

(19) 日本国特許庁(JP)

## (12) 公開特許公報(A)

(11) 特許出願公開番号

特開2012-113460

(P2012-113460A)

(43) 公開日 平成24年6月14日(2012.6.14)

(51) Int.Cl.

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

F 1

G06T 7/00 300B  
G06T 7/20 300A  
G06T 7/00 300F

テーマコード(参考)

5 L096

審査請求 未請求 請求項の数 7 O L (全 40 頁)

(21) 出願番号  
(22) 出願日特願2010-260869 (P2010-260869)  
平成22年11月24日 (2010.11.24)(71) 出願人 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 100082131  
弁理士 稲本 義雄  
(74) 代理人 100121131  
弁理士 西川 孝  
(72) 発明者 長谷川 雄一  
東京都港区港南1丁目7番1号 ソニー株  
式会社内  
(72) 発明者 周寧  
東京都港区港南1丁目7番1号 ソニー株  
式会社内

最終頁に続く

(54) 【発明の名称】情報処理装置および方法、並びにプログラム

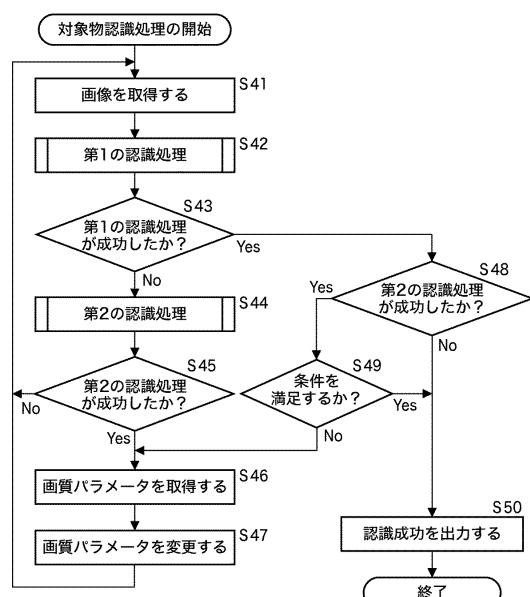
## (57) 【要約】

【課題】より確実に対象物を認識できるようにする。

【解決手段】第1の認識処理は、例えば、手を認識する。第2の認識処理は、例えば、人を認識する。取得した画像が手を認識するのに不適切な画像の場合、例えば、画像上の手が小さいため認識できない場合、第2の認識処理を行い、画像上から人を認識する。人が認識された場合、ズーム率などの画質パラメータが新たに取得される。そして、画質パラメータが変更された画像を取得し、再び第1の認識処理を行う。以上の処理を手が認識されるまで繰り返すことで、確実に手を認識する。本発明は、例えばパーソナルコンピュータに適用できる。

【選択図】図14

図14



**【特許請求の範囲】****【請求項 1】**

画像および画質パラメータを取得する取得手段と、  
前記画像から第1の対象物を認識する第1の認識手段と、  
前記画像から第2の対象物を認識する第2の認識手段と、  
前記第1の対象物と前記第2の対象物の認識に成功したかをそれぞれ判定する判定手段と、  
前記画像を前記画質パラメータに基づく画像に変更するように制御する制御手段とを備え、  
前記取得手段は、前記判定手段により前記第1の対象物の認識に失敗したと判定されるとともに、前記第2の対象物の認識に成功したと判定される場合、新たな前記画質パラメータを取得し、  
前記第1の認識手段は、新たな前記画質パラメータに基づき変更された前記画像から前記第1の対象物を認識する情報処理装置。

**【請求項 2】**

前記判定手段は、前記第1の対象物と前記第2の対象物との相対的な位置および相対的な大きさのうちの少なくとも一方が、予め設定された所定の条件を満足するかを判定する請求項1に記載の情報処理装置。

**【請求項 3】**

前記第1の対象物は、人、顔、および手のうちのいずれか1つであり、  
前記第2の対象物は、前記人、前記顔、および前記手のうちの前記第1の対象物とは異なるいずれか1つである  
請求項2に記載の情報処理装置。

**【請求項 4】**

前記人の姿勢を認識する姿勢認識手段をさらに備え、  
前記判定手段は、前記第1の対象物と前記姿勢認識手段により認識された前記姿勢の相対距離および相対角度のうち少なくとも一方が予め設定された所定の条件を満足するかを判定する

請求項3に記載の情報処理装置。

**【請求項 5】**

前記第1の対象物は、自動車のボディ、タイヤ、およびナンバープレートのうちのいずれか1つであり、

前記第2の対象物は、前記自動車の前記ボディ、前記タイヤ、および前記ナンバープレートのうち前記第1の対象物とは異なるのいずれか1つである

請求項2に記載の情報処理装置。

**【請求項 6】**

画像および画質パラメータを取得する取得ステップと、  
前記画像から第1の対象物を認識する第1の認識ステップと、  
前記画像から第2の対象物を認識する第2の認識ステップと、  
前記第1の対象物と前記第2の対象物の認識に成功したかをそれぞれ判定する判定ステップと、

前記画像を前記画質パラメータに基づく画像に変更するように制御する制御ステップとを含み、

前記取得ステップの処理は、前記判定ステップの処理により前記第1の対象物の認識に失敗したと判定されるとともに、前記第2の対象物の認識に成功したと判定される場合、新たな前記画質パラメータを取得し、

前記第1の認識ステップの処理は、新たな前記画質パラメータに基づき変更された前記画像から前記第1の対象物を認識する

情報処理方法。

10

20

30

40

50

**【請求項 7】**

コンピュータに、  
画像および画質パラメータを取得する取得ステップと、  
前記画像から第1の対象物を認識する第1の認識ステップと、  
前記画像から第2の対象物を認識する第2の認識ステップと、  
前記第1の対象物と前記第2の対象物の認識に成功したかをそれぞれ判定する判定ステップと、

前記画像を前記画質パラメータに基づく画像に変更するように制御する制御ステップと  
を実行させるためのプログラムであって、

前記取得ステップの処理は、前記判定ステップの処理により前記第1の対象物の認識に失敗したと判定されるとともに、前記第2の対象物の認識に成功したと判定される場合、新たな前記画質パラメータを取得し、

前記第1の認識ステップの処理は、新たな前記画質パラメータに基づき変更された前記画像から前記第1の対象物を認識する

プログラム。

**【発明の詳細な説明】****【技術分野】****【0001】**

本発明は情報処理装置および方法、並びにプログラムに関し、特に、より確実に対象物を認識するようにした情報処理装置および方法、並びにプログラムに関する。

20

**【背景技術】****【0002】**

図1に示すように、カメラ11を用いてユーザ1の手2を認識し、手2の動きのパターンによるジェスチャを使用して、テレビジョン受像機12などの電子機器を操作することが行われている。

**【0003】**

また、カメラを用いて撮像する場合、ユーザの顔を認識し、認識された顔の情報に基づいて、ピントや色のバランスなどのパラメータを自動的に制御することが行われている。

**【0004】**

このような処理を実行するための画像認識装置は、例えば、人、顔、手などを認識する（例えば特許文献1参照）。また、画像認識装置の中には、人の姿勢を認識するものもある（例えば特許文献2参照）。

30

**【先行技術文献】****【特許文献】****【0005】**

【特許文献1】特開2010-108475号公報

【特許文献2】特開平11-128535号公報

**【発明の開示】****【発明が解決しようとする課題】****【0006】**

しかしながら、画像認識装置は、取得した画像が適切な画像でない場合、対象物を認識することができないことがあった。

40

**【0007】**

図2と図3は、対象物の認識に適切でない画像の例を示している。図2Aは、明るすぎる画像の例を示している。図2Bは、暗すぎる画像の例を示している。図3Aは、ピントがぼけている画像の例を示している。図3Bは、認識する対象物（例えば、手5や顔6）が小さい例を示している。

**【0008】**

このような場合、画像認識装置は、手5や顔6などの対象物を認識することが困難であった。

50

**【0009】**

本発明は、このような状況に鑑みてなされたものであり、より確実に対象物を認識できるようにするものである。

**【課題を解決するための手段】****【0010】**

本発明の一側面の情報処理装置は、画像および画質パラメータを取得する取得手段と、前記画像から第1の対象物を認識する第1の認識手段と、前記画像から第2の対象物を認識する第2の認識手段と、前記第1の対象物と前記第2の対象物の認識に成功したかをそれぞれ判定する判定手段と、前記画像を前記画質パラメータに基づく画像に変更するように制御する制御手段とを備え、前記取得手段は、前記判定手段により前記第1の対象物の認識に失敗したと判定されるとともに、前記第2の対象物の認識に成功したと判定される場合、新たな前記画質パラメータを取得し、前記第1の認識手段は、新たな前記画質パラメータに基づき変更された前記画像から前記第1の対象物を認識する。

10

**【0011】**

前記判定手段は、前記第1の対象物と前記第2の対象物との相対的な位置および相対的な大きさのうちの少なくとも一方が、予め設定された所定の条件を満足するかを判定することができる。

**【0012】**

前記第1の対象物は、人、顔、および手のうちのいずれか1つであり、前記第2の対象物は、前記人、前記顔、および前記手のうちの前記第1の対象物とは異なるいずれか1つであるようにすることができる。

20

**【0013】**

前記人の姿勢を認識する姿勢認識手段をさらに備え、前記判定手段は、前記第1の対象物と前記姿勢認識手段により認識された前記姿勢の相対距離および相対角度のうち少なくとも一方が予め設定された所定の条件を満足するかを判定することができる。

**【0014】**

前記第1の対象物は、自動車のボディ、タイヤ、およびナンバープレートのうちのいずれか1つであり、前記第2の対象物は、前記自動車の前記ボディ、前記タイヤ、および前記ナンバープレートのうちの前記第1の対象物とは異なるいずれか1つであるようにすることができる。

30

**【0015】**

本発明の一側面の情報処理方法は、画像および画質パラメータを取得する取得ステップと、前記画像から第1の対象物を認識する第1の認識ステップと、前記画像から第2の対象物を認識する第2の認識ステップと、前記第1の対象物と前記第2の対象物の認識に成功したかをそれぞれ判定する判定ステップと、前記画像を前記画質パラメータに基づく画像に変更するように制御する制御ステップとを含み、前記取得ステップの処理は、前記判定ステップの処理により前記第1の対象物の認識に失敗したと判定されるとともに、前記第2の対象物の認識に成功したと判定される場合、新たな前記画質パラメータを取得し、前記第1の認識ステップの処理は、新たな前記画質パラメータに基づき変更された前記画像から前記第1の対象物を認識する。

40

**【0016】**

本発明の一側面のプログラムは、コンピュータに、画像および画質パラメータを取得する取得ステップと、前記画像から第1の対象物を認識する第1の認識ステップと、前記画像から第2の対象物を認識する第2の認識ステップと、前記第1の対象物と前記第2の対象物の認識に成功したかをそれぞれ判定する判定ステップと、前記画像を前記画質パラメータに基づく画像に変更するように制御する制御ステップとを実行させるためのプログラムであって、前記取得ステップの処理は、前記判定ステップの処理により前記第1の対象物の認識に失敗したと判定されるとともに、前記第2の対象物の認識に成功したと判定される場合、新たな前記画質パラメータを取得し、前記第1の認識ステップの処理は、新たな前記画質パラメータに基づき変更された前記画像から前記第1の対象物を認識する。

50

## 【0017】

本発明の一側面においては、画像および画質パラメータが取得され、画像から第1の対象物が認識されるとともに、第2の対象物が認識され、第1の対象物と第2の対象物の認識に成功したかがそれぞれ判定され、画像が画質パラメータに基づく画像に変更するように制御される。そして、第1の対象物の認識に失敗したと判定されるとともに、第2の対象物の認識に成功したと判定される場合、新たな画質パラメータが取得され、新たな画質パラメータに基づき変更された画像から第1の対象物が認識される。

## 【発明の効果】

## 【0018】

本発明の側面によれば、より確実に対象物を認識することが可能になる。

10

## 【図面の簡単な説明】

## 【0019】

【図1】カメラを用いたジェスチャ認識の例を示す図である。

【図2】認識に不適切な画像の例を示す図である。

【図3】認識に不適切な画像の例を示す図である。

【図4】本発明を適用したパーソナルコンピュータのハードウェアの構成例を示すブロック図である。

【図5】CPUの機能的構成例を示すブロック図である。

【図6】第2認識部の詳細な構成例を示すブロック図である。

20

【図7】輪郭特徴量計算部の詳細な構成例を示すブロック図である。

【図8】画像に対してフィルタ処理を行った結果を示す図である。

【図9】認識処理を説明するフローチャートである。

【図10】対象物特徴点の抽出について説明する図である。

【図11】輪郭特徴点の抽出について説明する図である。

【図12】輪郭特徴量計算処理を説明するフローチャートである。

【図13】認識結果の例を示す図である。

【図14】対象物認識処理を説明するフローチャートである。

【図15】認識結果の例を示す図である。

【図16】相対位置の条件の例を示す図である。

30

【図17】相対位置の条件の判定結果の例を示す図である。

【図18】相対的な大きさの比較の条件の例を示す図である。

【図19】相対的な大きさの比較の条件の判定結果の例を示す図である。

【図20】認識結果の例を示す図である。

【図21】姿勢認識部の詳細な構成例を示すブロック図である。

【図22】対象物認識処理の例を説明するフローチャートである。

【図23】認識結果の例を示す図である。

【図24】姿勢認識処理の例を説明するフローチャートである。

【図25】相対距離と相対角度の条件の例を示す図である。

【図26】相対距離の条件の判定結果の例を示す図である。

40

【図27】本発明を適用したテレビジョン受像機制御システムの構成例を示すブロック図である。

【図28】本発明を適用したデジタルカメラ制御システムの構成例を示すブロック図である。

## 【発明を実施するための最良の形態】

## 【0020】

以下、本発明を実施するための形態（以下、実施の形態という）について説明する。なお、説明は以下の順序で行う。

1. パーソナルコンピュータの構成
2. 第2認識部の構成
3. 輪郭特徴量計算部の構成

50

4 . 認識処理  
 5 . 輪郭特徴量計算処理  
 6 . 対象物認識処理 1  
 7 . 対象物認識処理 2  
 8 . 姿勢認識部の構成  
 9 . 対象物認識処理 3  
 10 . 姿勢認識処理  
 11 . テレビジョン受像機制御システムの構成  
 12 . デジタルカメラ制御システムの構成  
 13 . その他

10

【0021】  
 [パーソナルコンピュータの構成]  
 【0022】

図4は、本発明を適用したパーソナルコンピュータ21のハードウェアの構成例を示す  
 ブロック図である。

【0023】

情報処理装置としてのパーソナルコンピュータ21において、CPU(Central Processing Unit)41, ROM(Read Only Memory)42, RAM(Random Access Memory)43は、バス44により相互に接続されている。

20

【0024】  
 バス44には、さらに、入出力インターフェース45が接続されている。入出力インターフェース45には、入力部46、出力部47、記憶部48、通信部49、およびドライブ50が接続されている。

【0025】

入力部46は、キーボード、マウス、マイクロフォンなどの他、カメラ46Aを有する。  
 出力部47は、スピーカなどの他、ディスプレイ47Aを有する。

【0026】

カメラ46Aは、ディスプレイ47Aを見るユーザを撮影できる位置に配置されている。  
 ディスプレイ47Aは、各種のアプリケーションに対応する画像の他、必要に応じてカ  
 メラ46Aにより撮影された画像を表示する。

30

【0027】

記憶部48は、ハードディスクや不揮発性のメモリなどよりなる。通信部49は、ネット  
 ワークインターフェースなどよりなる。ドライブ50は、磁気ディスク、光ディスク、光  
 磁気ディスク、或いは半導体メモリなどのリムーバブルメディア51を駆動する。

【0028】

以上のように構成されるパーソナルコンピュータ21においては、CPU41が、例えば  
 、記憶部48に記憶されているプログラムを、入出力インターフェース45及びバス44を  
 介して、RAM43にロードして実行することにより、各種の処理が行われる。

【0029】

CPU41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51に記録して提供される。

40

【0030】

なお、パッケージメディアとしては、磁気ディスク(フレキシブルディスクを含む)、  
 光ディスク(CD-ROM(Compact Disc-Read Only Memory), DVD(Digital Versatile Disc)等)  
 、光磁気ディスク、もしくは半導体メモリなどが用いられる。

【0031】

また、プログラムは、ローカルエリアネットワーク、インターネット、デジタル衛星放  
 送といった、有線または無線の伝送媒体を介して提供することができる。

【0032】

パーソナルコンピュータ21においては、プログラムは、リムーバブルメディア51を

50

ドライブ 50 に装着することにより、入出力インターフェース 45 を介して、記憶部 48 にインストールすることができる。

【0033】

また、プログラムは、有線または無線の伝送媒体を介して、通信部 49 で受信し、記憶部 48 にインストールすることができる。その他、プログラムは、ROM 42 や記憶部 48 に、あらかじめインストールしておくことができる。

【0034】

図 5 は、CPU 41 の機能的構成を示すブロック図である。

【0035】

CPU 41 は、取得部 61、第 1 認識部 62、第 2 認識部 63、判定部 64、制御部 65、出力部 66、およびアプリケーション実行部 67 の機能ブロックを有している。なおCPU 41 の各ブロックは、必要に応じて相互に信号、データを授受することができる。

【0036】

取得部 61 は、各種の情報を取得する。第 1 認識部 62 および第 2 認識部 63 は、手、顔、人等、それぞれ異なる認識対象物の認識処理を行う。例えば、第 1 認識部 62 は認識対象物として手を認識し、第 2 認識部 63 は認識対象物として人を認識する。

【0037】

判定部 64 は、各種の情報を判定する。制御部 65 は、各種の制御処理を行う。出力部 66 は、各種の情報を出力する。アプリケーション実行部 67 は、アプリケーションを実行する。

【0038】

アプリケーションとしては、例えば、画像上の手のジェスチャを認識して、そのジェスチャに対応するコマンドを実行するアプリケーション、または画像上の顔を認識して顔の領域の画質を向上させるアプリケーションなどがある。

【0039】

すなわちアプリケーション実行部 67 は、画像上の対象物を認識し、その認識結果を用いて所定の処理を行うアプリケーションを実行する。

【0040】

[ 第 2 認識部の構成 ]

【0041】

手を認識する第 1 認識部 62 は、人を認識する第 2 認識部 63 と同様の構成を有している。ここでは代表して第 2 認識部 63 の構成について説明する。図 6 は、第 2 認識部 63 のより詳細な構成例を示すブロック図である。

【0042】

第 2 認識部 63 は、ピラミッド画像生成部 81、対象物特徴点抽出部 82、対象物特徴量計算部 83、輪郭特徴点抽出部 84、輪郭特徴量計算部 85、識別計算部 86、および認識結果出力部 87 から構成される。

【0043】

各部は必要に応じて相互に信号、データを授受することができる。このことは、後述する図 6、図 7、および図 21 においても同様である。

【0044】

ピラミッド画像生成部 81 は、取得された画像から、互いに解像度の異なる複数の画像をピラミッド画像として生成する。

【0045】

例えば、レベル L1 乃至レベル L8 までの 8 つの解像度の階層のピラミッド画像が生成される。レベル L1 のピラミッド画像が最も解像度が高く、レベル L1 からレベル L8 まで順番にピラミッド画像の解像度が低くなる。

【0046】

対象物特徴点抽出部 82 は、ピラミッド画像生成部 81 で生成されたピラミッド画像を

10

20

30

40

50

構成する各画像から、その画像の画素のいくつかを、対象物を認識するときに用いられる対象物特徴点として抽出する。

【0047】

対象物特徴量計算部83は、対象物特徴点抽出部82により抽出された各対象物特徴点について、その対象物特徴点と、他の対象物特徴点とを1つのペアとするペアリングを行う。

【0048】

また、対象物特徴量計算部83は、対象物特徴点が抽出された画像に基づいて、対象物特徴点のペア毎に、任意の2つの領域のテクスチャの距離を示す対象物特徴量を計算する。

10

【0049】

輪郭特徴点抽出部84は、ピラミッド画像生成部81で生成されたピラミッド画像を構成する各画像から、その画像の画素のいくつかを、対象物を認識するときに用いられる輪郭特徴点として抽出する。

【0050】

輪郭特徴量計算部85は、輪郭特徴点が抽出された画像に基づいて、例えばステアラブルフィルタ(Steerable Filter)を用いたフィルタ処理により、輪郭特徴点毎に、抽出された輪郭を示す輪郭特徴量を計算する。

20

【0051】

識別計算部86は、予め記憶部48に記憶されている識別用特徴量および統合識別器を読み出す。また、識別計算部86は、対象物特徴量計算部83により計算された対象物特徴量、および輪郭特徴量計算部85により計算された輪郭特徴量のうちの識別用特徴量に対応するものを、読み出した統合識別器に代入して演算を行う。

20

【0052】

ここで、統合識別器とは、対象物識別器と輪郭識別器とを統合して生成されたものである。

【0053】

対象物識別器および輪郭識別器は、予め統計学習により生成された、複数の弱識別器からなる強い識別器であり、対象物(例えば人)の特徴を利用して、入力された画像中に対象物の画像の領域が存在するか否かを識別するときに用いられる。

30

【0054】

また、識別用特徴量は、統合識別器を用いて対象物体の認識を行うときに用いられる対象物特徴点のペアの対象物特徴量、および輪郭特徴点の輪郭特徴量である。

【0055】

認識結果出力部87は、識別計算部86における演算結果に基づいて、対象物体が入力画像で認識されたか否かの識別結果を出力する。

【0056】

上述したように、第1認識部62は、第2認識部63と同様の構成を有し、同様の処理を行うので、その詳細な説明は省略する。

40

【0057】

[輪郭特徴量計算部の構成]

【0058】

図7は、図6の輪郭特徴量計算部85のより詳細な構成例を示すブロック図である。輪郭特徴量計算部85は、1次フィルタ処理部101、2次フィルタ処理部102、3次フィルタ処理部103、および特徴量生成部104から構成される。

【0059】

輪郭特徴点抽出部84により輪郭特徴点が抽出された画像は、1次フィルタ処理部101乃至特徴量生成部104に供給され、輪郭特徴点は、1次フィルタ処理部101乃至3次フィルタ処理部103に供給される。

50

【0060】

1次フィルタ処理部101は、輪郭特徴点抽出部84により抽出された輪郭特徴点毎に、輪郭特徴点に対してガウス関数Gの1次微分関数G<sub>1</sub>によりフィルタ処理を施して特徴量を抽出する。ここで、ガウス関数G、および1次微分関数G<sub>1</sub>は、次式(1)および式(2)により示される。

【0061】

【数1】

$$G = e^{-\frac{x^2+y^2}{2\sigma^2}} \quad \cdots(1)$$

【0062】

【数2】

$$G_1(\theta) = \cos(\theta) G_1(0^\circ) + \sin(\theta) G_1(90^\circ) \quad \cdots(2)$$

【0063】

式(1)において、 $\sigma$ はガウス幅を示している。式(2)において、 $\theta$ は任意の角度を示している。

【0064】

例えば、1次フィルタ処理部101は、ガウス関数Gのガウス幅 $\sigma$ を3つの所定値(例えば、ガウス幅1, 2, 3 = 1, 2, 4)に変化させ、ガウス幅 $\sigma$ 毎に所定の4方向(例えば、 $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ )について式(2)を計算する。

【0065】

なお、方向 $\theta$ は4方向に限らず、8方向、例えば $\theta$ を8方向に等分したときの各方向などとしてもよい。また、従来は、複数のガウス幅を用いて処理を行っていたが、本実施の形態においては、後述するように、ガウス幅は1つだけ用意しておけば良い。

【0066】

換言すれば、ガウス幅を変化させる必要がない。よって、上記では、「ガウス幅を3つの所定値に変化させ、ガウス幅 $\sigma$ 毎に所定の4方向について式(2)を計算する」と記載したが、本実施の形態においては、設定されているガウス幅 $\sigma$ において所定の方向の4方向について式(2)を計算するだけでよい。

【0067】

よって、複数のガウス幅毎に計算する必要がないため、計算量を低減させることが可能となる。このようなことは、他のフィルタ、例えば、2次フィルタ処理部102、3次フィルタ処理部103においても同様である。

【0068】

2次フィルタ処理部102は、抽出された輪郭特徴点毎に、輪郭特徴点に対してガウス関数Gの2次微分関数G<sub>2</sub>によりフィルタ処理を施して特徴量を抽出する。次式(3)は、2次微分関数G<sub>2</sub>を示しており、式(3)において $\theta$ は任意の角度を示している。

【0069】

【数3】

$$G_2(\theta) = k_{21}(\theta) G_2(0^\circ) + k_{22}(\theta) G_2(60^\circ) + k_{23}(\theta) G_2(120^\circ) \quad \cdots(3)$$

【0070】

また、式(3)における係数k<sub>2i</sub>( $\theta$ )(但し、 $i = 1, 2, 3$ )は、次式(4)で示される関数である。

【0071】

【数4】

$$k_{2i}(\theta) = \frac{1}{3} \{1 + 2\cos(2(\theta - \theta_i))\} \quad \cdots(4)$$

【0072】

10

20

30

40

50

例えば、2次フィルタ処理部102は、ガウス関数Gの所定のガウス幅において、所定の4方向（例えば、=1, 2, 3, 4）について式（3）を計算する。

【0073】

3次フィルタ処理部103は、抽出された輪郭特徴点毎に、輪郭特徴点に対してガウス関数Gの3次微分関数G<sub>3</sub>によりフィルタ処理を施して特徴量を抽出する。

【0074】

次式（5）は、3次微分関数G<sub>3</sub>を示しており、式（5）においてθは任意の角度を示している。

【0075】

【数5】

$$G_3(\theta) = k_{31}(\theta)G_3(0^\circ) + k_{32}(\theta)G_3(45^\circ) + k_{33}(\theta)G_3(90^\circ) + k_{34}(\theta)G_3(135^\circ)$$

…(5)

【0076】

また、式（5）における係数k<sub>3i</sub>（）（但し、i=1, 2, 3）は、次式（6）で示される関数である。

【0077】

【数6】

$$k_{3i}(\theta) = \frac{1}{4} \{ 2\cos(\theta - \theta_i) + 2\cos(3(\theta - \theta_i)) \} \quad \cdots(6)$$

20

【0078】

例えば、3次フィルタ処理部103は、ガウス関数Gの所定のガウス幅において、所定の4方向（例えば、=1, 2, 3, 4）について、式（5）を計算する。

【0079】

特徴量生成部104は、1次フィルタ処理部101、2次フィルタ処理部102、および3次フィルタ処理部103のそれぞれで、4つの方向について計算された各輪郭特徴点の特徴量の供給を受け、供給された合計12個（=3（次数）×4（方向））の特徴量を並べて輪郭特徴点における輪郭特徴量とする。

30

【0080】

また、各フィルタ処理部101, 102, 103には、ピラミッド画像生成部81から異なる解像度の複数の画像が供給されるため、各画像から4つの方向について計算された各輪郭特徴点の特徴量も供給される。

【0081】

この供給される特徴量は、ピラミッド画像生成部81が生成する画像の枚数に依存し、例えば、レベル1からレベル8までの8枚の画像が生成される場合、8枚分の4つの方向について計算された各輪郭特徴点の特徴量が供給されることになる。

【0082】

このように、輪郭特徴量計算部85では、ガウス関数を微分して得られる、方向に選択性を持つフィルタ（基底関数）が用いられて、微分の次数毎に異なる特徴量（輪郭）が抽出され、輪郭特徴量とされている。

40

【0083】

図8に、ガウス関数Gの微分関数により、人が写っている画像に対してフィルタ処理を行うことで得られた結果を示す。図8において、左側にはフィルタ処理の対象となる画像が示され、中央にはフィルタが示され、右側にはフィルタ処理後の画像が示されている。

【0084】

図8の左側に示した画像は、ピラミッド画像を構成する2枚の画像であり、異なる解像度の画像である。ピラミッド画像生成部81は、例えば、レベルL1乃至L8までの異なる解像度の画像を生成する。そのうちの例えば、レベルL1の画像131とレベルL2の

50

画像 132 が、図 8 の左側に示されている。

【0085】

図 8 の中央に示したフィルタは、1 次微分関数  $G_1$ 、2 次微分関数  $G_2$ 、および 3 次微分関数  $G_3$  のガウス幅  $\sigma$  が  $= 1$  のフィルタの一例を示している。

【0086】

図 8 の左側に示した画像 131, 132 に対して、図 8 の中央に示したフィルタのうち、例えば、3 次微分関数  $G_3$  のフィルタを用いて、フィルタ処理した場合、図 8 の右側に示した画像が生成される。

【0087】

すなわち、画像 131 に対して、3 次微分関数  $G_3$  のフィルタでフィルタ処理を行うと、画像 141-1 乃至 141-4 が生成される。また、画像 132 に対して、3 次微分のフィルタでフィルタ処理を行うと、画像 142-1 乃至 142-4 が生成される。

【0088】

画像 141-1 乃至 141-4 は、それぞれガウス幅  $\sigma$  が 1 ( $\sigma = 1$ ) の 3 次微分関数  $G_3$  のフィルタを用いて、フィルタ処理を行ったため、ガウス幅  $\sigma$  が 1 のフィルタでフィルタ処理が行われたときの画像となる。

【0089】

同様に、画像 142-1 乃至 142-4 は、それぞれガウス幅  $\sigma$  が 1 ( $\sigma = 1$ ) の 3 次微分関数  $G_3$  のフィルタを用いて、フィルタ処理を行ったため、ガウス幅  $\sigma$  が 1 のフィルタでフィルタ処理が行われたときの画像となる。

【0090】

しかしながら、画像 142-1 乃至 142-4 は、画像 132 をフィルタ処理した画像である。画像 132 は、画像 131 を縮小した画像である。

【0091】

このような場合、画像 142-1 乃至 142-4 は、ガウス幅  $\sigma$  が 2 ( $\sigma = 2$ ) の 3 次微分関数  $G_3$  のフィルタを用いてフィルタ処理を行った結果、生成される画像に相当する画像とすることができます。

【0092】

換言すれば、画像 131 に対して、ガウス幅  $\sigma$  が 2 ( $\sigma = 2$ ) の 3 次微分関数  $G_3$  のフィルタを用いてフィルタ処理したときに生成される画像に相当する画像が、画像 142-1 乃至 142-4 である。

【0093】

すなわち、ピラミッド画像に対して、フィルタ処理を行うことで、異なるガウス幅のフィルタを用いて、フィルタ処理したときと同等の画像を得ることが可能となる。

【0094】

例えば、ガウス幅毎にフィルタを予め用意し、フィルタ処理を行うときの演算量や処理負担と比較し、上記したように 1 つのガウス幅のフィルタを予め用意し、ピラミッド画像にフィルタ処理を行うときの演算量や処理負担は大幅に軽減されたものとなる。

【0095】

すなわち、ピラミッド画像を生成し、1 つのガウス幅でフィルタ処理を行うことで、処理時間を大幅に短縮することが可能となる。

【0096】

よって、このような手法で、画像から人などの対象物を検出するようにした場合、処理時間が短くなることで、リアルタイムに画像から対象物を検出することが可能となる。

【0097】

輪郭特徴量計算部 85 は、このようにして、異なるガウス幅  $\sigma$  のフィルタでフィルタ処理したときに相当する複数の画像を、さらに平均して得られる画像を生成する。その生成された平均の画像からは、人の輪郭を確認することができ、各フィルタを用いたフィルタ処理によって、画像から適切に人の輪郭が抽出される。

【0098】

10

20

30

40

50

## [認識処理]

## 【0099】

次に、図9乃至図13を参照して、対象物を認識する第1認識部62および第2認識部63の認識処理を説明する。

## 【0100】

なお、手を認識対象物とする第1認識部62の認識処理と人を認識対象物とする第2認識部63の認識処理とは認識対象物が異なるだけであり、処理は同様であるので、簡単のため人を認識対象物とする第2認識部61の認識処理についてだけ説明する。

## 【0101】

図9は、認識処理を説明するフローチャートである。図9の処理はカメラ46Aにより10画像が取得された時に実行される。

## 【0102】

ステップS1において、第2認識部63のピラミッド画像生成部81は、カメラ46Aにより取得された画像から、ピラミッド画像を生成する。

## 【0103】

上記したように、ピラミッド画像生成部81は、例えば、レベルL1乃至レベルL8までの8つの解像度の階層のピラミッド画像を生成する。

## 【0104】

対象物特徴点抽出部82と輪郭特徴点抽出部84は、それぞれ、供給されるピラミッド画像（異なる解像度の複数の画像）のうちの1つの画像を、処理対象の画像として、ステップS2以下の処理を実行し、複数の画像毎に繰り返しステップS2以下の処理を実行する。

## 【0105】

ステップS2において、対象物特徴点抽出部82は、処理対象とされた画像から対象物特徴点を抽出する。

## 【0106】

ステップS3において、対象物特徴量計算部83は、ステップS2で抽出された対象物特徴点から、未処理の対象物特徴点の1つを注目点として選択する。

## 【0107】

ステップS4において、対象物特徴量計算部83は、対象物特徴点抽出部82により抽出された対象物特徴点とその画像とに基づいて、各対象物特徴点について、対象物特徴点のペアリングを行う。

## 【0108】

ステップS5において、対象物特徴量計算部83は、ペアリングによりペアとされた対象物特徴点の各ペアについて対象物特徴量を計算する。

## 【0109】

例えば、図10に示す画像が対象物特徴点抽出部82により処理対象とされた場合、対象物特徴点抽出部82は、予め定められたマージン（のりしろ）と、サンプリングスキップ数とに基づいて、画像から対象物特徴点145を抽出する。

## 【0110】

なお、図10において、画像上の円は、対象物特徴点145とされた画素を示している。また便宜上、符号は、画像上の1つの円にのみ付されている。

## 【0111】

ここで、マージンとは、画像において、画像の端から対象物特徴点145の抽出の対象となる領域までの画素数である。また、サンプリングスキップ数とは、対象物特徴点145とされる画像上の画素と画素との間隔である。

## 【0112】

したがって、例えばマージンが5画素であり、サンプリングスキップ数が5画素である場合、対象物特徴点抽出部82は、画像の端から5画素以内の位置にある画素からなる領域を除外し、残りの内側の領域E11を対象物特徴点の抽出の対象とする。

10

20

30

40

50

## 【0113】

そして、対象物特徴点抽出部82は、領域E11内の画素のうち、互いに5画素だけ離れている位置の画素を対象物特徴点145として抽出する。

## 【0114】

すなわち、図10の縦方向または横方向に互いに隣り合う対象物特徴点145間の距離は5画素とされ、各対象物特徴点145は、領域E11内の画素とされる（ステップS2, S3の処理）。

## 【0115】

次に、対象物特徴量計算部83は、予め定められた最小半径および最大半径に基づいて、各対象物特徴点145のペアリングを行う。

10

## 【0116】

例えば、最小半径がR11であり、最大半径がR12である場合、所定の対象物特徴点KT1に注目したとき、対象物特徴量計算部83は、対象物特徴点KT1からの距離が、最小半径R11以上であり、かつ最大半径R12以内である全ての対象物特徴点145について、その対象物特徴点145と、対象物特徴点KT1とを1つのペアとする。

## 【0117】

例えば、対象物特徴点KT1からの距離が、最小半径R11以上であり、かつ最大半径R12以内である対象物特徴点145がN個存在する場合、N個の対象物特徴点145のペアが得られることになる。

20

## 【0118】

対象物特徴量計算部83は、全ての対象物特徴点145について、他の対象物特徴点145とのペアリングを行う（ステップS4の処理）。

## 【0119】

さらに、対象物特徴量計算部83は、ペアリングにより得られた対象物特徴点145の各ペアについて、ペアとなる各対象物特徴点145を中心とする所定の形の所定の大きさの領域同士のテクスチャの距離を対象物特徴量として計算する。

## 【0120】

例えば、図10に示す対象物特徴点KT1と対象物特徴点KT2とのペアについての対象物特徴量をSSD(Sum of Square Distance)により求める場合、対象物特徴量計算部83は、対象物特徴点KT1を中心とする所定の領域を領域Tx1とし、対象物特徴点KT2を中心とする領域Tx1と同じ大きさの領域を領域Tx2とする。

30

## 【0121】

そして、対象物特徴量計算部83は、領域Tx1内の画素の画素値と、その画素に対応する領域Tx2内の画素の画素値との差分の絶対値和を求め、求められた差分の絶対値和を対象物特徴量とする。

## 【0122】

なお、対象物特徴量は、SSDに限らず、SAD(Sum of Absolute Distance)や、正規化相関などとされてもよい（ステップS5の処理）。

## 【0123】

図9に戻り、ステップS6において、対象物特徴量計算部83は、全ての対象物特徴点について処理が終了したかを判定する。

40

## 【0124】

ステップS6において、全ての対象物特徴点145についてまだ処理が終了していないと判定された場合、処理はステップS3に戻り、次の対象物特徴点145が注目点として選択され、同様の処理が実行される。

## 【0125】

一方、ステップS6において、全ての対象物特徴点145について処理が終了したと判定された場合、処理はステップS7に進む。

## 【0126】

ステップS7において、輪郭特徴点抽出部84は、ステップS1で生成され、処理対象

50

とされた画像から輪郭特徴点を抽出する。

【0127】

例えば、輪郭特徴点抽出部84に図11Aに示す画像が入力された場合、輪郭特徴点抽出部84は、図11Bに示すように、画像において所定の間隔で並んでいる画素を、輪郭特徴点146として抽出する。

【0128】

なお、図11Bにおいて、画像上の円は輪郭特徴点146とされた画素を表している。また便宜上、符号は、画像上の1つの円にのみ付されている。

【0129】

図11Bに示す画像は、図中、横方向に32画素、縦方向に64画素からなる画像であり、輪郭特徴点抽出部84は、画像上の画素を、横方向および縦方向に2画素おきに輪郭特徴点146とする画素として選択する。

【0130】

これにより、横方向に12画素、縦方向に28画素、合計336(=12×28)画素が輪郭特徴点146として選択される。

【0131】

輪郭特徴点抽出部84は、画像から輪郭特徴点146を抽出すると、抽出した輪郭特徴点146と、入力された画像とを輪郭特徴量計算部85に供給する。

【0132】

ステップS8において、輪郭特徴量計算部85は、輪郭特徴量計算処理を行い、輪郭特徴点抽出部84から供給された輪郭特徴点146および画像に基づいて、各輪郭特徴点146の輪郭特徴量を計算する。

【0133】

[輪郭特徴量計算処理]

【0134】

ここで、図12のフローチャートを参照して、図9のステップS8の処理に対応する輪郭特徴量計算処理について説明する。

【0135】

ステップS21において、輪郭特徴量計算部85、より詳細には、輪郭特徴量計算部85の1次フィルタ処理部101、2次フィルタ処理部102、および3次フィルタ処理部103は、それぞれ輪郭特徴点抽出部84から供給されてきた輪郭特徴点のうち、未処理の輪郭特徴点の1つを注目画素として選択する。

【0136】

ステップS22において、輪郭特徴量計算部85は、方向qを示すカウンタqを1とする。これにより、方向qは1とされる。

【0137】

ステップS23において、1次フィルタ処理部101は、1次フィルタ処理を行う。

【0138】

すなわち、1次フィルタ処理部101は、処理対象となる注目画素の画素値に基づいて、ガウス幅を=1とし、かつ方向をqとして式(2)を演算し、フィルタ処理した結果を特徴量生成部104に供給する。

【0139】

すなわち、式(2)における方向がqとされて演算が行われ、輪郭が抽出される。

【0140】

なお、「ガウス幅を=1として」と記述したが、本実施の形態の場合、ガウス幅は、=1と固定されている(予め1つのガウス幅のフィルタが設定されている)ため、この「ガウス幅を=1として」という処理は省略することが可能である。

【0141】

すなわち、本実施の形態においては、ガウス幅が1のフィルタの方向をqとして式(2)を演算するという処理が、ステップS23において実行されることになる。

10

20

30

40

50

## 【0142】

また、ここでは、ガウス幅  $\sigma$  を  $\sigma = 1$  として説明を続けるが、予め用意されているフィルタのガウス幅は、 $\sigma = 1$  以外のガウス幅でも勿論良い。

## 【0143】

ステップ S 2 4 において、2 次フィルタ処理部 102 は、2 次フィルタ処理を行う。

## 【0144】

すなわち、2 次フィルタ処理部 102 は、注目画素の画素値に基づいて、ガウス幅  $\sigma = 1$  のフィルタの方向を  $q$  として式(3)を演算し、フィルタ処理した結果を特徴量生成部 104 に供給する。

## 【0145】

すなわち、式(3)における方向  $q$  が  $q$  とされて演算が行われ、輪郭が抽出される。

## 【0146】

ステップ S 2 5 において、3 次フィルタ処理部 103 は、3 次フィルタ処理を行う。

## 【0147】

すなわち、3 次フィルタ処理部 103 は、注目画素の画素値に基づいて、ガウス幅  $\sigma = 1$  のフィルタの方向を  $q$  として式(5)を演算し、フィルタ処理した結果を特徴量生成部 104 に供給する。

## 【0148】

すなわち、式(5)における方向  $q$  が  $q$  とされて演算が行われ、輪郭が抽出される。

## 【0149】

ステップ S 2 6 において、輪郭特徴量計算部 85 は、方向  $q$  が 4 であるか否か、すなわちカウンタ  $q = 4$  であるか否かを判定する。ステップ S 2 6 において、方向  $q$  が 4 でないと判定された場合、ステップ S 2 7 において、輪郭特徴量計算部 85 は、カウンタ  $q$  をインクリメントする。

## 【0150】

例えば、カウンタ  $q = 1$  であった場合、カウンタ  $q$  が 1 だけインクリメントされて  $q = 2$  とされ、これにより方向  $q$  は 2 とされる。カウンタ  $q$  が 1 だけインクリメントされると、処理はステップ S 2 3 に戻り、上述した処理が繰り返される。

## 【0151】

これに対して、ステップ S 2 6 において、方向  $q$  が 4 であると判定された場合、ステップ S 2 8 において、特徴量生成部 104 は、1 次フィルタ処理部 101、2 次フィルタ処理部 102、および 3 次フィルタ処理部 103 から供給された演算結果を輪郭特徴量として合成し、1 つの輪郭特徴点に対する輪郭特徴量を生成する。

## 【0152】

輪郭特徴量は、以下の式(7)または式(8)で求められる。

## 【0153】

## 【数7】

$$v_j = \sum_{x,y} |G_{d,\theta} \otimes I(x_i, y_i, s_i)| \quad \cdots (7)$$

## 【0154】

## 【数8】

$$v_j = \max_{x,y} |G_{d,\theta} \otimes I(x_i, y_i, s_i)| \quad \cdots (8)$$

## 【0155】

式(7)、式(8)において、 $G_{d,\theta}$  は、式(2)などと同じく、任意の角度  $\theta$  におけるガウス関数  $G$  の  $d$  次微分関数である。また、 $I(x_i, y_i, s_i)$  のうち、 $(x_i, y_i)$  は、処理対象とされている輪郭特徴点の画像内での座標を表し、 $(s_i)$  は、ピラミッド画像を構成する画像のうち、処理対象とされている画像のスケールを表す。

10

20

30

40

50

## 【0156】

式(7)は、任意の角度におけるガウス関数Gのd次微分関数と輪郭特徴量を畳込み演算し、その絶対値を $\max$ で総和を演算する式である。式(8)は、任意の角度におけるガウス関数Gのd次微分関数と輪郭特徴量を畳込み演算し、その絶対値を $\max$ で最大値をとる式である。

## 【0157】

式(7)と式(8)は、ともに、特徴量を算出する式であるが、式(7)は、局所的なエネルギーを計算する式であり、式(8)は、局所的な最大値を計算する式である。ここで、この式の意味について説明を加える。

## 【0158】

上記したような処理により、任意の角度における関数とスケールで抽出されたフィルタ係数を特徴量として、人などの対象物を検出する検出識別器を生成できる。

10

## 【0159】

しかしながら、この検出識別器では、例えば、人の着ている服装と背景の関係に依存する特徴量となってしまう。

## 【0160】

また、人のように歪みや変形の大きな認証対象に関しては、特徴量として選択性がありすぎる。よって、これらのこととを吸収して処理する必要があり、それぞれの特徴量を不变性のある特徴量にする必要がある。

20

## 【0161】

「人の着ている服装と背景に依存する特徴量」を、不变性のある特徴量にするには、フィルタ処理後の出力値の絶対値を演算することで解決することができる。絶対値を演算することで、人の輪郭に近い特徴量が抽出できる。

## 【0162】

さらに本実施の形態においては、1次微分関数、2次微分関数、さらに3次微分関数を演算し、それぞれ絶対値の演算を行っている。よって、1次微分関数による絶対値だけで演算を行う場合に比べて、はるかに精度を良くすることができ、不变性を有する特徴量を算出できるようになる。

30

## 【0163】

また、「人の着ている服装と背景に依存する特徴量として選択性がありすぎる」といったことに対しては、位置ずれによる不变演算を行うことで、そのようなこと吸収した特徴量を演算できるようになる。

## 【0164】

位置ずれによる不变演算とは、例えば、人の顔の輪郭を検出したとき、顔の形によらずその輪郭の長さはほぼ同じになるといったことを利用した演算である。

30

## 【0165】

換言すれば、輪郭の所定の部分に注目したとき、その部分が位置的にずれても、例えば、ほぼ丸顔の人の輪郭が位置的に移動し、細長い顔の人の輪郭に重なるようにしたときに、位置がずれただけで、その長さなどの値は不变であるとみなせる演算である。

40

## 【0166】

このような演算として、式(7)のように、総和が演算される。総和を演算することにより、例えば、人の顔の輪郭の総和が演算されることになる。

## 【0167】

または、式(8)のように、最大値が演算される。最大値を演算することにより、例えば、人の顔の輪郭のうちの最大値が演算されることになる。

## 【0168】

ここでは、総和と最大値という2つの演算を示した。換言すれば、上記したように、式(7)に基づき、局所的なエネルギーを計算する演算を示し、式(8)に基づき、局所的な最大値を計算する演算を示した。

50

## 【0169】

この他にも、局所的な最大値を有する点の周辺の局所的なエネルギーを計算する演算が行われるようにも良い。これは、式(8)の演算結果を受けて、式(7)の演算を行うようなイメージである。

【0170】

または、局所的なエネルギーの周辺の最大値を計算する演算が行われるよにしても良い。これは、式(7)の演算結果を受けて、式(8)の演算を行うようなイメージである。具体的な式は示さないが、このような演算で特徴量が算出されるようにも良い。

【0171】

このような演算により、各輪郭特徴点から特徴量が算出される。そして、ステップS29において、輪郭特徴量計算部85は、全ての輪郭特徴点について処理が終了したか否かを判定する。

【0172】

例えば、輪郭特徴点抽出部84から供給された全ての輪郭特徴点について、輪郭特徴量が求められた場合、処理が終了したと判定される。

【0173】

ステップS29において、全ての輪郭特徴点について処理がまだ終了していないと判定された場合、処理はステップS21に戻り、次の輪郭特徴点が注目画素として選択される。

【0174】

これに対して、ステップS29において、全ての輪郭特徴点について処理が終了したと判定された場合、処理は図9のステップS9に進む。

【0175】

ステップS9において、識別計算部86は、記憶部48に予め記憶されている識別用特徴量および統合識別器を読み出して、読み出した統合識別器に特徴量を代入して計算する。

【0176】

すなわち、識別計算部86は、対象物特徴量計算部83からの対象物特徴量、および輪郭特徴量計算部85からの輪郭特徴量のうちの識別用特徴量に対応するものを、統合識別器に代入して演算を行う。

【0177】

ここで、識別用特徴量とされる特徴量は、統計学習処理時において、統合識別器を構成する弱識別器の設定に用いられた特徴量である。すなわち誤り率が比較的低い所定の個数の弱識別器に対応する対象物特徴点のペアの対象物特徴量、および輪郭特徴点の輪郭特徴量である。

【0178】

また統合識別器は、対象物識別器と輪郭識別器とを統合して生成されている。対象物識別器は、対象物特徴量を用いて画像に対象物が含まれているかを識別する識別器である。輪郭識別器は、輪郭特徴量を用いて画像に対象物が含まれているかを識別する識別器である。

【0179】

所定の個数の弱識別器は、対応する対象物特徴点のペアの対象物特徴量が代入された場合、結果が正であるとき「+1」を出力し、負であるとき「-1」を出力する。そして対象物識別器は、所定の個数の弱判別器の多数決により認識しようとする対象物体の有無を出力する。

【0180】

輪郭識別器も同様に、輪郭特徴点の輪郭特徴量が弱識別器に代入され、所定の個数の弱判別器の多数決により認識しようとする対象物体の有無を出力する。

【0181】

具体的には、統合識別器は、以下のように生成される。

【0182】

10

20

30

40

50

まず、次式(9)を計算し、対象物識別器R(x)と、輪郭識別器T(x)との識別器和U(x)を求める。すなわち、対象物識別器R(x)と、輪郭識別器T(x)との線形結合により、識別器和U(x)が求められる。

【0183】

$$U(x) = \cdot R(x) + \cdot T(x) \quad \dots (9)$$

【0184】

なお、式(9)において、 $\cdot$ および $\cdot$ は所定の定数、すなわちチューニングパラメータであり、例えば統計学習処理に用いられる学習画像に対する識別率などにより求められる。

【0185】

求められた識別器和U(x)を用いて、次式(10)により示される統合識別器が生成される。

【0186】

$$\text{統合識別器} = \text{sign}(U(x)) \quad \dots (10)$$

【0187】

式(10)の演算が行われると、その演算の結果として、入力画像中に対象物体としての人が存在することを示す「+1」、または入力画像中に対象物体としての人が存在しないことを示す「-1」が得られる。識別計算部86は、統合識別器での演算結果を識別結果出力部87に供給する。

【0188】

ステップS10において、識別結果出力部87は、識別計算部86からの演算結果に基づいて、人の認識結果を出力し、認識処理は終了する。すなわち、対象物体が画像で認識されたか否かの識別結果が出力される。

【0189】

なお、以上においては、対象物体として人を検出すると説明したが、人に限らず、どのようなものであってもよい。

【0190】

例えば、対象物体が入力画像で認識されたか否かの識別結果として、図13に示すように、対象物体としての人、手、及び顔が検出された領域に枠が設定され、表示された画像が表示されるようにしてもよい。

【0191】

図13に示す画像は、対象物体として人151、手152、および顔153が表示されている画像であり、画像には、人151を囲む人認識枠161、手152を囲む手認識枠162、および顔153を囲む顔認識枠163が設定され、表示されている。

【0192】

このような表示を行う場合、認識結果出力部87にも画像が入力され、識別計算部86は、演算結果とともに入力画像における対象物体が検出された領域を示す情報を識別結果出力部87に供給する。

【0193】

認識結果出力部87は、識別計算部86からの演算結果および領域を示す情報をに基づいて、画像から対象物体が検出された場合には、画像とともに、対象物体(例えば人151)が検出された領域を囲む枠(例えば人認識枠161)を設定し、ディスプレイ47Aに表示させる。

【0194】

このようにして、第2認識部63は、画像から対象物特徴点を抽出して、対象物特徴点のペアの対象物特徴量を求るとともに、画像から輪郭特徴点を抽出して輪郭特徴量を求める。

【0195】

そして、第2認識部63は、求めた対象物特徴量および輪郭特徴量と、記憶部48に予め記憶されている統合識別器とを用いて、画像から対象物体として人を検出する。

10

20

30

40

50

## 【0196】

以上のようにして、第2認識部63は対象物体として人を認識する。第1認識部62が手を対象物として認識する処理も、対象物が異なるだけで、第2認識部63の処理と同様であり、その説明は繰り返しになるので省略する。

## 【0197】

## [対象物認識処理1]

## 【0198】

次に、図14乃至図19を参照して、第1認識部62と第2認識部63の2個の認識器を用いて対象物を認識する処理の例を説明する。なお認識器の数は、2個に限らず、3以上の複数個の認識器を用いてもよい。

10

## 【0199】

図14は、対象物認識処理を説明するフローチャートである。図14の例では、第1認識部62が認識対象物として手を認識し、第2認識部63が認識対象物として人を認識するものとする。

## 【0200】

ステップS41において、取得部61は、入力部46を構成するカメラ46Aから画像を取得する。例えば、図15Aに示す人151-1が手152-1を挙げている画像を取得される。

20

## 【0201】

ステップS42において、第1認識部62は、取得した画像に対して第1の認識処理を行う。すなわち、図15Aの画像から手を認識する処理が実行される。この処理の詳細は上述した通りである。

## 【0202】

図15Aの画像は、人151-1がカメラ46Aから離れた場所に立っている。このような場合、手152-1が小さくなるので、ステップS42の第1認識処理では図15Aの画像からは、手152-1を認識することが困難になる。

30

## 【0203】

ステップS43において、判定部64は、第1の認識処理が成功したかを判定する。すなわち、図15Aの画像から手152-1を認識できたかが判定される。

## 【0204】

ステップS43において第1の認識処理が失敗したと判定された場合、すなわち手を認識することができなかった場合、処理はステップS44に進む。

40

## 【0205】

ステップS44において、第2認識部63は、取得した画像に対して第2の認識処理を行う。すなわち、図15Aの画像に対して人を認識する処理が実行される。この処理の詳細は上述した通りである。

## 【0206】

図15Aの画像では、人151-1がカメラから離れた場所に立っている。しかし、人151-1の大きさは、手152-1より大きく、認識に充分な大きさである。ステップS44の第2の認識処理で人151-1が認識されたときの画像を図15Bに示す。

40

## 【0207】

図15Bは、人151-1が認識されたときの画像の例を示す図である。第2認識部63により人151-1が認識され、人認識枠161-1が設定され、表示されている。

## 【0208】

ステップS45において、判定部64は、第2の認識処理が成功したかを判定する。すなわち、図15Aの画像から人151-1を認識できたかが判定される。

## 【0209】

ステップS45において第2の認識処理が失敗したと判定された場合、処理はステップS41に戻る。すなわち、第1の認識処理と第2の認識処理のいずれも失敗した場合には、ステップS41乃至ステップS45の処理が繰り返される。

50

## 【0210】

ステップS45において第2の認識処理が成功したと判定された場合、すなわち図15Bの例のように、人151-1を認識することができた場合、すなわち、手の認識処理には失敗したが、人の認識処理には成功した場合、処理はステップS46に進む。

## 【0211】

ステップS46において、取得部61は、記憶部48から画質パラメータを取得する。画質パラメータは、カメラ制御パラメータと画像処理パラメータに分類できる。

## 【0212】

カメラ制御パラメータは、例えば、ズーム率、フォーカス位置、絞り値、シャッター時間、ゲイン値、色調整値、およびカメラ方向などカメラ本体の設定値を定めるパラメータである。画像処理パラメータは、例えば、明度、コントラスト、色変換、先鋭度値、ぼかし、およびデジタルズーム率など取得した画像を処理するためのパラメータである。

10

## 【0213】

このカメラ制御パラメータおよび画像処理パラメータのうちの少なくとも1種類が画質パラメータとして予め記憶部48に記憶されている。記憶部48に記憶されている画質パラメータは、その設定値が離散的に用意されている。

## 【0214】

例えば、画質パラメータとしてズーム率が設定されており、P個（Pは自然数）の設定値が用意されている場合、ズーム率  $a = \{a_1, a_2, \dots, a_P\}$  として記憶されている。取得部61は、 $\{a_1, a_2, \dots, a_P\}$  の中から新たな設定値を画質パラメータとして取得する。

20

## 【0215】

なお、例えば、ズーム率  $a$  に加え、明度  $b = \{b_1, b_2, \dots, b_Q\}$  ( $Q$ は自然数) を画質パラメータとして設定する場合、ズーム率  $a$  と明度  $b$  との設定値の組み合わせを画質パラメータとして提供することができる。

## 【0216】

また画質パラメータの種類が3種類以上の場合も同様である。そして、画像が認識されるまで全通りの組み合わせを試す方法が有効である。

## 【0217】

ステップS47において、制御部65は、現在の画質パラメータを、ステップS46において取得された新たな画質パラメータに変更するように制御する。ステップS47の処理の後、処理はステップS41に戻る。

30

## 【0218】

なお、画質パラメータに変えて、ユーザに対して「カメラに近づいて下さい」というような指示を取得し、ディスプレイ47Aに表示するようにしてもよい。

## 【0219】

ステップS43において第1の認識処理が成功したと判定されるまで、ステップS41乃至S47の処理が繰り返される。

## 【0220】

ステップS47の処理の後のステップS41の処理において、取得部61は、ステップS47において画質パラメータが変更された画像を取得する。図15Cを参照して、画質パラメータを変更した手を認識するのに適切な画像について説明する。

40

## 【0221】

図15Cは、画質パラメータとしてズーム率が変更された画像を示す図である。図15Cの図では、ズーム率がより望遠となる画質パラメータに変更された場合の例を示している。すなわち、人151-2が画像上により大きく表示されている。

## 【0222】

ステップS42において、第1認識部62は、第1の認識処理を実行する。画質パラメータが変更された図15Cの画像は、手152-2の大きさが、図15Aに示される場合より大きく、認識に充分な大きさである。従って、ステップS42の処理で、第1認識部

50

6 2 により手 1 5 2 - 2 が認識され、手認識枠 1 6 2 - 2 が設定され、表示される。

【0 2 2 3】

ステップ S 4 3 において第 1 の認識処理が成功したと判定された場合、すなわち手を認識することができた場合、処理はステップ S 4 8 に進む。

【0 2 2 4】

ステップ S 4 8 において、判定部 6 4 は、ステップ S 4 4 の第 2 の認識処理が成功したかを判定する。すなわち、人を認識する処理が成功したかが判定される。

【0 2 2 5】

ステップ S 4 8 において第 2 の認識処理が成功したと判定された場合、処理はステップ S 4 9 に進み、ステップ S 4 9 において、判定部 6 4 は、条件を満足するかを判定する。

【0 2 2 6】

条件は、相対的な位置関係と相対的な大きさの関係との 2 つの条件がある。まず図 1 6 と図 1 7 を参照して、相対的な位置関係の条件について説明する。

【0 2 2 7】

図 1 6 A は、人認識枠 1 6 1 と手認識枠 1 6 2 との相対的な位置関係の例を示す図である。検出領域 1 8 1 は、人認識枠 1 6 1 の所定の位置に予め設定されている。そして判定部 6 4 により手認識枠 1 6 2 が、検出領域 1 8 1 内にあるかが条件として判定される。

【0 2 2 8】

具体的には、例えば、人認識枠 1 6 1 の左上の頂点の座標 (x<sub>b</sub>, y<sub>b</sub>) に対して、検出領域 1 8 1 の左上の頂点の座標 (x<sub>1</sub>, y<sub>1</sub>) と、検出領域 1 8 1 の右下の頂点の座標 (x<sub>2</sub>, y<sub>2</sub>) は所定の関係になるように予め設定される。

【0 2 2 9】

この場合、判定部 6 4 は、手認識枠 1 6 2 の左上の頂点の座標 (x<sub>h</sub>, y<sub>h</sub>) が次式を満足するかを判定する。

【0 2 3 0】

【数 9】

$$\left\{ \begin{array}{l} X_1 \leq X_h \leq X_2 \\ y_1 \leq y_h \leq y_2 \end{array} \right. \cdots (11)$$

【0 2 3 1】

図 1 6 B は、人認識枠 1 6 1 と手認識枠 1 6 2 との相対的な位置関係の他の例を示す図である。検出領域 1 8 2 は、人認識枠 1 6 1 に対する手認識枠 1 6 2 の相対位置の分布に基づいて設定される。

【0 2 3 2】

具体的には、手の位置の座標 (x<sub>h</sub>, y<sub>h</sub>) の平均  $\mu_{x, y}$  と共に分散  $\sigma_{x, y}$  が与えられることで定義される 2 次元正規分布を考えたときの、共分散  $\sigma_{x, y}$  の値が閾値 w 0 以上となる所定の範囲が規定される。

【0 2 3 3】

つまり、手認識枠 1 6 2 の左上の頂点の座標を (x<sub>h</sub>, y<sub>h</sub>) を取得したとき、判定部 6 4 は、その座標 (x<sub>h</sub>, y<sub>h</sub>) が次式を満足するかを判定する。すなわち、座標が x y 平面内の平均  $\mu_{x, y}$  から所定の距離の範囲内に位置するかが判定される。

【0 2 3 4】

【数 1 0】

$$2\text{次元正規分布 } w = N(X_h | \mu_{x,y}, \sum_{xy})$$

$$w > w_0 \cdots (12)$$

【0 2 3 5】

なお、図 1 6 の例では、手認識枠 1 6 2 の左上の頂点の座標 (x<sub>h</sub>, y<sub>h</sub>) を基準の座

10

20

30

40

50

標としたが、基準とする座標は、手認識枠 162 の他の頂点、辺上、または重心などでもよい。

【0236】

また、図 16B の例では 2 次元正規分布を使用したが、ロジスティック分布、双曲線正割分布、レイリー分布、レビイ分布、またはコーシー分布など他の確率分布を同様に規定して使用してもよい。

【0237】

図 17A は、相対的な位置関係の条件を満足する例を示す図である。人認識枠 161-3 の左上の頂点の座標 ( $x_b, y_b$ ) に対する手認識枠 162-3 の左上の頂点の座標 ( $x_h, y_h$ ) は、検出領域 181-3 内に存在する。従って、判定部 64 は、相対的な位置関係の条件を満足すると判定する。

10

【0238】

図 17B は、相対的な位置関係の条件を満足しない例を示す図である。図 17B に示すように、手認識枠 162-4 の左上の頂点の座標 ( $x_h, y_h$ ) が検出領域 181-4 の範囲外である人認識枠 161-4 の左下の位置に存在する。

20

【0239】

このような場合、例えば手がユーザのひざ付近に位置していることになるので、その認識結果は誤認識の可能性が高い。従って、判定部 64 は、相対的な位置関係の条件を満足しないと判定する。

【0240】

次に図 18 と図 19 を参照して、相対的な大きさの関係の条件について説明する。図 18A は、人認識枠 161 の大きさを示す図である。図 18A の例では、人認識枠 161 の横の長さを人認識枠 161 の大きさ  $H$  としている。

30

【0241】

図 18B は、手認識枠 162 の大きさを示す図である。図 18B の例では、手認識枠 162 の横の長さを手認識枠 162 の大きさ  $h$  としている。

【0242】

そして、この人認識枠 161 の大きさ  $H$  と手認識枠 162 の大きさ  $h$  との比が、予め設定された所定の範囲内にあるかが判定される。

【0243】

具体的には、判定部 64 は、次式が満足されるかを判定する。

30

【0244】

【数 11】

$$a_1 < \frac{h}{H} < a_2 \quad (0 < a_1, a_2 < 1, a_1, a_2 \text{ は有理数}) \quad \cdots (13)$$

【0245】

また、他の例として、人認識枠 161 の大きさ  $H$  に対する手認識枠 162 の大きさ  $h$  の比  $s$  に関する平均値  $\mu_s$  と分散  $\sigma^2$  が与えられることで定義される正規分布を考えたときの、分散  $\sigma^2$  の値が閾値  $s_0$  以上となる所定の範囲が規定される。

40

【0246】

つまり、手認識枠 162 の大きさ  $h$  と、人認識枠 161 の大きさ  $H$  が得られたとき、判定部 64 は、次式が満足されるかを判定する。すなわち、検出された手認識枠 162 の大きさ  $h$  と人認識枠 161 の大きさ  $H$  に対する比  $s$  の分散  $\sigma^2$  が、閾値  $s_0$  以上であるか、つまり、比  $s$  が平均  $\mu_s$  から所定の範囲内にあるかが判定される。

【0247】

【数12】

$$\text{正規分布 } S = N\left(\frac{h}{H} \mid \mu_s, \sigma^2\right) \quad \cdots(14)$$

$$S > S_0$$

【0248】

図19Aは、相対的な大きさの関係の条件を満足する例を示す図である。

【0249】

人認識枠161-11の大きさHに対する手認識枠162-11の大きさhの比が、予め設定された範囲内である場合、すなわち手認識枠162-11の大きさhが、人認識枠161-11の大きさHに対して適切な大きさである場合、判定部64は、相対的な大きさの関係の条件を満足すると判定する。 10

【0250】

図19Bは、相対的な大きさの関係の条件を満足しない例を示す図である。図19Bに示すように、手認識枠162-12の大きさh'は、人認識枠161-12の大きさH'に較べて小さ過ぎる。

【0251】

このような場合、手または人の認識結果は誤認識の可能性が高いので、判定部64は、相対的な大きさの関係の条件を満足しないと判定する。これにより、より正確に対象物を認識することができる。 20

【0252】

なお、条件として相対的な位置関係と相対的な大きさの関係との2つの条件について説明したが、この2つの条件を組み合わせて判断してもよい。

【0253】

すなわち、式(12)と式(13)とを組み合わせて、判定部64は、次式が満足されるかを判定するようにしてもよい。

【0254】

【数13】

$$Z = WS$$

$$Z > Z_0 \quad \cdots(15)$$

30

【0255】

図14に戻り、ステップS49において条件が満足されていないと判定された場合、処理はステップS46に進む。すなわち、手と人の両方が認識されたとしても条件が満足されていない場合、画質パラメータが取得され、新たな画像から認識処理が行われる。

【0256】

ステップS49において条件が満足されていると判定された場合、すなわち、手と人の両方が認識され、条件も満足されている場合、処理はステップS50に進む。 40

【0257】

また、ステップS48において第2の認識処理が成功していないと判定された場合、すなわち、ステップS41において最初に取得した、画質パラメータを変更していない画像から手が認識されたが、人は認識することができなかった場合にも、処理はステップS50に進む。

【0258】

ステップS50において、出力部66は、認識成功を出力する。例えば、出力部66は、アプリケーション実行部67に認識成功を出力する。

【0259】

アプリケーション実行部67は、出力部66から認識成功が通知されたとき、所定のア

40

50

プリケーションを実行する。例えば、アプリケーション実行部 6 7 により、ジェスチャ認識に基づくコマンドが実行される。

【0260】

ステップ S 5 0 の処理の後、図 1 4 の対象物認識処理は終了する。

【0261】

このように、手と人とをそれぞれ認識する 2 つの認識器を用いることで、手を認識する認識器単体では認識が困難な画像でも、より確実に手を認識することができるようになる。

【0262】

なお、図 1 4 の例では、第 1 認識部 6 2 が認識対象物として手を認識し、第 2 認識部 6 3 が認識対象物として人を認識するものとしたが、第 1 認識部 6 2 および第 2 認識部 6 3 の認識対象物は、これらに限られない。

10

【0263】

[ 対象物認識処理 2 ]

【0264】

本実施の形態の対象物認識処理の他の例として、第 1 認識部 6 2 が認識対象物として顔を認識し、第 2 認識部 6 3 が認識対象物として人を認識する場合の例を図 1 4 と図 2 0 を参照して説明する。なお、図 1 4 と図 1 5 を参照して説明した例と同様の処理は繰り返しになるので簡単に説明する。

20

【0265】

図 1 4 のステップ S 4 1 において、取得部 6 1 は、図 2 0 A の画像を取得する。図 2 0 A の画像は、ピントがぼけている画像の例を示す図である。

【0266】

ステップ S 4 2 において、第 1 認識部 6 2 は、取得した図 2 0 A の画像に対して第 1 の認識処理、すなわち顔を認識する処理を行う。しかし、図 2 0 A の画像は、ピントがぼけているため顔 1 5 3 - 1 5 を認識することが困難である。

【0267】

ステップ S 4 3 において、判定部 6 4 は、第 1 の認識処理、すなわち顔 1 5 3 - 1 5 の認識に成功したか判定する。図 2 0 A の画像では、顔 1 5 3 - 1 5 を認識することができないので、ステップ S 4 3 において第 1 の認識処理が失敗したと判定され、処理はステップ S 4 4 に進む。

30

【0268】

ステップ S 4 4 において、第 2 認識部 6 3 は、取得した図 2 0 A の画像に対して第 2 の認識処理、すなわち人を認識する処理を行う。図 2 0 A の画像は、ピントがぼけているが、人 1 5 1 - 1 5 の画像は認識に充分な画像となっている。

【0269】

ステップ S 4 4 の第 2 の認識処理が実行され、人 1 5 1 - 1 5 が認識されると、図 2 0 B に示すように、人認識枠 1 6 1 - 1 5 が設定され、表示される。

【0270】

ステップ S 4 5 において、判定部 6 4 は、第 2 の認識処理が成功したかを判定する。図 2 0 A の画像が認識される場合、ステップ S 4 5 において第 2 の認識処理が成功したと判定され、処理はステップ S 4 6 に進む。

40

【0271】

ステップ S 4 6 において、取得部 6 1 は、新たな画質パラメータを取得する。図 2 0 の例の場合、画質パラメータとしてフォーカス位置と先鋭度値とが記憶されており、フォーカス位置と先鋭度値との設定値の組み合わせから新たな設定値を取得する。

【0272】

ステップ S 4 7 において、制御部 6 5 は、現在の画質パラメータを、ステップ S 4 6 において取得された新たな画質パラメータに変更するように制御する。ステップ S 4 7 の処理の後、処理はステップ S 4 1 に戻る。

50

## 【0273】

なお、ステップS43において第1の認識処理が成功したと判定されるまで、ステップS41乃至S47の処理が繰り返される。

## 【0274】

ステップS41において、顔を認識するのに適切な画像を取得したとすると、ステップS42において、第1の認識処理が実行されることにより、図20Cの画像が出力される。ステップS42の処理で、顔153-16が認識されると、顔認識枠163-16が設定、表示される。

## 【0275】

そして、ステップS43において、第1の認識処理が成功したと判定され、処理はステップS48に進む。ステップS48以降の処理は、図14乃至図19を参照して説明した例と同様であるので省略する。

10

## 【0276】

なお、図20の例の場合、アプリケーション実行部67は、ステップS50の処理により認識成功が出力されると、顔をよりはっきり且つきれいに表示できるように、ピントや色を調整するアプリケーションを実行する。勿論、画質パラメータとして、ピントを変更するようにしてもよい。

## 【0277】

## [姿勢認識部の構成]

## 【0278】

20

条件が満足されているかの判定に人の姿勢を利用することができる。図21は、このような場合に、CPU41にさらに含まれる姿勢認識部201の詳細な構成例を示すブロック図である。

## 【0279】

姿勢認識部201は、遅延画像取得部221、演算部222、信号強度検出部223、および検出結果出力部224から構成されている。

## 【0280】

遅延画像取得部221は、RAM43に所定時間前に記憶された画像、または図示せぬ遅延回路を介して供給される所定時間前の画像を取得し、演算部222に供給する。

30

## 【0281】

演算部222は、原画像と遅延画像の差分を演算し、演算結果を信号強度検出部223に供給する。信号強度検出部223は、信号強度が所定の値より高い領域を検出し、検出結果を検出結果出力部224に供給する。

## 【0282】

検出結果出力部224は、検出結果を画像上に重畳して出力する。

## 【0283】

## [対象物認識処理3]

## 【0284】

図22乃至図26を参照して、本実施の形態の対象物認識処理のさらに他の例を示す。図22においてステップS61乃至S68、S70、S71の処理は、図14のステップS41乃至S50の処理に対応する処理であり、これらの処理は繰り返しになるので簡単に説明する。

40

## 【0285】

図22の例では、第1認識部62が認識対象物として手を認識し、第2認識部63が認識対象物として顔を認識する場合の例を示している。

## 【0286】

ステップS61において、取得部61は、図23Aの画像を取得する。図23Aの画像は、カメラ46Aの近くに入151-21がいるか、またはカメラ46Aのズーム率のパラメータが広角に設定されているため、手152-21がカメラの画角におさまりきらない画像の例である。

50

## 【0287】

ステップS62において、第1認識部62は、取得した図23Aの画像に対して第1の認識処理、すなわち手152-21を認識する処理を行う。しかし、図23Aの画像は、手152-21が一部しか表示されておらず、手152-21は認識することができない。

## 【0288】

ステップS63において、判定部64は、第1の認識処理、すなわち手152-21の認識に成功したか判定する。図23Aの画像では、手152-21を認識することができないので、ステップS63において第1の認識処理が失敗したと判定され、処理はステップS64に進む。

10

## 【0289】

ステップS64において、第2認識部63は、取得した図23Aの画像に対して第2の認識処理、すなわち顔153-21を認識する処理を行う。図23Aの画像は、顔153-21がはっきり表示されている。

## 【0290】

ステップS64の第2の認識処理が実行され、顔153-21が認識されると、図23Bに示すように、顔認識枠163-21が設定、表示される。

## 【0291】

ステップS65において、判定部64は、第2の認識処理が成功したかを判定する。この場合、ステップS65において第2の認識処理が成功したと判定され、処理はステップS66に進む。

20

## 【0292】

ステップS66において、取得部61は、新たな画質パラメータを取得する。図23の例の場合、画質パラメータとしてズーム率が記憶されており、記憶されているズーム率の設定値から新たな設定値が取得される。

## 【0293】

ステップS67において、制御部65は、現在の画質パラメータを、ステップS66において取得された新たな画質パラメータに変更するように制御する。ステップS67の処理の後、処理はステップS61に戻る。

30

## 【0294】

以上のようにして、ステップS63において第1の認識処理が成功したと判定されるまで、ステップS61乃至S67の処理が繰り返される。

## 【0295】

ステップS61において、手を認識するのに適切な画像が取得されたとすると、ステップS62において、第1の認識処理が実行されることにより、図23Cの画像に示すように、手152-22の周りに、手認識枠162-22が設定、表示される。

## 【0296】

そして、ステップS63において、第1の認識処理が成功したと判定され、処理はステップS68に進む。

40

## 【0297】

ステップS68において、判定部64は、第2の認識処理が成功したかを判定する。図23の例の場合、ステップS68において第2の認識処理が成功したと判定され、すなわち、手と顔が認識されたと判定され、処理はステップS69に進む。

## 【0298】

ステップS69において、姿勢認識部201は、姿勢認識処理を実行する。図24を参照して姿勢認識処理について説明する。

## 【0299】

## [姿勢認識処理]

## 【0300】

図24は、姿勢認識処理の例を説明するフローチャートである。

50

## 【0301】

ステップS81において、遅延画像取得部221は、遅延画像を取得する。原画像の所定時間前の画像である遅延画像は、隨時RAM43に記憶されている。

## 【0302】

ステップS82において、演算部222は、原画像と遅延画像の差分を演算する。ステップS83において、信号強度検出部223は、差分が演算された画像から信号強度が高い領域を検出する。つまり、差分の値の絶対値が大きい領域が検出される。

## 【0303】

例えば、ユーザが手を動かしたとすると、ユーザの動作前の手および腕の位置と、動作後の手および腕の位置の領域が検出される。

10

## 【0304】

ステップS84において、検出結果出力部224は、信号強度検出部223により検出された領域を原画像上に重畳して出力する。検出された領域を原画像上に重畳して出力する例を図23Cを参照して説明する。

## 【0305】

図23Cの画像には、姿勢164-1と姿勢164-2が設定され、出力されている。姿勢164-1は、ユーザの動作前、すなわち遅延画像におけるユーザの手および腕の位置を示している。姿勢164-2は、ユーザの動作後、すなわち原画像におけるユーザの手および腕の位置を示している。

## 【0306】

図24のステップS84の処理の後、処理は図22のステップS70に進む。ステップS70において、判定部64は、条件を満足するかを判定する。図25と図26を参照して、手認識枠162と姿勢164の相対的な距離の関係と相対的な角度の関係の条件について説明する。

20

## 【0307】

図25は、手と腕の相対的な距離の関係と相対的な角度の関係の条件の例を説明する図である。

## 【0308】

まず、相対的な距離の関係の条件について説明する。相対的な距離は、例えば、手認識枠162の4つの辺のうち姿勢164の先端に最も近い辺の中点の座標と、手認識枠162に最も近い姿勢164の先端の座標との間の距離dを演算して算出される。距離dは、式(16)を用いて算出される。

30

## 【0309】

## 【数14】

$$d = \sqrt{dx^2 + dy^2} \quad \cdots (16)$$

## 【0310】

$d_x$ は、手認識枠162の所定の辺の中点Pmのx座標と、姿勢164の先端Pf(この場合、腕の先端)のx座標の差により算出される。同様に $d_y$ は、手認識枠162の所定の辺の中点Pmのy座標と、姿勢164の先端Pfのy座標の差により算出される。

40

## 【0311】

そして判定部64は、次式が満足されるかを判定する。すなわち距離dが、予め設定された値d1とd2の範囲内にあるかが判定される。

## 【0312】

## 【数15】

$$d1 \leq d < d2 \quad (d_1, d_2 \text{ は } 0 \text{ 以上の実数}) \quad \cdots (17)$$

## 【0313】

また、他の例として、距離dに関する平均値 $\mu_d$ と分散 $\sigma_d^2$ が与えられることで定義

50

される正規分布を考えたときの、分散  $\sigma_d^2$  の値が閾値  $e_0$  以上となる所定の範囲が規定される。

【0314】

そして、判定部 64 は、次式が満足されるかを判定する。すなわち、検出された距離  $d$  が平均  $\mu_d$  から所定の範囲内にあるかが判定される。

【0315】

【数16】

正規分布  $e = N(d | \mu_d, \sigma_d^2)$

$$e > e_0 \quad \cdots (18)$$

10

【0316】

次に、相対的な角度の関係の条件について説明する。相対的な角度は、手認識枠 162 の所定の辺と平行な平行線 191 と姿勢 164 との間にできる角度  $d$  により表わされる。

【0317】

平行線 191 に対応する辺としては、手認識枠 162 の辺のうち、角度  $d$  がより小さくなる辺が選択される。図 25 の例においては、鉛直方向に近い辺が選択されている。

【0318】

角度  $d$  に関する平均値  $\mu_d$  と分散  $\sigma_d^2$  が与えられることで定義される正規分布を考えたときの、分散  $\sigma_d^2$  の値が閾値  $f_0$  以上となる所定の範囲が規定される。

20

【0319】

そして、判定部 64 は、次式が満足されるかを判定する。すなわち、検出された角度  $d$  が平均  $\mu_d$  から所定の範囲内にあるかが判定される。

【0320】

【数17】

正規分布  $f = N(d\theta | \mu\theta, \sigma_\theta^2)$

$$f > f_0 \quad \cdots (19)$$

30

【0321】

なお、条件として相対的な位置距離と相対的な角度の関係との 2 つの条件について説明したが、この 2 つの条件を組み合わせて判断してもよい。

【0322】

例えば、式(18)と式(19)とを組み合わせて、判定部 64 は、次式を満足するかが判定される。

【0323】

【数18】

$q = ef$

$$q > q_0 \quad \cdots (20)$$

40

【0324】

図 26 は、相対的な距離の判定結果の例を示す図である。

【0325】

図 26 A は、相対的な距離の関係の条件を満足する例を示す図である。手認識枠 162 - 31 と姿勢 164 - 31 は接触しており、距離  $d = 0$  となる。従って、手認識枠 162 - 31 と姿勢 164 - 31 との距離は、予め設定された範囲内にあるので、判定部 64 は、相対的な大きさの関係の条件を満足すると判定する。

【0326】

図 26 B は、相対的な距離の関係の条件を満足しない例を示す図である。図 26 B に示

50

すように、手認識枠 162-32 の 4 つの辺のうち姿勢 164-32 の先端に最も近い辺の中点の座標と、手認識枠 162-32 に最も近い姿勢 164-32 の先端の座標との間の距離  $d$  は、 $d'$  となっている。

【0327】

手認識枠 162-32 と姿勢 164-32 との距離  $d$  が、予め設定された範囲外であるとすると、手の認識結果は誤認識の可能性が高いので、判定部 64 は、相対的な距離の関係の条件を満足しないと判定する。

【0328】

これにより、図 26B に示すように、顔認識枠 163-32 に対する手認識枠 162-32 の相対位置が、検出領域 181-11 内であり、相対位置の関係の条件を満足しても、手認識枠 162-32 と姿勢 164-32 との相対位置の関係の条件を満足していないので、判定部 64 は、条件を満足していないと判定する。

【0329】

このように、図 14 の対象物認識処理に対して、さらに姿勢認識（例えば、腕）の認識結果を使用して対象物（例えば、手）を認識できたかを判定するので、より確実に対象物を認識することができるようになる。

【0330】

図 22 に戻り、ステップ S70 において条件を満足していないと判定された場合、処理はステップ S66 に進む。

【0331】

ステップ S70 において条件を満足すると判定された場合、処理はステップ S71 に進む。ステップ S71 において、出力部 66 は、認識成功を出力する。ステップ S71 の処理の後、図 22 の対象物認識処理は終了する。

【0332】

【テレビジョン受像機制御システムの構成】

【0333】

次に図 27 を参照して、本実施の形態におけるテレビジョン受像機制御システム 301 について説明する。図 27 は、本発明を適用したテレビジョン受像機制御システム 301 の構成例を示すブロック図である。

【0334】

情報処理装置としてのテレビジョン受像機制御システム 301 は、例えば、図 14 と図 15 を参照して上述した対象物認識処理 1 と同様の処理を実行する。テレビジョン受像機制御システム 301 は、カメラ 321、制御装置 322、画像表示装置 323、および RAM 324 から構成されている。

【0335】

カメラ 321 は、設定された画質パラメータに基づく画像を取得する。またカメラ 321 は、画像表示装置 323 を見るユーザを撮影できる位置に配置されている。

【0336】

画像表示装置 323 は、地上デジタル放送、BS 放送、および CS 放送などの画像の他、必要に応じてカメラ 321 により取得された画像を表示する。

【0337】

RAM 324 は、制御装置 322 において各種の処理を実行するのに必要なデータ、プログラム等を記憶する。

【0338】

制御装置 322 において、カメラ制御エンジン 341、人認識エンジン 342、手認識エンジン 343、認識結果判定エンジン 344、ジェスチャ認識エンジン 345、テレビジョン受像機メインエンジン（以下、TV メインエンジンと称する）346、および制御用 CPU 348 は、画像信号用バス 347 およびバス 349 により相互に接続されている。

【0339】

カメラ制御エンジン 341 は、カメラ 321 から画像を取得する。また、カメラ制御工

10

20

30

40

50

ンジン 3 4 1 は、画像が上述した画質パラメータの設定値に基づく画像になるようにカメラを制御する。

【 0 3 4 0 】

人認識エンジン 3 4 2 は、画像信号用バス 3 4 7 を介してカメラ制御エンジン 3 4 1 から取得した画像から認識対象物として人を認識する。具体的には、人認識エンジン 3 4 2 の処理は、対象物認識処理 1 における第 2 認識部 6 3 の処理と同様の処理を行う。

【 0 3 4 1 】

手認識エンジン 3 4 3 は、画像信号用バス 3 4 7 を介してカメラ制御エンジン 3 4 1 から取得した画像から認識対象物として手を認識する。具体的には、手認識エンジン 3 4 3 の処理は、対象物認識処理 1 における第 1 認識部 6 2 の処理と同様の処理を行う。

10

【 0 3 4 2 】

認識結果判定エンジン 3 4 4 は、画像信号用バス 3 4 7 またはバス 3 4 9 を介して取得した人認識エンジン 3 4 2 および手認識エンジン 3 4 3 の認識結果に基づいて、対象物を認識できたかを判定し、判定結果を出力する。具体的には、認識結果判定エンジン 3 4 4 は、図 1 4 のステップ S 4 3 , S 4 5 , S 4 8 , S 4 9 , S 5 0 の処理と同様の処理を実行する。

【 0 3 4 3 】

ジェスチャ認識エンジン 3 4 5 は、画像信号用バス 3 4 7 を介して取得した人認識エンジン 3 4 2 または手認識エンジン 3 4 3 の認識結果から、人または手のジェスチャの種類を認識する。

20

【 0 3 4 4 】

TV メインエンジン 3 4 6 は、バス 3 4 9 を介して取得したジェスチャ認識エンジン 3 4 5 のジェスチャの認識結果、または制御用 CPU 3 4 8 からの指示に基づいて、画像表示装置 3 2 3 への画像の出力を制御する。

【 0 3 4 5 】

制御用 CPU 3 4 8 は、カメラ制御エンジン 3 4 1 、人認識エンジン 3 4 2 、手認識エンジン 3 4 3 、認識結果判定エンジン 3 4 4 、ジェスチャ認識エンジン 3 4 5 、および TV メインエンジン 3 4 6 の動作を、バス 3 4 9 を介して制御する。

【 0 3 4 6 】

これにより、テレビジョン受像機制御システム 3 0 1 は、ユーザの手の動きのジェスチャの画像をカメラ 3 2 1 により取得し、そのジェスチャの種類を認識するので、ユーザがリモートコントローラなどを使用しないで、チャネルの切り替えなどの操作をすることができる。その動作は上述した場合と同様なので、その説明は省略する。

30

【 0 3 4 7 】

[ デジタルカメラ制御システムの構成 ]

【 0 3 4 8 】

次に図 2 8 を参照して、本実施の形態におけるデジタルカメラ制御システム 4 0 1 について説明する。図 2 8 は、本発明を適用したデジタルカメラ制御システム 4 0 1 の構成例を示すプロック図である。

40

【 0 3 4 9 】

情報処理装置としてのデジタルカメラ制御システム 4 0 1 は、例えば、図 1 4 と図 2 0 を参照して上述した対象物認識処理 2 と同様の処理を実行する。デジタルカメラ制御システム 4 0 1 は、カメラ 4 2 1 、制御装置 4 2 2 、画像表示装置 4 2 3 、RAM 4 2 4 、および記録メディア 4 2 5 から構成されている。

【 0 3 5 0 】

カメラ 4 2 1 は、画像を取得する。画像表示装置 4 2 3 は、カメラ 4 2 1 により取得された画像、または画像処理を施した画像を表示する。

【 0 3 5 1 】

RAM 4 2 4 は、制御装置 4 2 2 において各種の処理を実行するのに必要なデータ、プログラム等を記録する。記録メディア 4 2 5 は、制御装置 4 2 2 により制御され、カメラ

50

421により撮影された画像、その他の情報を記憶する。

【0352】

制御装置422において、カメラ信号処理エンジン441、人認識エンジン442、顔認識エンジン443、認識結果判定エンジン444、画像先鋒度調整エンジン445、JPEG (Joint Photographic Experts Group) エンジン446、画像表示用エンジン447、および制御用CPU449は、画像信号用バス448およびバス452により相互に接続されている。

【0353】

また、制御用CPU449と外部メディアコントローラ451は、データバス450を介して相互に接続されている。データバス450は、画像信号用バス448にも接続されている。

【0354】

カメラ信号処理エンジン441は、カメラ421により取得されたアナログデータの画像をデジタルデータの画像に変換して取得する。

【0355】

人認識エンジン442は、画像信号用バス448を介してカメラ信号処理エンジン441から取得した画像から認識対象物として人を認識する。具体的には、人認識エンジン442の処理は、上述した対象物認識処理2における第2認識部63の処理と同様である。

【0356】

顔認識エンジン443は、画像信号用バス448を介してカメラ信号処理エンジン441から取得した画像から認識対象物として顔を認識する。具体的には、顔認識エンジン443の処理は、上述した対象物認識処理2における第1認識部62の処理と同様である。

【0357】

認識結果判定エンジン444は、画像信号用バス448またはバス452を介して取得した人認識エンジン442および顔認識エンジン443の認識結果に基づいて、対象物を認識できたかを判定し、判定結果を出力する。具体的には、図14のステップS43, S45, S48, S49, S50の処理と同様の処理を実行する。

【0358】

画像先鋒度調整エンジン445は、画像信号用バス448を介して取得したカメラ信号処理エンジン441の画像の先鋒度を調整する画像処理を行う。

【0359】

JPEGエンジン446は、画像信号用バス448を介して取得したカメラ信号処理エンジン441または画像先鋒度調整エンジン445の画像に対してJPEGの規格に基づく圧縮、伸張処理を行う。

【0360】

画像表示用エンジン447は、画像信号用バス448を介して取得した各種画像を画像表示装置423に出力し、表示させる。

【0361】

制御用CPU449は、カメラ信号処理エンジン441、人認識エンジン442、顔認識エンジン443、認識結果判定エンジン444、画像先鋒度調整エンジン445、JPEGエンジン446、および画像表示用エンジン447の動作を、バス452を介して制御する。

【0362】

外部メディアコントローラ451は、記録メディア425を制御して、カメラ421により撮影された画像データを記憶させたり、記憶された画像データを読み出し、画像表示装置423に表示させる。

【0363】

これにより、デジタルカメラ制御システム401は、カメラ421により取得された画像からユーザの顔などを認識し、認識された顔の領域の先鋒度を向上させ、記録メディア425に記録させるので、ユーザがより好適な画像を撮像し、記録再生することができる

10

20

30

40

50

。この場合の動作も上述した場合と同様なので、その説明は省略する。

【0364】

以上においては、認識対象を人およびその部位としたが、その他の物体を認識対象物とすることもできる。例えば、自動車を認識する情報処理装置において、「ボディ認識」、「タイヤ認識」、および「ナンバープレート認識」などを組み合わせて同様の認識処理を実行することが可能である。

【0365】

[その他]

【0366】

本明細書において、プログラムのステップは、その順序に沿って時系列的に行われる処理はもちろん、必ずしも時系列的に処理されなくとも、並列的あるいは個別に実行される処理をも含むものである。 10

【0367】

本発明の実施の形態は、上述した実施形態に限定されるものではなく、本発明の要旨を逸脱しない範囲において種々の変更が可能である。

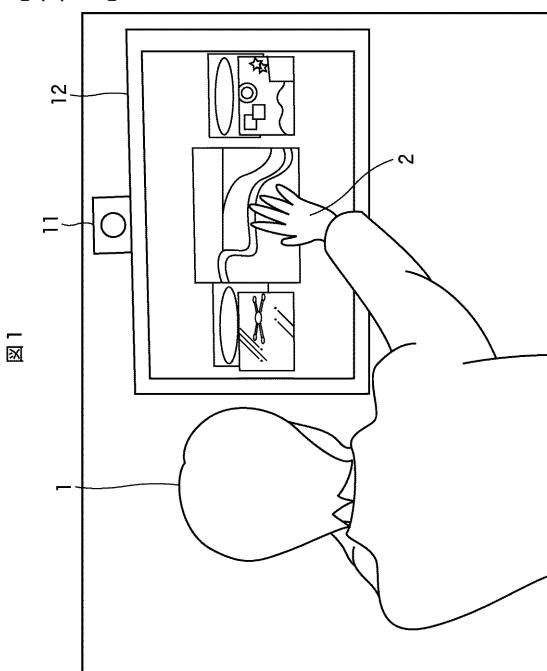
【符号の説明】

【0368】

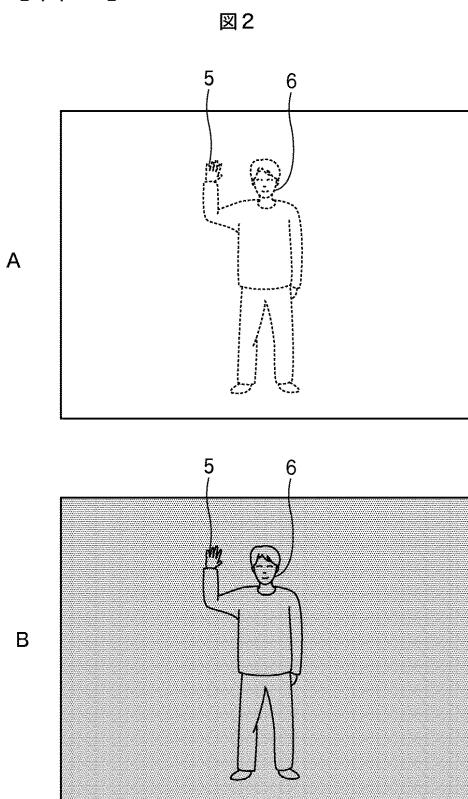
21 コンピュータ, 61 取得部, 62 第1認識部, 63 第2認識部,

64 判定部, 65 制御部

【図1】

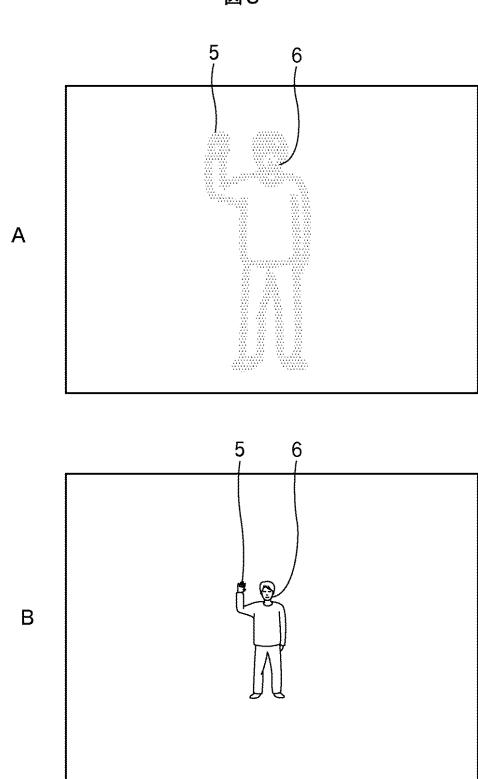


【図2】

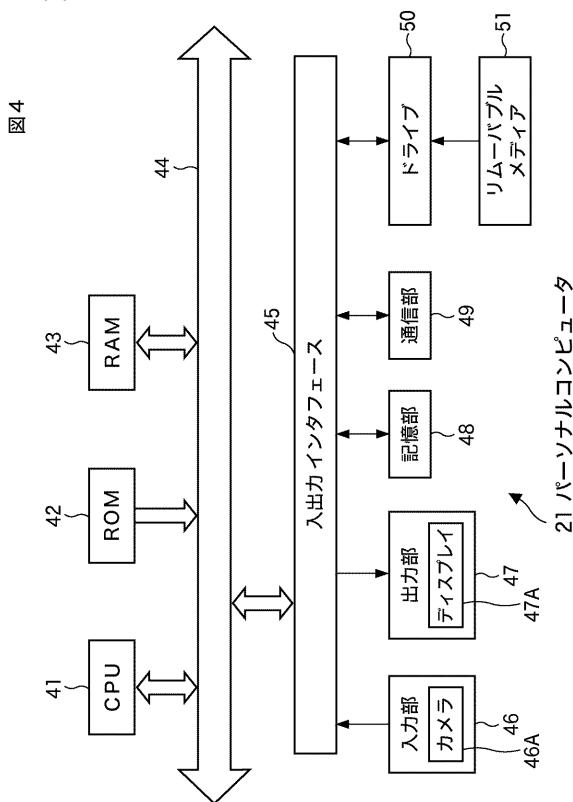


【図3】

図3

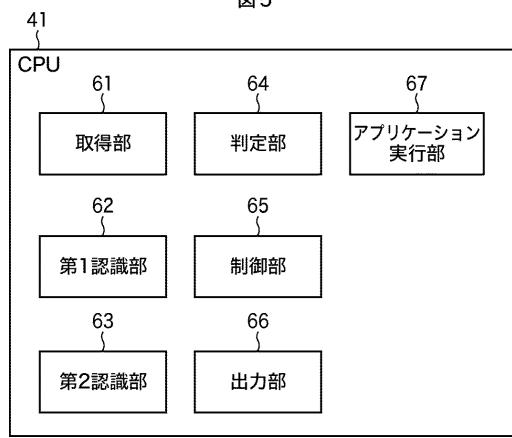


【図4】



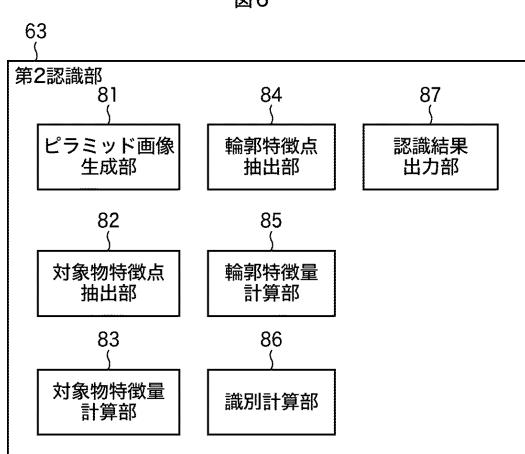
【図5】

図5



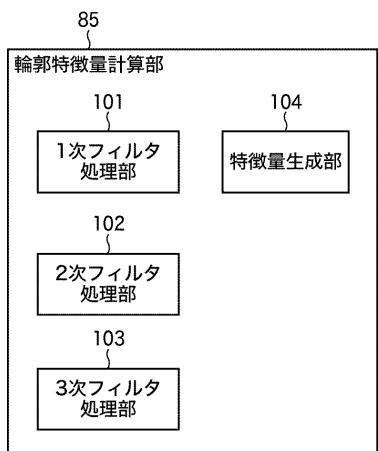
【図6】

図6

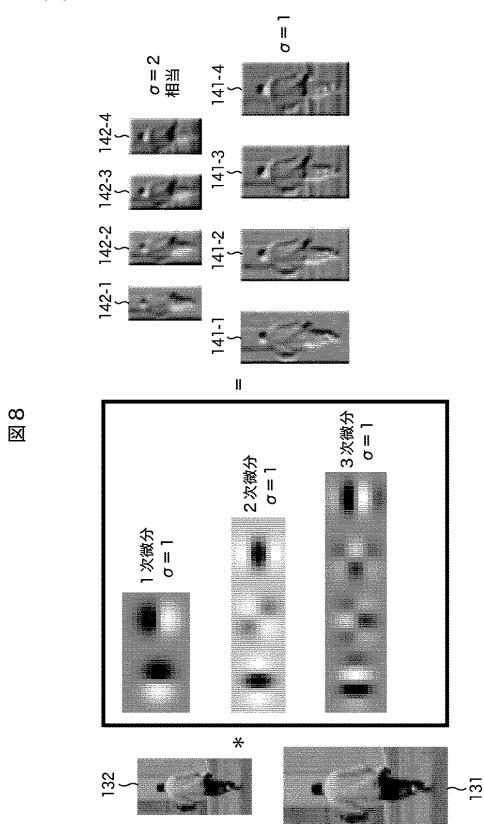


【図7】

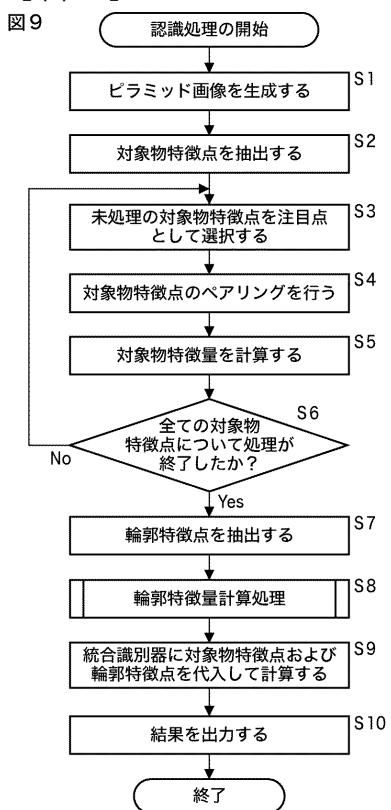
図7



【図8】

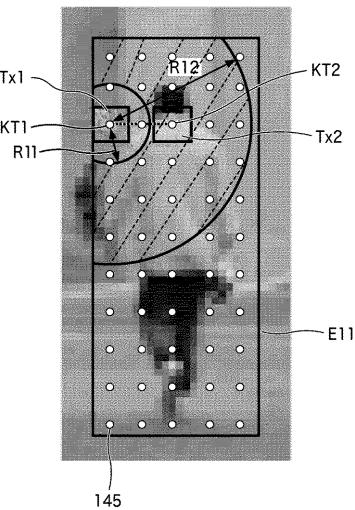


【図9】



【図10】

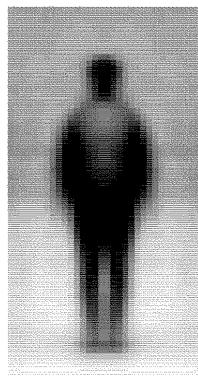
図10



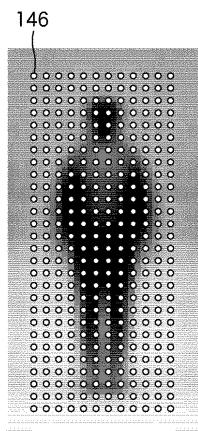
【図11】

図11

A

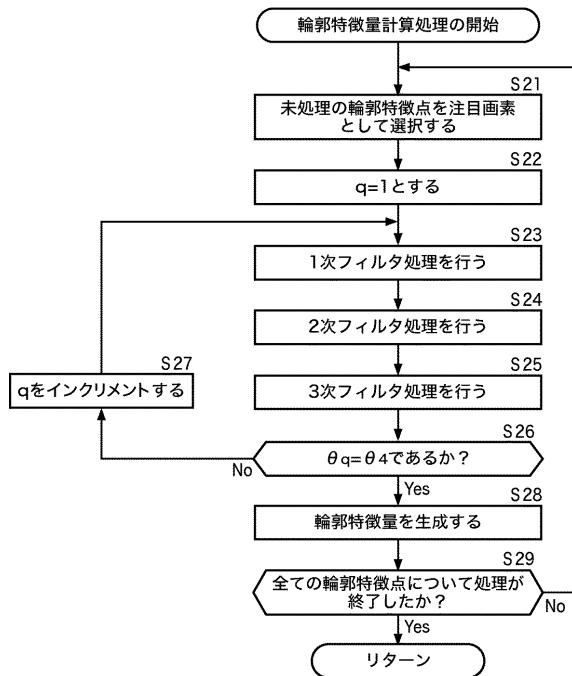


B



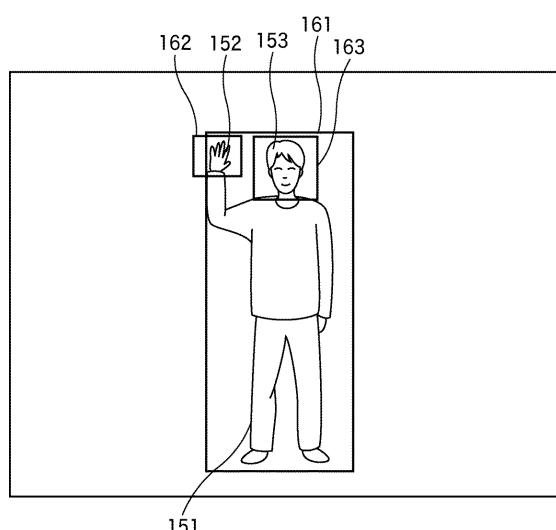
【図12】

図12



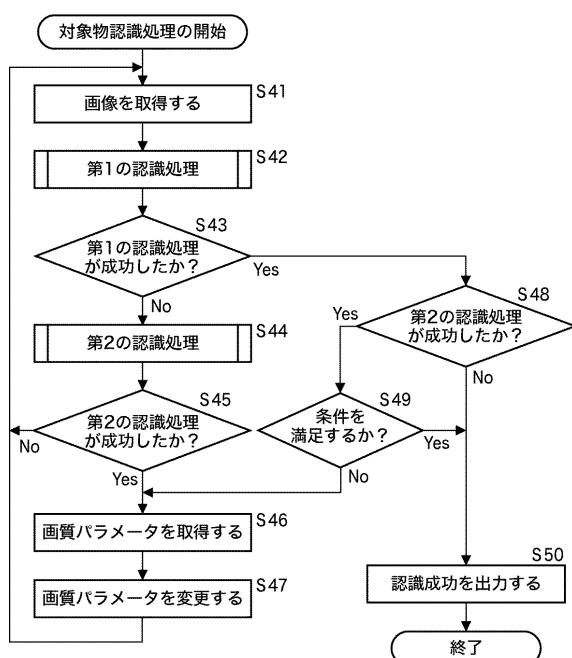
【図13】

図13

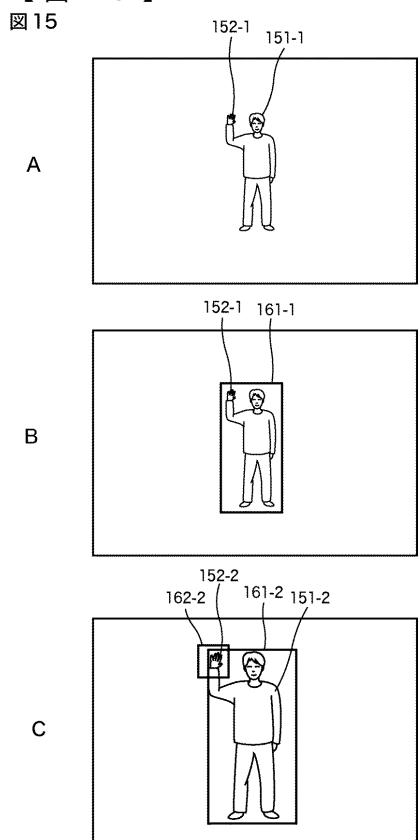


【図14】

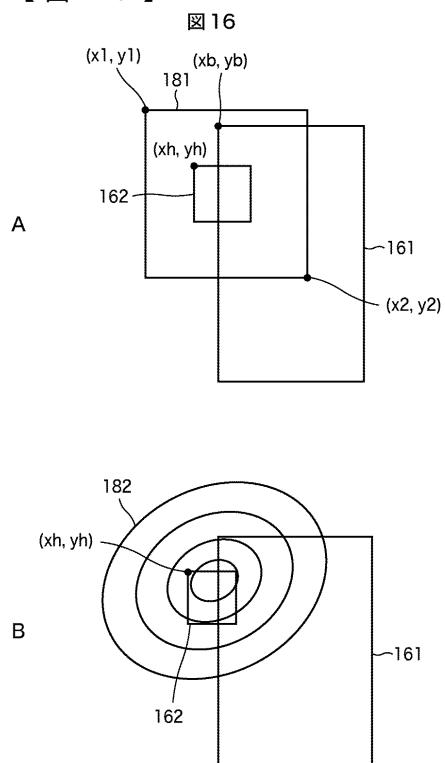
図14



【図15】

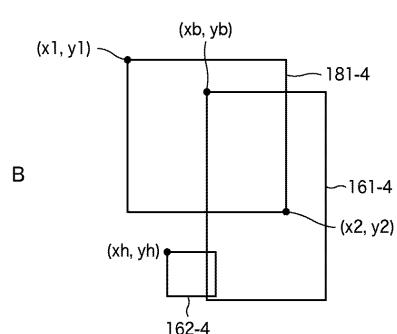
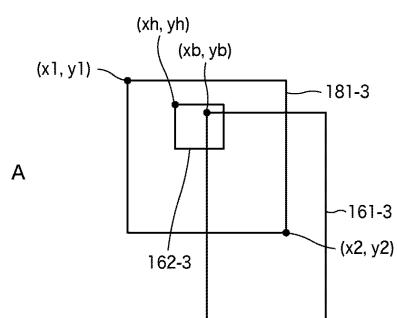


【図16】



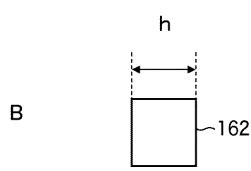
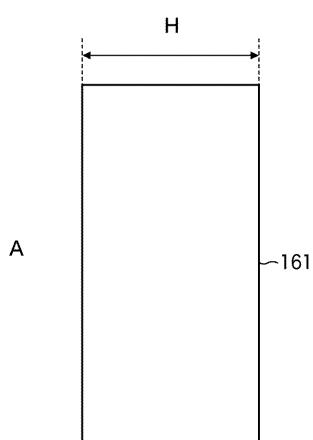
【図17】

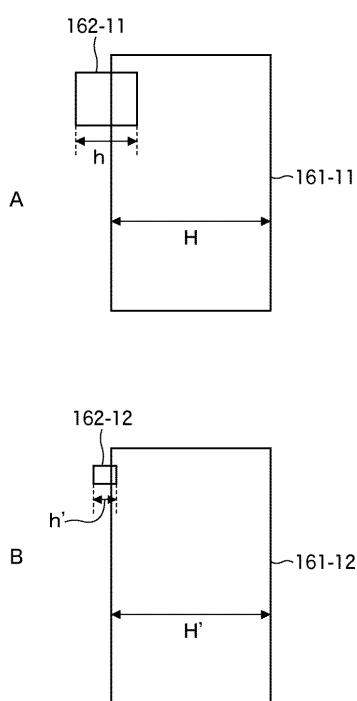
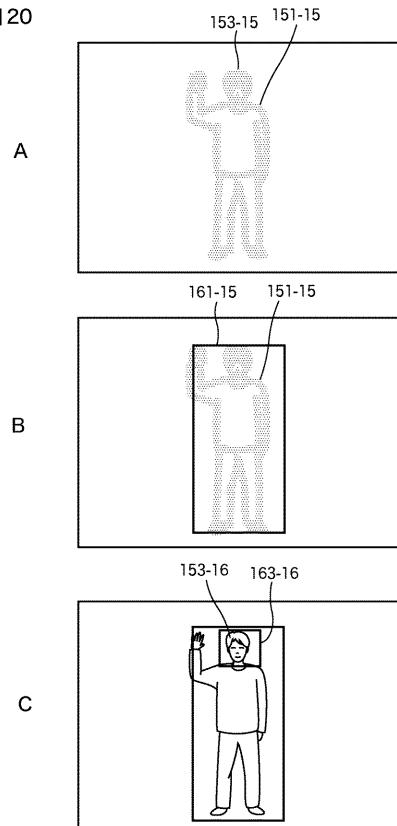
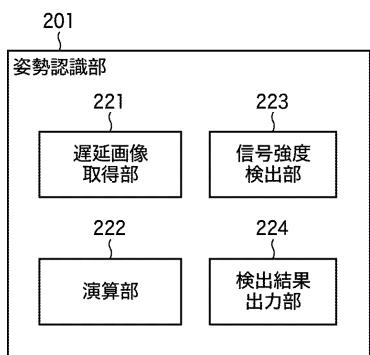
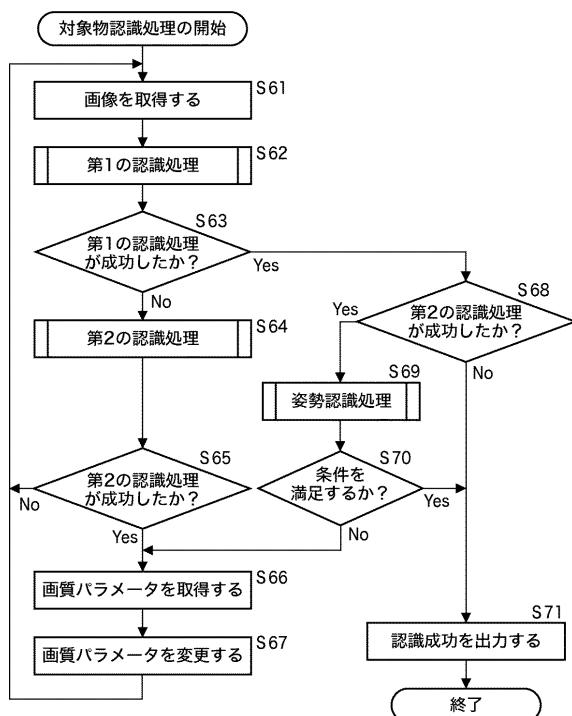
図17



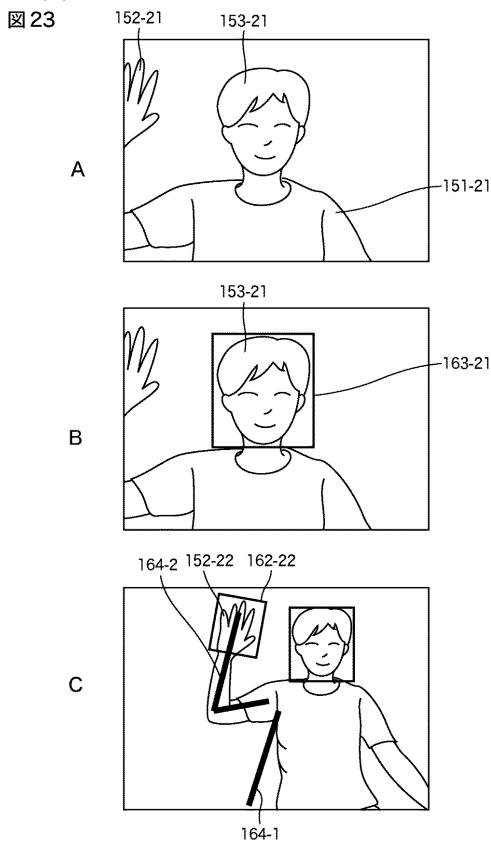
【図18】

図18

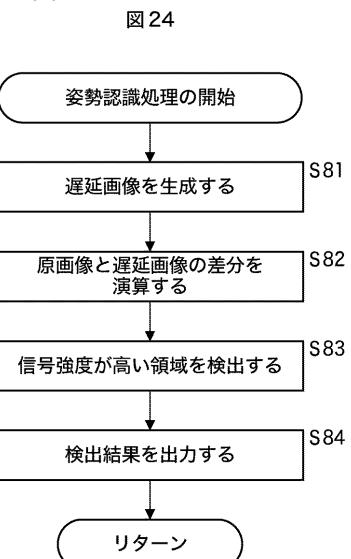


【図19】  
図19【図20】  
図20【図21】  
図21【図22】  
図22

【図23】

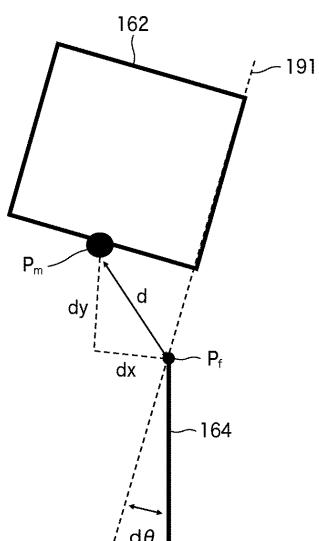


【図24】



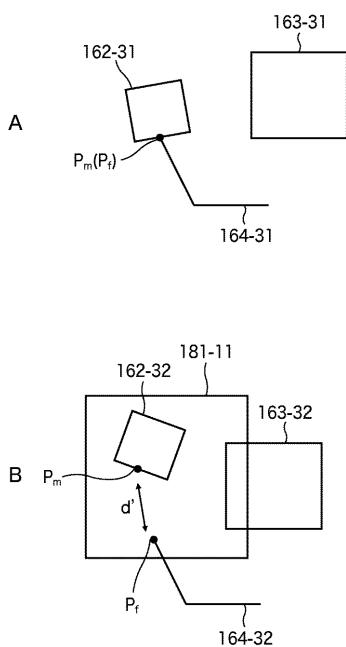
【図25】

図25



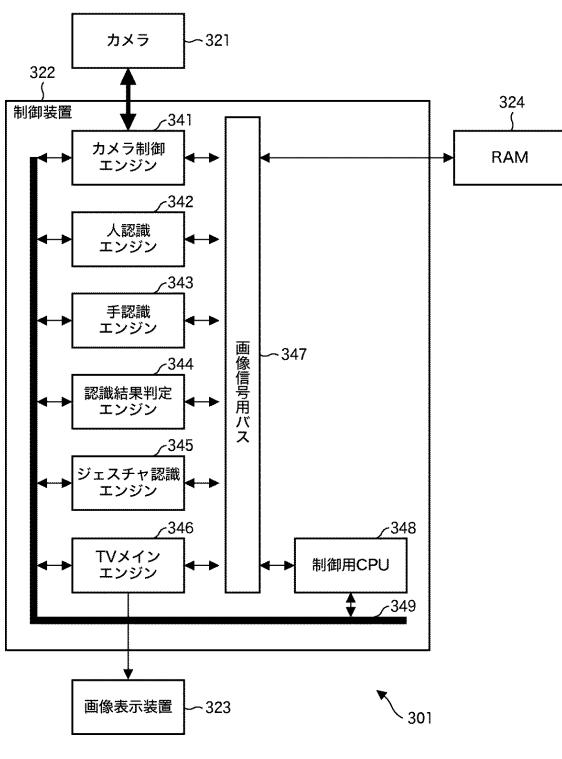
【図26】

図26



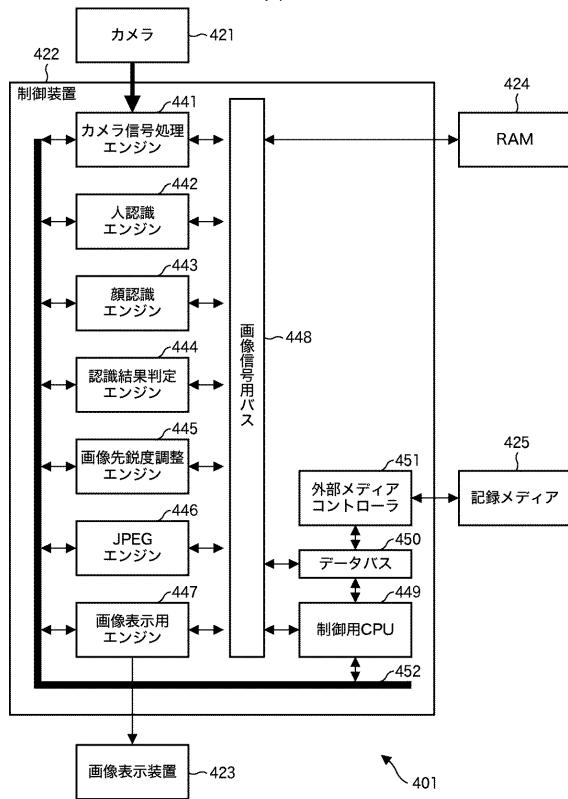
【図27】

図27



【図28】

図28



---

フロントページの続き

F ターム(参考) 5L096 AA06 BA08 BA18 CA02 CA06 DA02 EA03 FA06 FA09 FA59  
FA66 FA67 FA69 GA08 GA51 GA55 HA01 HA08 JA03 JA13  
KA04