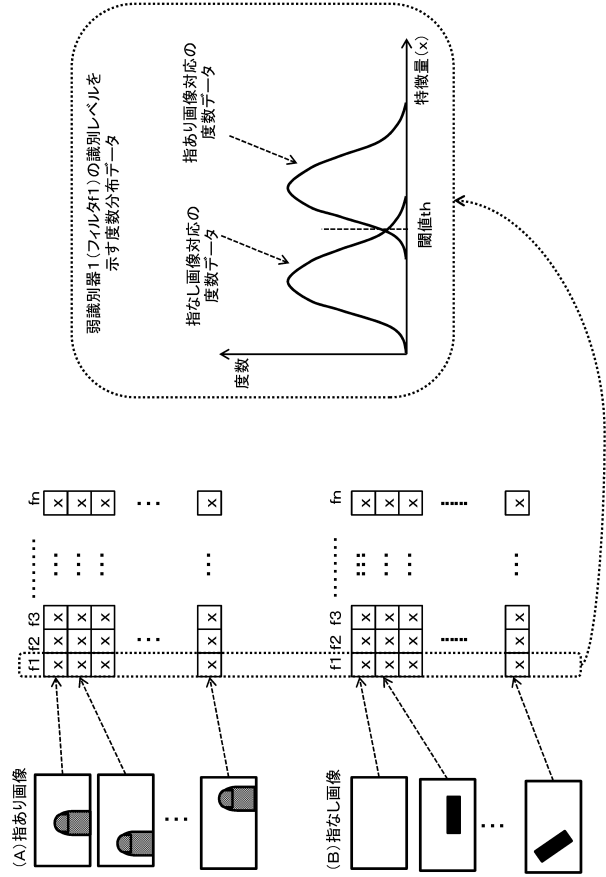
【図 7】



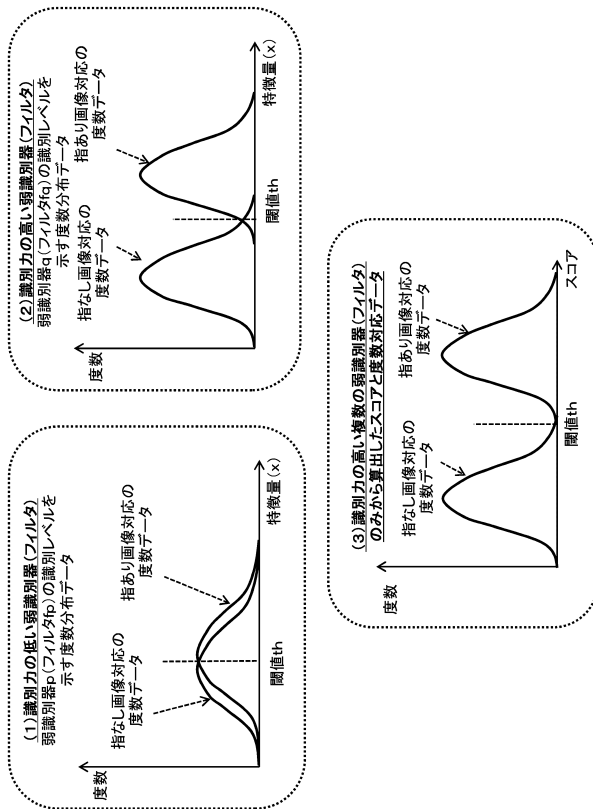
【図 8】



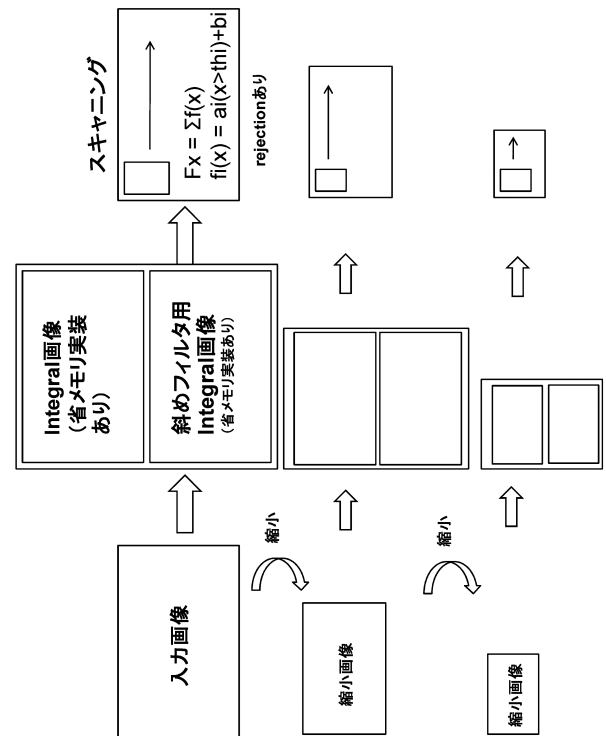
【図 9】



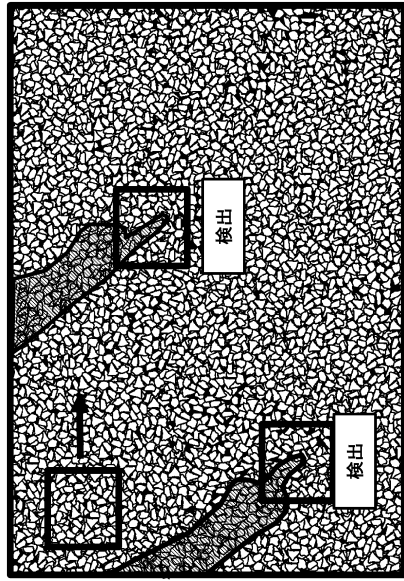
【図 10】



【図 11】

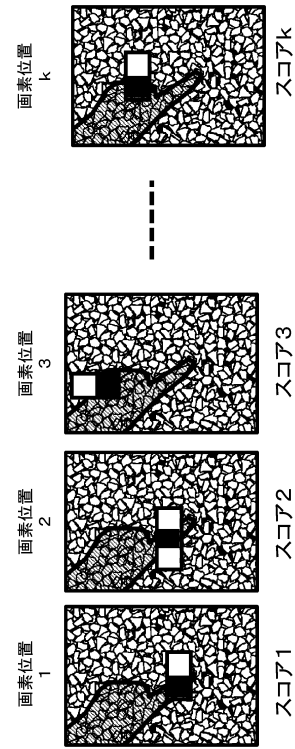


【図 1 2】



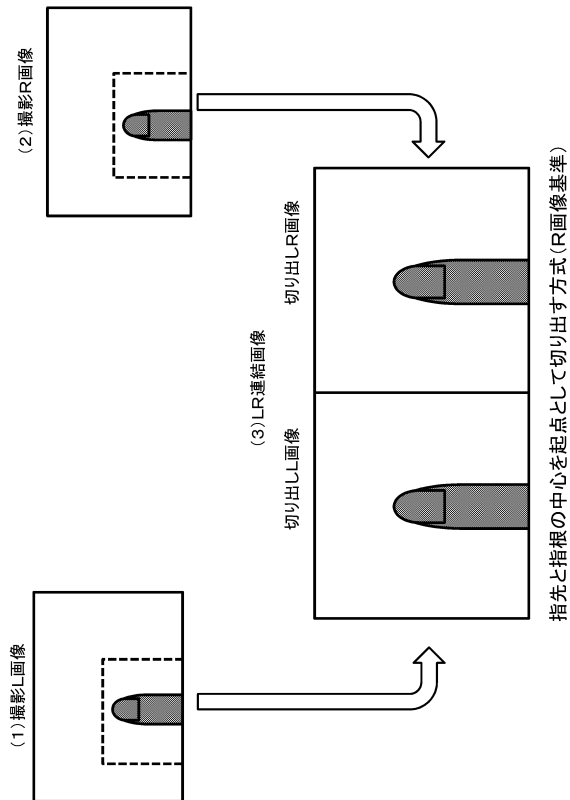
各画素点で、「指か指でないか」の判別(スコアと閾値との比較)を実行しながらスキミングしていく  
スコアが閾値を超えたら検出されたとする

【図 1 3】

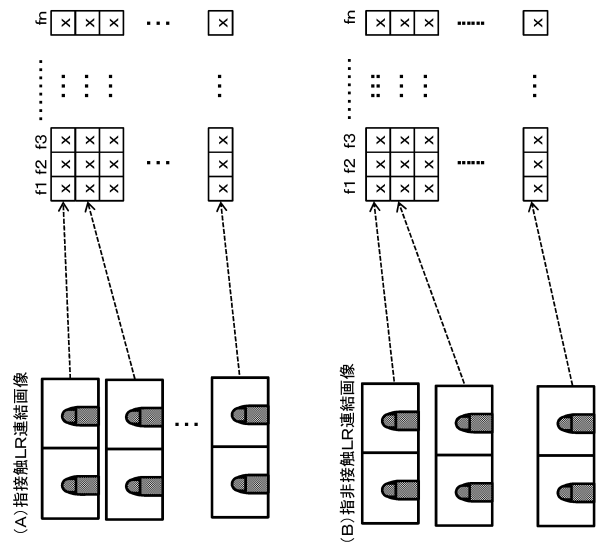


総合スコア = スコア1 + スコア2 + スコア3 + ... + スコアk

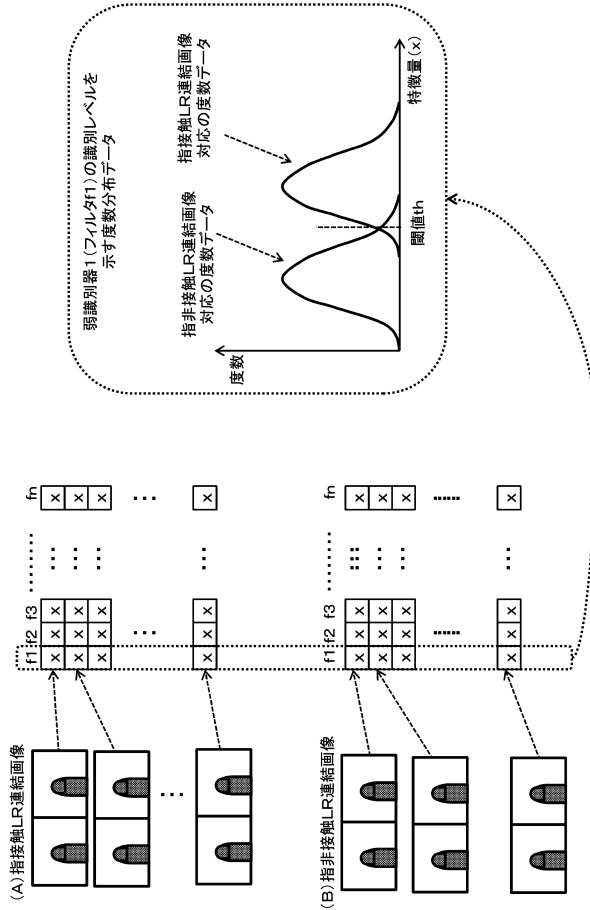
【図 1 5】



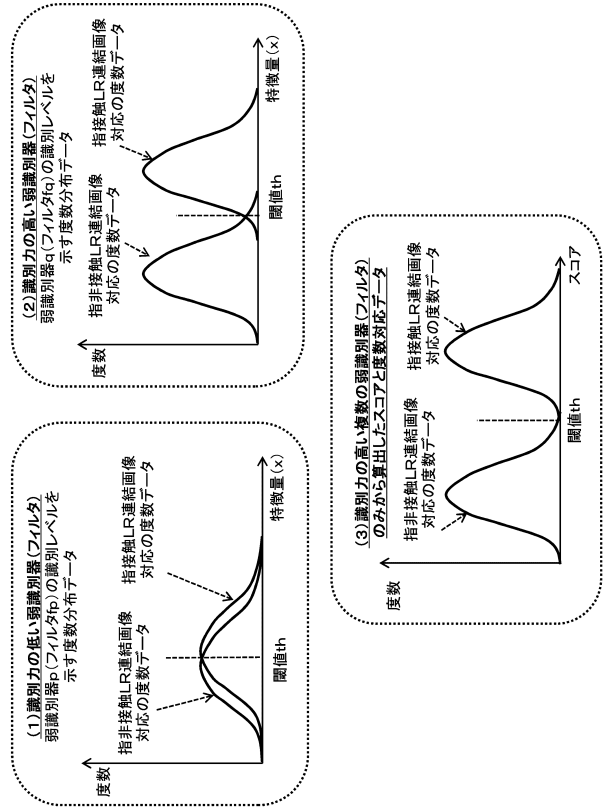
【図 1 7】



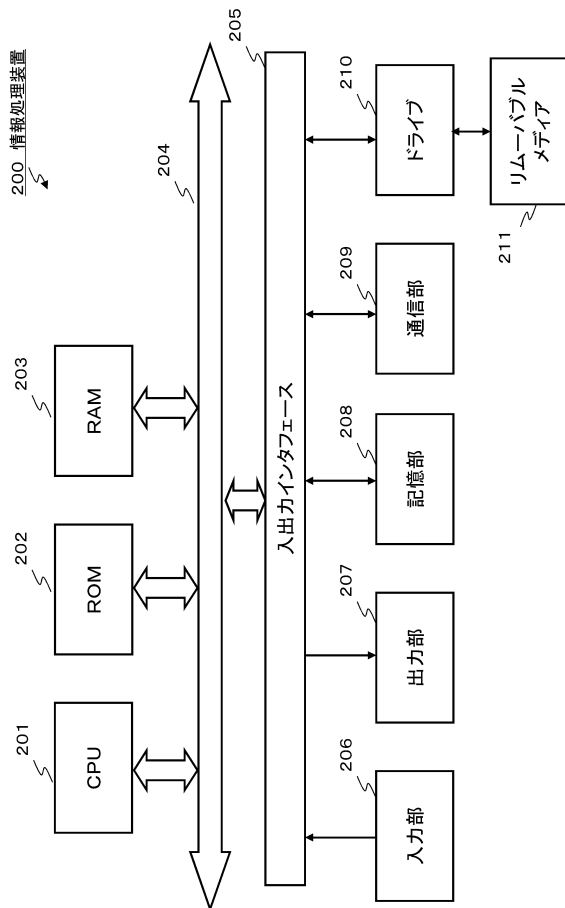
【図 18】



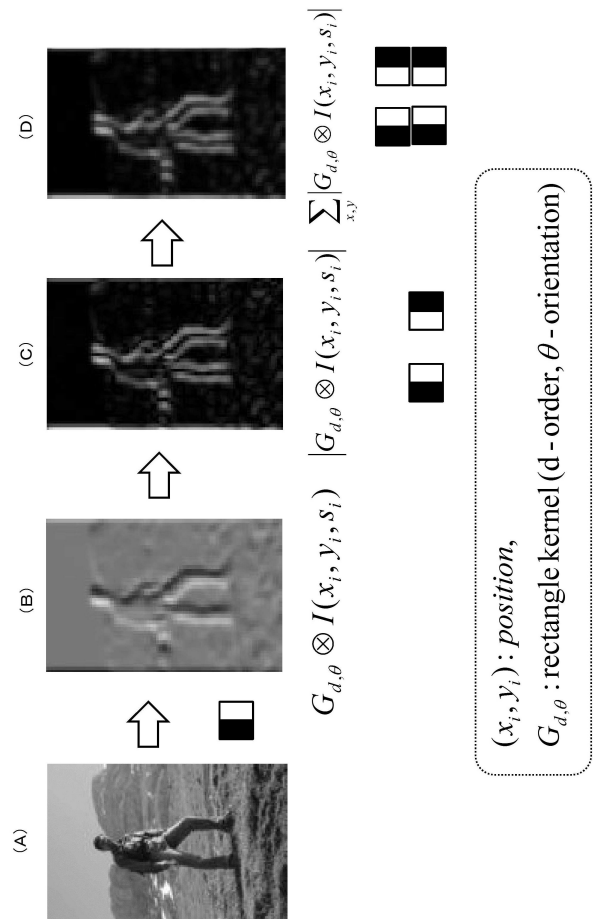
【図 19】



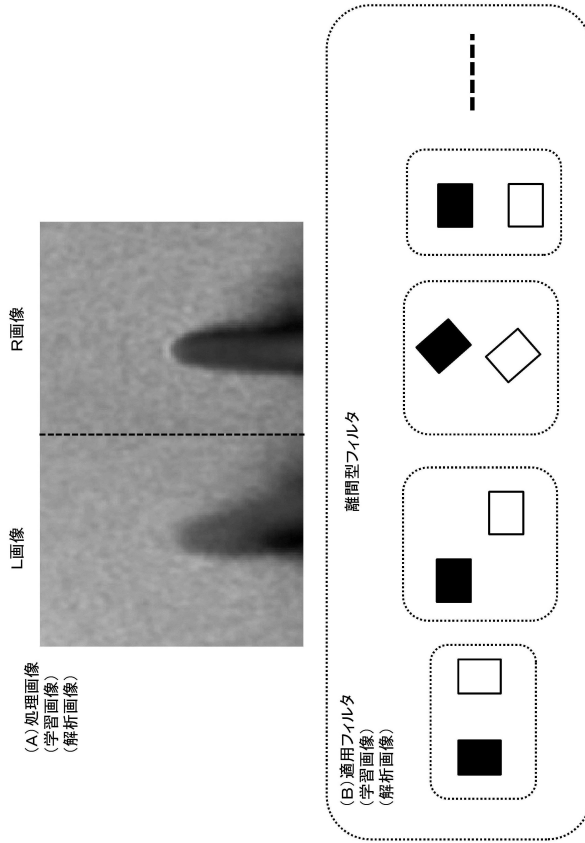
【図 23】



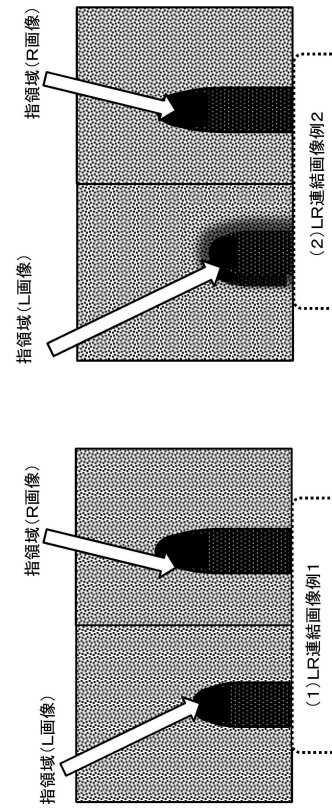
【図 6】



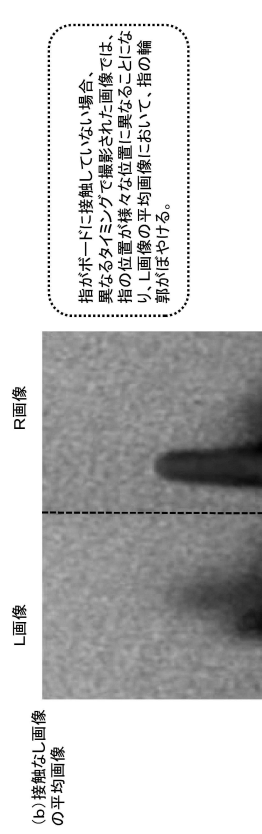
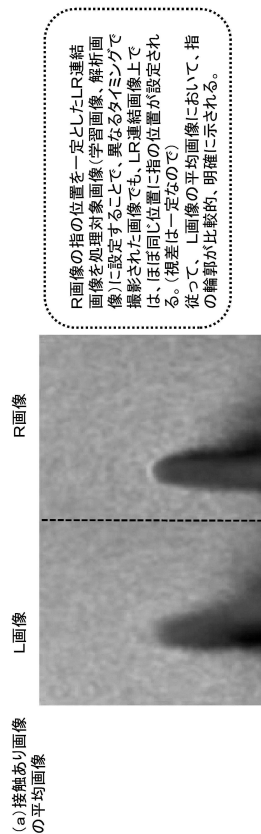
【図 14】



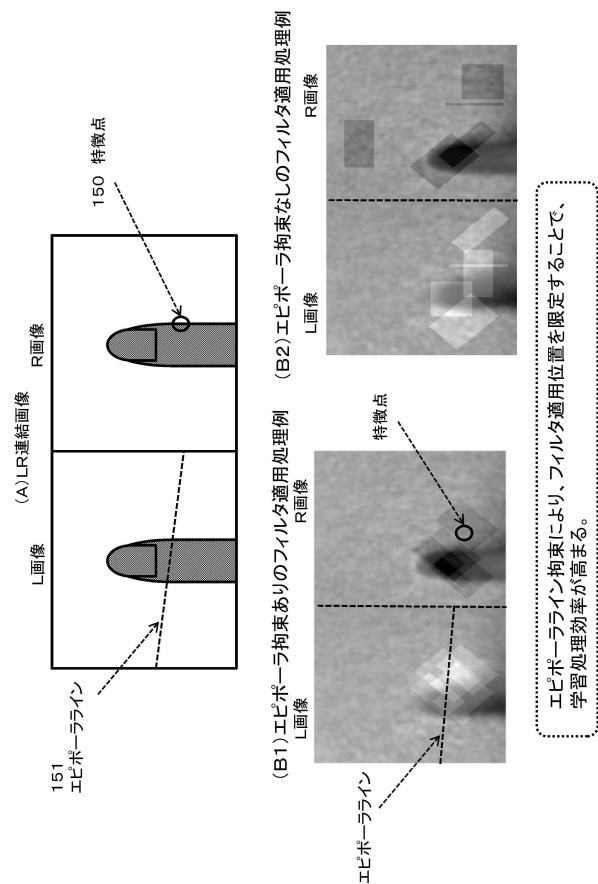
【図 16】



【図 20】


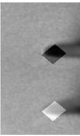

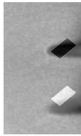

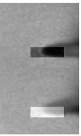

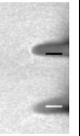

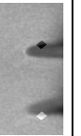


【図 21】





【図 22】

	フィルタの種類	フィルタと 画像との対応例	特徴量データに反映される主な画像特性
(1)	フィルタ1 		視差情報を反映した特徴量が得られる
(2)	フィルタ2 		被写体形状を反映した特徴量データが得られる
(3)	フィルタ3 		視差情報を反映した特徴量が得られる
(4)	フィルタ4 		被写体形状を反映した特徴量データが得られる
(5)	フィルタ5 		被写体形状を反映した特徴量データが得られる

---

フロントページの続き

(72)発明者 横野 順  
東京都港区港南1丁目7番1号 ソニー株式会社内

審査官 佐田 宏史

(56)参考文献 特開2003-028635(JP,A)  
特開2013-029995(JP,A)  
特開2013-161142(JP,A)  
特開2011-118694(JP,A)  
特開2006-047252(JP,A)  
特開2006-338082(JP,A)

(58)調査した分野(Int.Cl., DB名)  
G06T 1/00, 7/00-7/90  
G06F 3/01, 3/0346