

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4481663号
(P4481663)

(45) 発行日 平成22年6月16日(2010.6.16)

(24) 登録日 平成22年3月26日(2010.3.26)

(51) Int.Cl.

F 1

G06T 7/20 (2006.01)
G06T 1/00 (2006.01)G06T 7/20 300A
G06T 1/00 340B

請求項の数 12 (全 24 頁)

(21) 出願番号 特願2004-7924 (P2004-7924)
 (22) 出願日 平成16年1月15日 (2004.1.15)
 (65) 公開番号 特開2005-202653 (P2005-202653A)
 (43) 公開日 平成17年7月28日 (2005.7.28)
 審査請求日 平成18年12月7日 (2006.12.7)

(73) 特許権者 000001007
 キヤノン株式会社
 東京都大田区下丸子3丁目30番2号
 (74) 代理人 10012524
 弁理士 別役 重尚
 (72) 発明者 真継 優和
 東京都大田区下丸子3丁目30番2号 キ
 ャノン株式会社内
 (72) 発明者 森 克彦
 東京都大田区下丸子3丁目30番2号 キ
 ャノン株式会社内
 (72) 発明者 石井 美絵
 東京都大田区下丸子3丁目30番2号 キ
 ャノン株式会社内

最終頁に続く

(54) 【発明の名称】動作認識装置、動作認識方法、機器制御装置及びコンピュータプログラム

(57) 【特許請求の範囲】

【請求項 1】

画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段と、
 前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定
 されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列
 画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに
 関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに
 重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人
 物を特定する特定手段と、

前記時系列画像データから、前記特定手段で特定された人物の動作パターンを検出する
 状態検出手段と、

前記検出された動作パターンのデータを前記人物特有の動作パターンに関するモデルデ
 ータとして、前記特定された人物に固有の意味的情報に対応付けて保存する学習手段と、
 を備えることを特徴とする動作認識装置。

【請求項 2】

前記状態検出手段で検出された前記人物の動作パターンに対応するモデルデータが保存
 されているか否か判別する判別手段と、

前記判別手段で前記モデルデータが保存されていると判別した場合、当該モデルデータ
 に対応付けて保存されている前記人物に固有の意味的情報を出力する出力手段と、を更に

備え、

前記学習手段は、前記判別手段が前記モデルデータが保存されていないと判別した場合に、前記検出された動作パターンのデータを、前記人物特有の動作パターンに関するモデルデータとして、前記特定された人物に固有の意味的情報に対応付けて保存することを特徴とする請求項1記載の動作認識装置。

【請求項3】

画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段と

前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段と、

10

前記時系列画像データから前記特定手段で特定された人物が一定時間以上接触する物体を検出し、当該特定された人物が前記検出された物体に接触する頻度が所定の閾値より高い場合に、当該物体を当該特定された人物が常用する常用物として判定する常用物検出手段と、を備えることを特徴とする動作認識装置。

【請求項4】

画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段と

20

前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段と、

前記時系列画像データから、前記特定手段で特定された人物の居場所と動作パターンとを検出する状態検出手段と、

前記状態検出手段で所定の居場所と所定の動作パターンとが検出された頻度が、所定閾値以上である場合、当該所定の動作パターンを前記特定された人物の習慣的行動として判定する習慣的行動判定手段と、備えることを特徴とする動作認識装置。

30

【請求項5】

画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段と

前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段と、を備えることを特徴とする動作認識装置。

40

【請求項6】

所定の機器を制御する機器制御装置であって、

画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段と

前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人

50

物を特定する特定手段と、

前記時系列画像データから前記特定手段で特定された人物の動作パターンを検出する状態検出手段と、

前記検出された動作パターンのデータを前記人物特有の動作パターンに関するモデルデータとして、前記特定された人物に固有の機器制御情報であって前記所定の機器を制御するための機器制御情報に対応付けて保存する学習手段と、

前記状態検出手段で検出された前記人物の動作パターンに対応するモデルデータが保存されている場合、当該モデルデータに対応付けて保存されている前記機器制御情報を出力する出力手段と、を備えることを特徴とする機器制御装置。

【請求項 7】

10

認識装置の動体検出手段が、画像入力装置により入力された時系列画像データから動領域を検出する動体検出ステップと、

認識装置の特定手段が、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定ステップと、を備えることを特徴とする動作認識方法。

【請求項 8】

20

コンピュータを、

画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段、

前記時系列画像データから、前記特定手段で特定された人物の動作パターンを検出する状態検出手段、

30

前記検出された動作パターンのデータを前記人物特有の動作パターンに関するモデルデータとして、前記特定された人物に固有の意味的情報に対応付けて保存する学習手段、として機能させるためのコンピュータプログラム。

【請求項 9】

40

コンピュータを、

画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段、

前記時系列画像データから前記特定手段で特定された人物が一定時間以上接触する物体を検出し、当該特定された人物が前記検出された物体に接触する頻度が所定の閾値より高い場合に、当該物体を当該特定された人物が常用する常用物として判定する常用物検出手段、として機能させるためのコンピュータプログラム。

【請求項 10】

50

コンピュータを、

画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段と、

前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段、

前記時系列画像データから、前記特定手段で特定された人物の居場所と動作パターンとを検出する状態検出手段、

前記状態検出手段で所定の居場所と所定の動作パターンとが検出された頻度が、所定閾値以上である場合、当該所定の動作パターンを前記特定された人物の習慣的行動として判定する習慣的行動判定手段、として機能させるためのコンピュータプログラム。

【請求項 1 1】

コンピュータを、

画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段、として機能させるためのコンピュータプログラム。

10

20

【請求項 1 2】

コンピュータを、

画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段、

前記時系列画像データから前記特定手段で特定された人物の動作パターンを検出する状態検出手段、

30

前記検出された動作パターンのデータを前記人物特有の動作パターンに関するモデルデータとして、前記特定された人物に固有の機器制御情報であって前記所定の機器を制御するための機器制御情報に対応付けて保存する学習手段、

前記状態検出手段で検出された前記人物の動作パターンに対応するモデルデータが保存されている場合、当該モデルデータに対応付けて保存されている前記機器制御情報を出力する出力手段、として機能させるためのコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、動作認識装置、動作認識方法、機器制御装置及びコンピュータプログラムに関する。

40

【背景技術】

【0 0 0 2】

従来、画像から人物の行動パターンを検出する手法としては、隠れマルコフモデルを用い、標準的な行動の確率モデルを自動的に習得する行動検出システム（例えば、特許文献1，2）や、いわゆるジェスチャ認識方法の技術（例えば、特許文献3）が提案されている。

【0 0 0 3】

また、得られた行動パターンや動作の情報を利用した応用技術が提案され（例えば、特

50

許文献 4 ~ 11)、一方、動作パターンから個人認証を行う方法が提案されている(例えば、特許文献 12 ~ 14)。

【0004】

また、パスワード解析や指紋判定等から操作員を個別に識別し、その操作員にあった操作画面を備えるインターフェースがある(例えば、特許文献 15)。

【特許文献 1】特開平 11-259643 号公報

【特許文献 2】米国特許第 6212510 号明細書

【特許文献 3】特開平 10-162151 号公報

【特許文献 4】特開平 8-257017 号公報

【特許文献 5】特開平 9-81309 号公報

10

【特許文献 6】特開平 9-161102 号公報

【特許文献 7】特開平 11-249773 号公報

【特許文献 8】特開 2000-293685 号公報

【特許文献 9】特開 2000-101898 号公報

【特許文献 10】特許第 2801362 号公報

【特許文献 11】特許第 3140911 号公報

【特許文献 12】特開 2000-222361 号公報

【特許文献 13】特開 2001-34764 号公報

【特許文献 14】特開 2001-202524 号公報

【特許文献 15】特開平 10-149207 号公報

20

【発明の開示】

【発明が解決しようとする課題】

【0005】

しかしながら、上記の行動パターンを検出する手法及び検出された行動パターン情報を利用する方法では、対象となる人物等の動物体を個別に特定し、その特定された動物体に固有な動作や状態を認識し、学習することはできない。例えば、上記特許文献 6 に開示される方式においては、歩行パターンから個人の識別を行うので、歩行パターンが類似する人物の間での識別は困難である。従って、個人別にジェスチャ等の動作パターンに応じた環境設定、各種機器の操作方法又は機器の動作モードの設定(パーソナル化)を自動的に行なうことは困難であり、行動の癖や習慣的な行動や日常的に常用している物のカテゴリを個人別に自動学習することは困難であった。

30

【0006】

動作パターンから個人認証をする方法に関する上記従来例では、認証すべき対象の数が増えるほど該当する動作を厳格に規定しておかなければならず、一方では認証のために登録した動作が忠実に再現されるように利用者側が注意して動作を行わなければならないという点で負担が大きいものであった。そのため動作パターンのみでは個人認証を行うことは困難であった。

【0007】

本発明の目的は、対象別の動作の学習を簡便に行なうことができ、高い信頼度で対象を認識することができ、対象の動作に基づいて機器操作制御することができる動作認識装置、動作認識方法、機器制御装置及びコンピュータプログラムを提供することにある。

40

【課題を解決するための手段】

【0008】

上述の目的を達成するために、請求項 1 記載の動作認識装置は、画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段と、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第 1 類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第 2 類似度を評価し、前記第 1 類似度と前記第 2 類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段と

50

、前記時系列画像データから、前記特定手段で特定された人物の動作パターンを検出する状態検出手段と、前記検出された動作パターンのデータを前記人物特有の動作パターンに関するモデルデータとして、前記特定された人物に固有の意味的情報に対応付けて保存する学習手段と、を備えることを特徴とする。

【0009】

請求項2記載の動作認識装置は、請求項1記載の動作認識装置において、前記状態検出手段で検出された前記人物の動作パターンに対応するモデルデータが保存されているか否か判別する判別手段と、前記判別手段で前記モデルデータが保存されていると判別した場合、当該モデルデータに対応付けて保存されている前記人物に固有の意味的情報を出力する出力手段と、を更に備え、前記学習手段は、前記判別手段が前記モデルデータが保存されていないと判別した場合に、前記検出された動作パターンのデータを、前記人物特有の動作パターンに関するモデルデータとして、前記特定された人物に固有の意味的情報に対応付けて保存することを特徴とする。

10

【0010】

請求項3記載の動作認識装置は、画像入力装置により入力された時系列画像データから動領域を検出手段と、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段と、前記時系列画像データから前記特定手段で特定された人物が一定時間以上接触する物体を検出し、当該特定された人物が前記検出された物体に接触する頻度が所定の閾値より高い場合に、当該物体を当該特定された人物が常用する常用物として判定する常用物検出手段と、を備えることを特徴とする。

20

【0012】

請求項4記載の動作認識装置は、画像入力装置により入力された時系列画像データから動領域を検出手段と、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段と、前記時系列画像データから、前記特定手段で特定された人物の居場所と動作パターンとを検出する状態検出手段と、

30

前記状態検出手段で所定の居場所と所定の動作パターンとが検出された頻度が、所定閾値以上である場合、当該所定の動作パターンを前記特定された人物の習慣的行動として判定する習慣的行動判定手段と、備えることを特徴とする。

【0014】

請求項5記載の動作識装置は、画像入力装置により入力された時系列画像データから動領域を検出手段と、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段と、を備えることを特徴とする。

40

【0016】

請求項6記載の機器制御装置は、所定の機器を制御する機器制御装置であって、画像入力装置により入力された時系列画像データから動領域を検出手段と、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データ

50

タの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段と、前記時系列画像データから前記特定手段で特定された人物の動作パターンを検出する状態検出手段と、前記検出された動作パターンのデータを前記人物特有の動作パターンに関するモデルデータとして、前記特定された人物に固有の機器制御情報であつて前記所定の機器を制御するための機器制御情報に対応付けて保存する学習手段と、前記状態検出手段で検出された前記人物の動作パターンに対応するモデルデータが保存されている場合、当該モデルデータに対応付けて保存されている前記機器制御情報を出力する出力手段と、を備えることを特徴とする。

10

【0017】

請求項7記載の動作認識方法は、認識装置の動体検出手段が、画像入力装置により入力された時系列画像データから動領域を検出する動体検出ステップと、認識装置の特定手段が、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段と、前記

20

【0018】

請求項8記載のコンピュータプログラムは、コンピュータを、画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段、前記時系列画像データから、前記特定手段で特定された人物の動作パターンを検出する状態検出手段、前記検出された動作パターンのデータを前記人物特有の動作パターンに関するモデルデータとして、前記特定された人物に固有の意味的情報に対応付けて保存する学習手段、として機能させる。

30

【0019】

請求項9記載のコンピュータプログラムは、コンピュータを、画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段、前記時系列画像データから前記特定手段で特定された人物が一定時間以上接触する物体を検出し、当該特定された人物が前記検出された物体に接触する頻度が所定の閾値より高い場合に、当該物体を当該特定された人物が常用する常用物として判定する常用物検出手段、として機能させる。

40

【0020】

請求項10記載のコンピュータプログラムは、コンピュータを、画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段と、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力さ

50

れる検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段、前記時系列画像データから、前記特定手段で特定された人物の居場所と動作パターンとを検出する状態検出手段、前記状態検出手段で所定の居場所と所定の動作パターンとが検出された頻度が、所定閾値以上である場合、当該所定の動作パターンを前記特定された人物の習慣的行動として判定する習慣的行動判定手段、として機能させる。

【0021】

請求項1_1記載のコンピュータプログラムは、コンピュータを、画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段、として機能させる。

10

【0022】

請求項1_2記載のコンピュータプログラムは、コンピュータを、画像入力装置により入力された時系列画像データから動領域を検出する動体検出手段、前記動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記時系列画像データの前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、前記第1類似度と前記第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定する特定手段、前記時系列画像データから前記特定手段で特定された人物の動作パターンを検出する状態検出手段、前記検出された動作パターンのデータを前記人物特有の動作パターンに関するモデルデータとして、前記特定された人物に固有の機器制御情報であって前記所定の機器を制御するための機器制御情報に対応付けて保存する学習手段、前記状態検出手段で検出された前記人物の動作パターンに対応するモデルデータが保存されている場合、当該モデルデータに対応付けて保存されている前記機器制御情報を出力する出力手段、として機能させる。

20

【発明の効果】

30

【0029】

本発明によれば、検出された動作パターンのデータを特定された人物に固有の意味的情報に対応付けて学習するので、対象別の動作の学習を簡便に行うことができる。また、動領域の画像データから抽出される幾何学的特徴と予め対象となる人物ごとに設定されている幾何学的特徴に関するモデルデータとの間の第1類似度を評価し、前記動領域から抽出される動作パターンの特徴と人物特有の動作パターンに関するモデルデータとの間の第2類似度を評価し、第1類似度と第2類似度とに重み付けを行って出力される検出レベルに基づいて前記動領域の画像データに対応する人物を特定するので、高い信頼度で対象を認識することができる。更に、検出された動作パターンのデータを人物特有の動作パターンに関するモデルデータとして、特定された人物に固有の機器制御情報であって所定の機器を制御するための機器制御情報に対応付けて学習し、検出された人物の動作パターンに対応するモデルデータが保存されている場合、このモデルデータに対応付けて保存されている機器制御情報であって前記所定の機器を制御するための機器制御情報を出力するので、対象の動作に基づいて機器操作制御することができる。

40

【発明を実施するための最良の形態】

【0032】

以下、本発明の実施の形態を図面を参照しながら説明する。

【0033】

図1は、本発明の第1の実施の形態に係る動作認識装置の概略構成を示す図である。

【0034】

50

図1において、動作認識装置は、不図示の撮像装置又はデータベースから動画像データを入力する画像入力装置1と、入力された動画像データから動きベクトルその他の特徴を用いて動領域を検出して出力する動体検出装置2と、検出された動領域にある人物や物体等の対象の認識を行って、その認識結果を出力する被写体認識装置3と、被写体認識装置3が認識した対象となる人物の動作や状態を検出する状態検出装置4と、状態検出装置4が検出した動作又は状態がその対象となる人物に対応した固有の意味を有する内容を学習し、所定の記憶手段に格納する学習装置5と、動体検出装置2、被写体認識装置3、状態検出装置4、及び学習装置5が夫々行う処理の起動、停止及び処理モードの変更等を制御する処理制御装置6とを備える。

【0035】

10

ここで、動作とは、対象が人物である場合、身振り、手振り、目付き、表情の変化等、身体の動きを伴う変化を指す。また、状態とは、「座っている」、「倒れている」等の殆ど静止している所定の姿勢、又は「食事をしている」、「笑っている」等一連の動作の変化の組み合わせであって、特に何らかの意味的解釈ができる場合をいう。以下で用いる動作又は状態カテゴリは、「食事をしている」、「歩いている」等複数の身体部分を同時に動かす意味的動作のほか、「手を振る」、「髪の毛を触る」、「手を挙げる」等のように身体の限られた部分を使った運動を表すものであってもよい。

【0036】

図2は、図1の動作認識装置によって実行される機器制御処理のフローチャートである。

20

【0037】

図2において、まず、動体検出装置2は、画像入力装置1により入力された時系列画像データから動領域を検出し(ステップS101)、被写体認識装置3は、後述する図3の被写体認識処理を実行して、動領域の画像データを用いて動作を行う主体である対象となる人物の認識を行い(ステップS102)、状態検出装置4は、後述する図4の動作認識処理を実行して、その対象となる人物の動作に対応した動作カテゴリを抽出し(ステップS103)、この対象となる人物及び動作カテゴリに固有の機器制御情報を学習装置5内の不図示の辞書データ作成部から抽出する(ステップS104)。

【0038】

30

次いで、処理制御装置6は、対象となる人物に対し、ディスプレイへのメッセージ表示又は音声により、抽出された機器制御情報の内容の確認動作(例えば、身振り、表情、額き又は首振りの動作、音声による肯定又は否定等)を要求し(ステップS105)、対象となる人物が抽出した機器制御情報に対して肯定する確認動作をしたか否かを判別し(ステップS106)、肯定する確認動作をしたときは、抽出された機器制御情報を機器制御信号として出力して(ステップS107)、本処理を終了し、肯定する確認動作をしなかつたときは、機器制御信号を出力することなく、本処理を終了する。対象となる人物からの確認動作は、対象となる人物毎に予め設定してもよく、又は、共通のサインやメッセージにてもよい。

【0039】

40

図2の処理によれば、対象となる人物及び動作カテゴリに固有の機器制御情報を学習装置5内の不図示の辞書データ作成部から抽出する(ステップS104)ので、対象となる人物の動作に基づいて機器操作制御することができる。

【0040】

機器制御情報は、例えば、左手の指を握り、部屋の中のある機器に向かって突き出すというある対象となる人物の動作に対して、「その機器を起動する」という機器制御内容を設定することができる。機器操作内容は、「テレビをONにする(OFFにする)」、「テレビの音量を上げる(下げる)」、「照明をONにする(OFFにする)」、「お湯を沸かす」、「コンピュータを起動する」、「コンピュータの特定のプログラムを起動する」等、多種多様に設定することができる。なお、動作によって制御すべき機器が特定されるので、必ずしも制御したい機器に対してその対象となる人物が注目している状態にある

50

ことを示す行為を行わなくてもよい。

【0041】

また、機器制御内容は、上記のような単一の操作制御カテゴリに分類付けられる内容以外に、例えば、「テレビをONにして、照明を暗くする」のように、一連の制御の組み合せであってもよい。

【0042】

このようにして対応付けられる動作と機器制御内容とは、個々の対象となる人物に応じて設定され、上述したように姿と動作とから認証された人物以外は操作又は制御を行えないように設定することもできる。

【0043】

また、機器制御情報を抽出する為の動作は、日常的にその対象がとる動作とは明らかに異なる動作であればよい。

【0044】

図3は、図2のステップS102の被写体認識処理のフローチャートである。

【0045】

本処理において、学習装置5は、予め特定人物の複数の姿勢に対応した画像データのモデルデータ、及び人物特有の動作に関するモデルデータを被写体カテゴリと対応付けて不図示の記録装置に記録しているものとする。

【0046】

図3において、まず、動領域の画像データから画像データの幾何学的特徴又は色成分に関する特徴を抽出し(ステップS201)、抽出した画像データの特徴と予め対象となる人物ごとに設定されたモデルデータとの類似度を評価する(ステップS202)。この類似度は、モデルデータとの相関係数の値、又は特定被写体カテゴリの顕著度として評価する。顕著度は、例えばモデルデータとの相関係数最大値 C_{max} と次の順位の相関係数値 C_{next} との比 C_{max} / C_{next} 、又は被写体カテゴリ別の相関係数分布の尖度即ちkurtosisを用いて評価する。ここに分布aのkurtosisは、aの分散 σ^2 と平均 \bar{a} 及びサンプル数Nを用いて

$$kurtosis(a) = (1/N) \cdot (a - \bar{a})^4 / (\sigma^2)^2 - 3$$

で与えられる。平均 \bar{a} はaの平均を示す。

【0047】

幾何学的特徴の類似度は、例えば、局所的な領域で検出される複数の特徴を比較的単純なパターン(例えば、所定の方向成分を有する線分レベル、その組み合わせで得られるV字、L字、T字型コーナパターンレベル等)、単純なパターンの組み合わせ又は相対位置関係等により記述される中程度の複雑さを有するパターン、及び中程度の複雑さを有するパターンの組み合わせにより得られる上位の複雑パターンと、階層的に用意しておき、各階層レベルにおける局所的なテンプレートマッチングすることにより評価する(本出願人による特許第3078166号公報)。

【0048】

次いで、被写体認識装置3は、状態検出装置4から抽出された動作パターンの特徴を入力し、動作パターンの特徴と、動作に関するモデルデータとの類似度評価を行い(ステップS203)、画像データの特徴と動作パターンの特徴についての信頼度を算出し(ステップS204)、算出した信頼度に応じて各特徴の類似度を重み付けして出力し(ステップS205)、出力された類似度を検出して最大検出レベルが閾値以上であるか否かを判別し(ステップS206)、最大検出レベルが閾値以上であるときは、最大検出レベルを与えるモデルデータに対応する被写体カテゴリを被写体認識結果として出力し(ステップS207)、最大検出レベルが閾値以上でないときは、画像データ又は動作パターンの特徴に対応付けた被写体カテゴリの学習を行って(ステップS208)、本処理を終了する。

【0049】

本処理は、例えば、人物の上半身、特に、頭部の検出を行って得られる頭部の画像とそ

10

20

30

40

50

の人物の動作とから行う。学習装置5は、予め特定人物の複数の姿勢に対応した人物頭部のモデルデータ、及び人物特有の動作パターンに関するモデルデータを不図示の記録装置に格納している。なお、モデルデータの生成を行うべき人物の姿勢は、離散的に所定の角度幅でサンプリングしておけばよく、連続的に殆ど全ての姿勢状態をカバーするようにモデルデータを大量に用意する必要はない。

【0050】

ここで行うマッチングの処理方式としては、既存のモデルデータのサイズを動体検出装置2で検出される動領域のサイズに応じて変えて生成し、テンプレートマッチングを行う方式、又はマッチングの処理性能が被写体サイズの変化に対して不变であるような方式(本出願人による特許第3078166号公報)のいずれを用いてもよい。

10

【0051】

また、マッチング結果の信頼度が低いと予測される場合(例えば、対象が後ろ姿である場合、照明の影響で頭部の顔画像のコントラストが低い場合等)には、被写体カテゴリの抽出は、動作カテゴリとの整合性判定により行う。具体的には、画像ベースで推定された対象に固有の動作特徴(例えば、歩行パターン、姿勢)を有しているか否かの判定を行う。この場合、被写体認識装置3は学習装置5内部の辞書データを参照し、その対象が顕著な動作特徴を有していない場合には、対象の判別が困難であることを出力し、これを受けた処理制御装置6は以降の処理を中断する。

【0052】

画像データの特徴についての信頼度は、被写体カテゴリの候補を検出するのに必要な特徴の類似度の検出レベルと、その画像データに対応する視点位置(正面位置か背面位置か等)と、検出された被写体の向きとによって重み付けをして評価し、被写体の認識に最適な視点位置又は方向で最も重み係数が高く、逆に認識に適さない視点位置からの画像については重み係数を低くする。

20

【0053】

例えば、人物の認証のために正面の顔を所定サイズ以上検出できたときは、最大の信頼度とし、後ろ向きの姿は正面と比べて一般的にパターンとしての識別性が低い場合が多いので、低い信頼度を与えるように予め設定しておく。ただし、後ろ向きの姿であってもその特定人物に顕著な特徴を有しており、その特徴が検出された場合には、信頼度は高くなるように修正する必要がある。従って信頼度は、顕著度(個性の指標)のファクタ-を乗算した値を用いる。なお、顕著度がある閾値より低い場合には、所定の定数(例えば1)となるようにしておく。

30

【0054】

顕著度(個性の指標)の算出手順としては、例えば、平均的な人物頭部の後ろ姿(男女別)を画像ベクトル又は特徴抽出後の特徴ベクトルにより求め、その特徴ベクトルに関する共分散行列から固有ベクトルを求める。次に、顕著度を固有ベクトル空間での座標値の平均ベクトルからの偏差(差の絶対値)として抽出する。

【0055】

一方、動作パターンの特徴の信頼度は、時系列特徴データのSN比、時系列特徴データのモデルデータに対する類似度と閾値の比、又は所定時間幅で算出される最大類似度の分散値等により評価される。また、検出された動作パターンが対象となる人物に特有なものか否かを表す指標(即ち、個性の程度)を予めその動作特徴カテゴリについて評価しておき、その指標を信頼度に乗じてもよい。

40

【0056】

動作パターンを検出する目的は、機器制御のための意味的情報を抽出することと個体の認証を行うことの2つがある。画像データで姿のみから被写体認識を行って動作パターンを検出して後述するような機器制御を行うか、又は、動作パターンのみから被写体認識を行い、更に異なる動作パターンから機器制御情報を抽出することにより、機器制御を対象となる人物の動作に基づいてパーソナライズすることができる。また、画像パターンの特徴と動作パターンの特徴とから被写体の認証を行い、特定の動作パターンから機器制御を

50

行うこともできる。

【0057】

上述した認証目的の動作検出においても信頼度を上げるために、対象となる人物の動作パターンと画像データとを組み合わせて認証を行う。例えば、対象となる人物の歩行パターンデータと頭部の画像とにより行う。歩行パターンの検出は画像からシルエットパターンを抽出して、そのシルエットの輪郭線から特徴量（水平・垂直方向のヒストグラム等）を抽出し、その特徴量の変化のパターンを照合する方式等を用いればよい。

【0058】

動作の個性に関する特徴量として、動作パターン検出に使われる時系列特徴データから平均的な動作パターンを抽出し、更に時系列特徴データの共分散行列から固有動作ベクトルを抽出し、その固有ベクトル空間内の座標値の平均ベクトルからの偏差を個性として抽出してもよい。

10

【0059】

図3の処理によれば、姿と動作とを組み合わせて認証を行うことにより（ステップS202～ステップS203）、高い信頼度で対象となる人物を認識することができる。

【0060】

図4は、図2のステップS103の動作認識処理のフローチャートである。

【0061】

本処理は、上記図2の被写体認識処理（ステップS102）の終了後に実行され、特定された人物が意図する所定の意味を持った動作又は状態の検出を行う。ここでは、動作に関する身体各部（顔、目付き、腕、手先、足等）の検出、検出された特徴点又は特徴量の抽出、及び特徴点又は特徴量の時系列的変化の分析を行うが、それ以外の方法で行ってもよい。

20

【0062】

学習装置5内部の記録装置には、画像データから抽出される特徴点又は特徴量のモデルデータと対応付けられた動作カテゴリが予め格納されているものとする。

【0063】

図4において、被写体認識装置3から認識結果である被写体カテゴリを入力し（ステップS301）、動作パターンの特徴を抽出する前段階としての被写体に関するプリミティブな特徴量を抽出する（ステップS302）。ここでは、動作パターンの検出に用いる特徴量として、背景画像との差分画像から抽出される幾何学的特徴又は色成分ベースの特徴量を用いるか、又はエッジ抽出を画像フレーム毎に行い、そのフレーム間差分により得られるエッジ特徴を用いる。

30

【0064】

次いで、被写体の画像に形状モデルを当てはめる（ステップS303）。例えば、被写体が人物である場合には、動作パターン検出のための特徴抽出は、予め単純化された3次元パーツモデル（円筒、楕円体、super-quadrics等）により近似された人体各部の形状モデルを用いた人体モデルに基づき、人体各部のパーツ検出と各パーツの特徴量を検出することにより行うか、又は、人体頭部の検出を先に行い、次に頭部の位置・姿勢を基準として腕、手、足等の存在範囲を推定する。後者の場合、該当する領域で肌色検出又は動きベクトル分布の検出を行うことにより各パーツの検出を行ってもよい。

40

【0065】

次いで、対象となる人物を構成するパーツごとの特徴量（局所的かつ幾何学的な特徴、動きベクトル、色成分特徴等）をパーツの代表点付近での平均的特徴量として所定の記憶手段に格納し、パーツ単位での代表点位置の時系列画像データを抽出し（ステップS304）、抽出した時系列画像データとモデルデータとの照合を行い（ステップS305）、時系列画像データに対応するモデルデータが存在する否かを判別し（ステップS306）、時系列画像データに対応するモデルデータが存在するときは、モデルデータとの類似度が最も高い動作カテゴリを出力して（ステップS307）、本処理を終了し、時系列画像データに対応するモデルデータが存在しないときは、時系列画像データに対応付けた動作

50

カテゴリの学習を行って(ステップS308)、本処理を終了する。

【0066】

図4の処理によれば、時系列画像データに対するモデルデータが存在しないときは(ステップS306でNO)、時系列画像データに対応付けた動作カテゴリの学習を行う(ステップS308)ので、対象となる人物別の動作の学習を簡便に行うことができる。

【0067】

なお、人物動作の抽出技術の詳細については、文献(星野、「人物動作における個性の抽出と再構成」画像電子学会誌第30巻、pp.631-640, 2001)等に記載の技術を用いればよい。モデルデータとのマッチング処理の基本的な方法としては、入力画像の遷移系列とモデルとの対応付けをDynamic time warpingという手法で行う方法(T. J. DARELL and A.P. PENTLAND, 1993 "Space-Time Gestures,")、動作の状態遷移を確率モデルで行う隠れマルコフモデルを用いた手法(J. Yamato, J. Ohya, and K. Ishii, 1992 "Recognizing Human Action in Time-Sequential Images Using Hidden Markov Model," Proc. CVPR, pp. 379-385)、時空間エッジ画像ベクトル列の連続動的計画法(Continuous Dynamic Programming)による方法(西村、岡「ジェスチャ動画像の逆時間動作や静止動作をスポットティング認識するためのNon-monotonic連続DP」、信学技報、PRMU96-32、pp. 49-56, 1996)等がある。

10

【0068】

状態検出装置4で行われる動作の認識は、上述した方法以外に、動作カテゴリに対応するモデル特徴量の時系列パターンと検出された特徴量の時系列パターンとの所定の特徴量空間での距離を求め、最小距離を与えるような動作カテゴリを認識結果としてもよい。また、その最小距離が所定の閾値以下であることを条件とし、最小距離がその閾値より大となる場合には、未分類の動作カテゴリが検出されたことを表すようにしてもよい。

20

【0069】

学習装置5で行われる動作パターンの学習処理(ステップS308)は、動作パターンを表す特徴量系列とモデル特徴量系列データとの距離が一定閾値より小さくならない場合は、検出された動作に対応する特徴点又は特徴量に関する時系列画像データ又は時系列画像データを保存し、またその動作パターンに対して仮のラベルが与えられ、該当する動作パターンの意味的情報を登録する為の学習シーケンスが処理制御装置6により起動される。学習シーケンスとしては、音声ベースでの対話により行う方法を用いてもよい。

30

【0070】

また、動作カテゴリの照合ができずに記録保持された動作パターンの画像データを所定のディスプレイに再生し、その動作パターンを人物が確認した後で該当する機器の操作を行うと動作パターンと機器操作との対応付けが自動的に行われるようにもよい。なお、この確認は、音声、動作、キーボード操作、ボタンスイッチの操作等、予め定められた方法で行えばよい。

【0071】

さらに、動作カテゴリの照合ができなかったことを人物に示す表示をディスプレイ上の表示又は所定のランプの点灯、又は音声等により行うことにより、システムが直前に行われた動作に対応する意味的情情報を登録可能な状態にあることを人物に知らせたうえで、その人物が機器操作を行うと、その機器操作内容が直前の動作と自動的に対応付けられるようにもよい。

40

【0072】

登録済みの動作パターンと機器制御情報との関連付けに対して、修正や再登録を行うこともできる。修正とは、同じ動作パターンであっても異なる機器制御情報に変更することであり、再登録とは、所定の機器制御情報と対応付けられた所定の動作を類似の動作パターンに変更することを意味するものとする。

【0073】

その方法としては、修正又は再登録モードを意味するものとして対象となる人物によらずに予め標準的に定められる動作又は状態をその対象となる人物が本発明に係る動作認識

50

装置に対して示すか、動作認識装置が埋め込まれた機器に付属する所定の端末上のボタン操作等により修正又は再登録モードに自動的にセットされるように構成しておけばよい。

【0074】

更に、本実施の形態で説明した動作認識装置を用いれば、同一機器を複数の人物が動作により操作することができると共に、人物間で優先順位をつけておくこともできる。また、人物別の優先順位を時間帯ごとに設定してもよい。

【0075】

本実施の形態で説明した動作認識装置をマンマシンインターフェースに用いる場合には、身振り等の動作と対応付けられた操作意志の伝達に先立って機器操作を意図しているか否かにかかわらず、対象となる人物の認証が自動的に行われる。また対象となる人物の行う動作は機器操作情報と対応付けられているので、人物は操作対象の注目意志の伝達を前もって行わなくてよいことはいうまでもない。

10

【0076】

上述した動作認識装置は、自律的にユーザの意図を学習するマンマシンインターフェース装置として家庭又はオフィス内の機器（例えば、テレビ、ステレオ、パソコン、ディスプレイ、ロボット、照明器具、エアコン、冷蔵庫）や、机、ロボット等に搭載し、機器制御に用いることができる。また、建物内部の天井、壁面、扉等に埋め込んで身振り、表情等から個々の人物の意図する機器操作内容を学習し、検出する自動機器制御システムを構成することができる。

20

【0077】

上記第1の実施例によれば、図4の処理により対象別の動作の学習を簡便に行うことができ、図3の処理により高い信頼度で対象を認識することができ、図2の処理により対象の動作に基づいて機器操作制御することができる。

【0078】

図5は、本発明の第2の実施の形態に係る動作認識装置の概略構成を示す図である。

【0079】

図5の動作認識装置は、学習装置5が後述する図6に示す構成を有する点、状態検出装置4及び学習装置5に接続され、常用物の検出及び登録を行う常用物検出装置7を備えている点で図1の装置と異なるのみであって、他の構成要素は図1のものと同じである。常用物とは、対象となる人物が日ごろ使う頻度の高い道具、器具、装置等の物体（例えば、対象となる人物が日常的に用いるコーヒーカップ、寝具、歯ブラシ、本）をいう。

30

【0080】

本実施の形態の動作認識装置は、建物内に設置された状態監視システム（複数台のカメラと一つ又は複数の本実施の形態に係る動作認識装置とを含む）に組み込むことにより、特定人物の常用物の所在を特別な発信機やマーカを取り付けずに短時間で検出することができる。

【0081】

図6は、図5における学習装置5の内部構成を示すブロック図である。

【0082】

図6において、学習装置5は、カテゴリを学習して記録する物体カテゴリ辞書データ生成記録装置50と、音声又はディスプレイ上の表示により対象となる人物に質問をするダイアログプロセス制御装置51とを備える。

40

【0083】

図7は、図5における常用物検出装置7によって実行される常用物検出処理のフローチャートである。

【0084】

図7において、常用物の検出に先立って、特定人物を検出し（ステップS401）、後述する図8の常用物候補抽出処理を行って、その特定人物が一定時間（例えば3秒）以上接觸する物体を検出し（ステップS402）、学習装置5の内部にある物体カテゴリ辞書データ生成記録装置50を検索して、既に登録されている常用物候補のモデルデータと検

50

出された特徴データとの類似度の評価を行い(ステップS403)、検出された特徴データに対応する物体カテゴリーが物体カテゴリー辞書データ生成記録装置50に登録されているか否かを判別する(ステップS404)。

【0085】

ステップS402において、接触とは、手、足、又は体全体が所定の物体に触れる行為(例えば、把持する、掴む、寝る、履く、撫でる、叩く、着る)をいい、この接触状態を検出するためには、接触する身体の各部の検出と接触対象である物体の検出を行う必要がある。ここでは接触の例として物体を把持する場合を考えているが、この場合には手首から先の運動と一体的に動く物体を接触された物体として検出する。

【0086】

ステップS404の判別の結果、検出された特徴データに対応する物体カテゴリーが登録されていないか、又は判別が困難であるときは、常用物候補の学習を行った(ステップS405)後、検出された物体の特徴データに対応する物体カテゴリーが登録されているときは、ステップS405の学習を行うことなく、常用物候補が特定人物の手の先と接触する頻度を週単位又は月単位で算出し(ステップS406)、算出した頻度が所定の閾値(例えば週5回等)より高ければ常用物と判定し(ステップS407)、本処理を終了する。

【0087】

ステップS405の学習処理は、学習装置5の内部にあるダイアローグプロセス制御装置51が起動され、対象となる人物に対して音声又はディスプレイ上の表示により、その物体の名称を示すよう要求する。例えば、常用物候補の画像を所定のディスプレイに表示して、「この物は何ですか?」という質問を音声又は文字表示により行う。この質問に対する音声又は文字入力等による応答を認識、解析して得られる物体の名称を、既に格納されている画像データ及び関連情報と共に物体カテゴリー辞書に登録する。例えば、接触を行っている主体である特定人物の名称とその常用物カテゴリー名称又は常用物候補の画像データ、及び接触の平均頻度又は累積回数とを記録保存すればよい。

【0088】

図8は、図7のステップS402の常用物候補抽出処理のフローチャートである。

【0089】

図8において、まず、動領域としての人物頭部を検出する(ステップS501)。頭部の検出は、画像中の動きを伴う橙円近似可能な動領域であって、顔の検出が行われること、又は、毛髪色の検出(後頭部の場合)又は肌色成分の検出(正面から側面の範囲で検出される場合)がされることを条件として行われる。

【0090】

次いで、手の先(又は手のひら、手の甲、指先等)を検出する(ステップS502)。手の先の検出は、頭部近傍にある肌色成分を有する動領域であって頭部ではない領域の先端領域を手領域候補として検出し、その候補領域について手首から先の部分に関するテンプレートマッチング等を適用して行う。ここで先端領域は手首から指先までの領域を指し、該当する領域の3次元形状モデルを用い、その形状モデルを用いて透視変換により生成される画像データと実際の画像データとの類似度評価又は対応付けを行うことにより抽出する。手領域候補の大きさは検出された頭部サイズを基準に定められ、頭部サイズより著しく大きいか又は小さい領域は候補から除かれる。手の先の検出としては、このような方法に特に限定せず他の方法を用いてもよいことは言うまでもない。

【0091】

次いで、予め動きに関する一体性を評価するために、手先領域の動きベクトルを検出し(ステップS503)、手の領域(手先を含む)に隣接する領域であって動きベクトルがほぼ手領域と一致する領域を動きベクトルの分布に基づいて領域分割することにより抽出し(ステップS504)、その領域の物体を常用物の候補とし(ステップS505)、その物体領域の追跡を行う(ステップS506)。把持動作後の非接触状態、即ち手の先が候補物体から離れた後、この検出された物体の画像データを所定の記憶手段に格納し(ステップS507)、常用物候補を認識するための特徴抽出を行い(ステップS508)、

10

20

30

40

50

本処理を終了する。

【0092】

上記第2の実施の形態によれば、図7において、図8の常用物候補抽出処理を行って特定人物が一定時間以上接触する物体を検出し(ステップS402)、検出された特徴データに対応する物体カテゴリが物体カテゴリ辞書データ生成記録装置50に登録されていないか、又は判別が困難であるときは、常用物候補の学習を行う(ステップS405)ので、特定人物の常用物の検出をその人物と対応付けて学習することにより、ある物体が不特定多数の人により使われるのか否かを判定することができる。

【0093】

また、本実施の形態に係る動作認識装置を監視システム又はロボット等に組み込むことにより、特定人物が常用する物体を第三者が使おうとした場合に警告を発することができる。

10

【0094】

図9は、本発明の第3の実施の形態に係る動作認識装置の概略構成を示す図である。

【0095】

図9の動作認識装置は、状態検出装置4に接続され、後述する図12の注視物体カテゴリ抽出処理を行う注視対象検出装置9と、状態検出装置4に接続され、対象となる人物の習慣的行動の検出及び学習を所定の視点位置からの時系列画像を用いて行う習慣的行動判定装置10とを備えている点で図5の装置と異なるのみであり、他の構成要素は図5のものと同じである。

20

【0096】

習慣的行動とは、ある対象となる人物がとる行動のうち場所と時間帯の少なくとも一方に関連付けられ、かつ、頻度の比較的高いものをいう。このような習慣的行動を学習し、検出することにより、人物の認証への適用や習慣的行動パターンに対応した機器や環境の自動設定を行うことができる。人物の認証へ適用する場合には、このような習慣的行動パターンの検出を主とし、当該人物の顔その他の姿からの認証を補助的に行うことにより、視点位置から比較的遠方に居る人物の認証を行う目的等に使うことができる。

【0097】

状態検出装置4により検出される状態とは、一般的には前述したように殆ど静止の所定の姿勢又は動きの変化が小さい場合である。ここでは、状態検出装置4は後述する図10の状態カテゴリ20に対応する時系列画像と特徴量データを記録し、対象となる人物の姿勢、表情、基本動作カテゴリ21の検出、及びその対象となる人物が注視している又は接触している目的物体カテゴリ23の検出により、意味解釈のできる状態カテゴリの抽出を行う。

30

【0098】

意味的解釈ができる状態とは、「テレビを見ている」、「本を読んでいる」、「食事をしている」、「瞑想をしている」等のように目的物体(テレビ、本等)と動作(見る、する等)との対応付けが可能な場合、又は、「食事」、「瞑想」等のように目的物体が物理的実体ではなく、かつ一義的に動作との対応付けは存在せず、予め意味のある状態として表情又は手足の動作に関する複数の動作パターン及び場所又は時間帯が学習装置5の辞書データ作成記録手段50内部の記憶手段に記憶されている場合をいう。

40

【0099】

図10は、状態カテゴリの辞書データ内部での表現形式を示す図であり、(a)は行為の対象となる目的物体を有する場合、(b)は目的物体を有しない場合を示す。

【0100】

図10(b)において、目的物体を有する状態カテゴリ20は、姿勢、表情、基本動作カテゴリ21、観測される居場所又は時間帯22、及び目的物体カテゴリ23により記述される。目的物体カテゴリ23は、接触物体カテゴリ24及び注視物体カテゴリ25の少なくとも1つから成る。

【0101】

50

例えば、「テレビを見ている」という状態カテゴリ20は、「注視している」という基本動作カテゴリ21、及び「テレビ」という注視物体カテゴリの2つのカテゴリを検出することにより状態が記述される。

【0102】

図10(a)において、目的物体を有しない状態カテゴリ20(例えば、「瞑想している」等)は、姿勢、表情、基本動作カテゴリ21、及び観測される居場所又は時間帯22により記述される。

【0103】

これにより、注視物体も接触物体も有しない状態(例えば、「瞑想」については、目を閉じて体を起こした姿勢を所定時間継続した状態として定義)についても意味的解釈のある状態として予め簡単な複数の構成要素(例えば、「瞑想」については、「目を閉じている」、「上半身の(半)直立固定姿勢」、「前方固定の顔の向き」等のように)の組み合わせとして表現しておくことができる。

【0104】

本実施の形態においては、殆どの場合、少なくとも接触物体及び注視物体の少なくとも1つからなる物体カテゴリ、及び動作カテゴリを検出することにより特定可能な状態カテゴリが検出される。状態カテゴリを特定する際に、前述したように観測される場所又時間帯に関する情報22を更に抽出してもよい。

【0105】

これにより、例えば、ダイニングテーブルを前にして(居場所に関する情報)接触又は注視対象として「料理の盛られた食器」が検出され、動作カテゴリとして「手を口の近傍に動かす」が検出された場合を「食事をする」の状態カテゴリとする。なお、同じ状態カテゴリであっても接触対象(注視対象)カテゴリ及び動作カテゴリがそれぞれ特定の一つのカテゴリが検出されるような場合にのみ限定する必要はない。即ち、ある状態カテゴリを記述するとき、接触物体カテゴリ及び動作カテゴリをそれぞれ複数個用意しておきそのうちのいずれかが検出されることにより、その状態カテゴリを特定するようにしてもよい。

【0106】

図11は、図9における習慣的行動判定装置10によって実行される習慣的行動判定処理のフローチャートである。

【0107】

図11において、状態カテゴリと時系列特徴データとの対応付けを行う(ステップS601)。状態カテゴリと時系列特徴データとの対応付けは、前述したように、学習装置5内部の辞書データ生成装置を用いて対話的に行うか、予め標準的な状態カテゴリに対応する動作パターンに関する特徴量の時系列画像データとして辞書データ作成装置の記憶手段に格納されているものとする。次いで、入力された動画像から対象となる人物を検出し、前記実施の形態と同様の被写体認識処理を人物の示す動作パターン又は姿に基づいて行う(ステップS602)。

【0108】

次に、その特定された人物の動作パターンの特徴量時系列画像データ(対象となる人物を構成するパートの代表点ごとの時系列位置情報等)を検出し(ステップS603)、状態検出装置4はその時系列画像データと辞書内部の状態カテゴリとの照合を行う(ステップS604)。ステップS604では、前述した注視物体カテゴリ又は接触物体カテゴリの抽出、及び動作カテゴリの抽出を行い、次にそれらのデータを用いた状態カテゴリの分類をニューラルネットワーク、統計的処理、又はファジー推論等の技術を用いて行う。

【0109】

次いで、時系列画像データと辞書内部の状態カテゴリの対応付けが可能であるか否かを判別し(ステップS605)、時系列画像データと辞書内部の状態カテゴリの対応付けが可能でないか、又は困難であるときは、状態カテゴリの学習を行い(ステップS606)、時系列画像データと辞書内部の状態カテゴリの対応付けが可能であるときは、さらに状

10

20

30

40

50

態カテゴリと関連付けられる場所、時間帯の少なくとも一方と共に所定の記憶装置に格納する（ステップS607）。

【0110】

ステップS606の学習は、学習装置5により時系列特徴量データと動作・状態カテゴリとの対応付けの辞書データを作成することによって行われる。学習処理はユーザヒシステムとの間で対話的に行ってもよい。

【0111】

次いで、抽出された状態カテゴリの観測頻度を算出して、例えば、週に3回以上の頻度の状態又は動作を習慣的行動とすることにより、習慣性の判定を行って（ステップS608）、本処理を終了する。

10

【0112】

図11の処理によれば、その特定された人物の動作パターンの特徴量時系列画像データを検出し（ステップS603）、時系列画像データと辞書内部の状態カテゴリの対応付けが可能でないか、又は困難であるときは、状態カテゴリの学習を行う（ステップS606）ので、状態カテゴリを学習し、その状態カテゴリの頻度を算出することにより習慣性の判定を行うことができる。

【0113】

次に、図10における注視物体カテゴリ25の検出処理の流れについて説明する。先ず、対象となる人物の視線方向、又は顔の向きの検出を行う。その検出された方向が一定時間以上ほぼ同じ方向（視野角にして±10度以内）に留まっていることが検出された場合に「注視している」という状態にあると判定される。但し、顔の向きはほぼ一定であっても、視線方向は定まっていないことが検出できる場合には「注視している」の状態には該当しないものとする。

20

【0114】

なお、顔の向きは判別できても、正面の顔画像を検出できないために目が開いているか分からぬ場合には、目が開いているものとして処理するが、「注視している」という状態に対応する動作や姿勢に関する状態は必ずしも一つに定まらないので、他の動作・姿勢を検出してもよいことはいうまでもない。

【0115】

注視対象検出装置9が行う注視物体カテゴリの抽出処理は、視線方向又は顔の向きの方向にある物体であって、注視対象となり得る物体を注視物体カテゴリとして状態認識装置4内部の記憶部に記憶し、画像データを記憶された注視物体カテゴリと照合して最大類似度を与える注視物体カテゴリを抽出することにより行う。

30

【0116】

図12は、図9における注視対象検出装置9によって実行される注視物体カテゴリ抽出処理を説明するのに用いられる図であり、（a）は観測位置が人物の上方である場合、（b）は観測位置が人物の後方である場合を示す。

【0117】

図12（a）に示すように、観測位置が人物の上方であるとき、例えば、カメラが天井に設置される等、人物の存在位置より充分に高い所にあって下を見下ろすようなときは、注視対象は、人物の顔の向き（又は視線方向）を中心とする所定範囲（例えば、30度が目安となる）の視野角内にある注視物体カテゴリに属する物体をいう。顔の向きは頭部中心と鼻の頂点（又は上から見た頭部の凸部頂点）を結ぶ方向とする。

40

【0118】

図12（b）に示すように、観測位置が人物の後方であり、かつ人物の注視対象がその人物により遮蔽されていないときは、画像中にある物体であって人物の（後ろ姿の）近傍領域にあるものから該当する物体カテゴリを抽出する。この注視対象の探索範囲は一般的には、動作認識装置の視点位置から見た人物及び注視対象となり得る物体の相対距離に依存する。ここでは、動作認識装置の視点位置から見た人物頭部のサイズから人物までの距離を推定し、その人物を中心とした所定の視野角範囲を探索範囲とする。

50

【0119】

この場合、可能であれば動作認識装置は後頭部の画像から予め行う画像ベースでの対応付け（キャリブレーション）等により顔の向きを大まかに推定し、その推定された方向を中心とする範囲であって人物位置を基準とする所定の視野角範囲を探索範囲としてもよい。

【0120】

注視物体カテゴリの特定は、複数の観測位置からの画像を用いて行ってもよい。例えば、人物の上方及び後方の観測位置から候補となる注視物体カテゴリをそれぞれ抽出し、共通するカテゴリから特定してもよい。また、上記いずれにも該当しない位置に動作認識装置の画像入力手段（撮像装置）が存在する場合、例えば、人物の横顔が観測されるような場合には、その観測位置からの注視対象の推定を行わない。

10

【0121】

図12の処理によれば、対象が注視する物体の注視物体カテゴリを抽出することができる。

【0122】

上記第3の実施の形態によれば、図12の処理により注視物体カテゴリを抽出し、図11の処理により状態カテゴリを学習し、その状態カテゴリの頻度を算出することにより習慣性の判定を行うことができる。

【0123】

上記第1～第3の実施の形態によれば、画像データ及び動作パターンの両方からそれぞれの信頼度に基づいて人物等の対象となる人物を特定することができ、対象となる人物が比較的遠方にいる場合や照明条件等が大きく変動してもその対象となる人物の特定を高い精度で行うことができる。また、動作の意味をその特定の対象となる人物と対応付けて学習することにより、機器制御のパーソナライズ化を行うことができる。更に、特定の人物の常用物、習慣的行動パターンの検出及び学習を簡便に行うことができる。

20

【0124】

また、本発明の目的は、上記実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体（又は記録媒体）を、システム又は装置に供給し、そのシステム又は装置のコンピュータ（又はCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出して実行することによっても達成されることは言うまでもない。

30

【0125】

この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施の形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。

【0126】

また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施の形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム(OS)などが実際の処理の一部又は全部を行い、その処理によって前述した実施の形態の機能が実現される場合も含まれることは言うまでもない。

40

【0127】

さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わるCPU等が実際の処理の一部又は全部を行い、その処理によって前述した実施の形態の機能が実現される場合も含まれることは言うまでもない。

【0128】

また、上記プログラムは、上述した実施の形態の機能をコンピュータで実現することができればよく、その形態は、オブジェクトコード、インタプリタにより実行されるプログラム、OSに供給されるスクリプトデータ等の形態を有するものでもよい。

50

【0129】

プログラムを供給する記録媒体としては、例えば、RAM、NV-RAM、フロッピー(登録商標)ディスク、光ディスク、光磁気ディスク、CD-ROM、MO、CD-R、CD-RW、DVD(DVD-ROM、DVD-RAM、DVD-RW、DVD+RW)、磁気テープ、不揮発性のメモリカード、他のROM等の上記プログラムを記憶できるものであればよい。又は、上記プログラムは、インターネット、商用ネットワーク、若しくはローカルエリアネットワーク等に接続される不図示の他のコンピュータやデータベース等からダウンロードすることにより供給される。

【図面の簡単な説明】

【0130】

10

【図1】本発明の第1の実施の形態に係る動作認識装置の概略構成を示す図である。

【図2】図1の動作認識装置によって実行される動作認識処理のフローチャートである。

【図3】図2のステップS102の被写体認識処理のフローチャートである。

【図4】図2のステップS103の動作パターン抽出処理のフローチャートである。

【図5】本発明の第2の実施の形態に係る動作認識装置の概略構成を示す図である。

【図6】図5における学習装置5の内部構成を示すブロック図である。

【図7】図5における常用物検出装置7によって実行される常用物検出処理のフローチャートである。

【図8】図7のステップS402の常用物候補抽出処理のフローチャートである。

【図9】本発明の第3の実施の形態に係る動作認識装置の概略構成を示す図である。

20

【図10】状態カテゴリーの辞書データ内部での表現形式を示す図であり、(a)は行為の対象となる目的物体を有しない場合、(b)は目的物体を有する場合を示す。

【図11】図9における習慣的行動判定装置10によって実行される習慣的行動判定処理のフローチャートである。

【図12】図9における注視対象検出装置9によって実行される注視物体カテゴリー抽出処理を説明するのに用いられる図であり、(a)は観測位置が人物の上方である場合、(b)は観測位置が人物の後方である場合を示す。

【符号の説明】

【0131】

30

1 画像入力装置

2 動体検出装置

3 被写体認識装置

4 状態検出装置

5 学習装置

6 処理制御装置

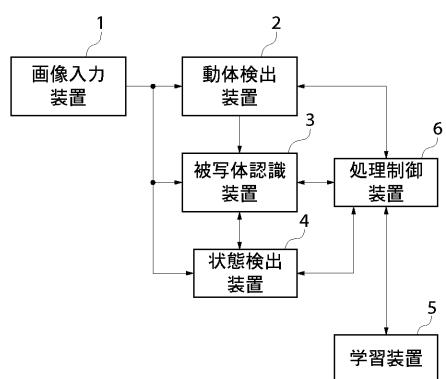
7 常用物検出装置

8 表示装置

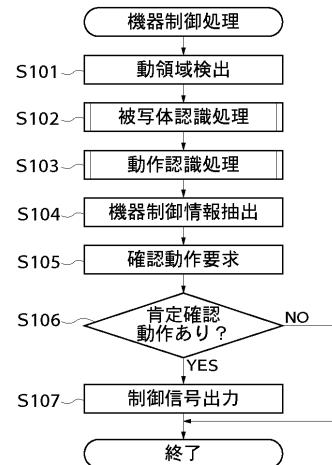
9 注視対象検出装置

10 習慣的行動判定装置

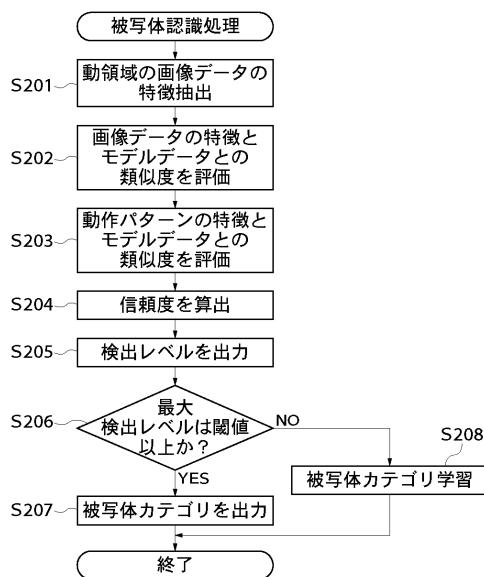
【図1】



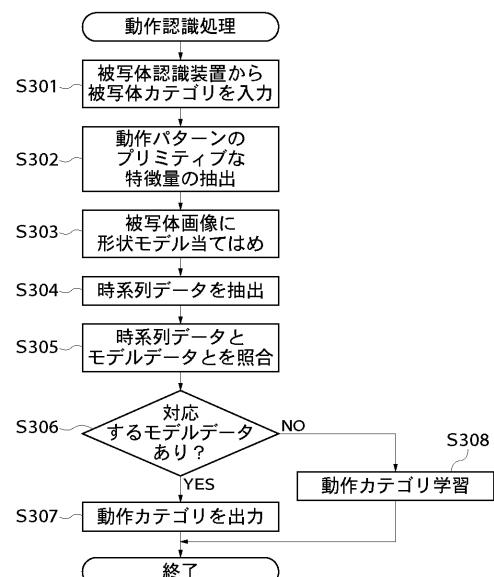
【図2】



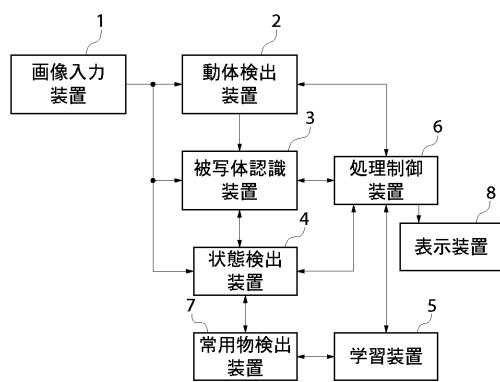
【図3】



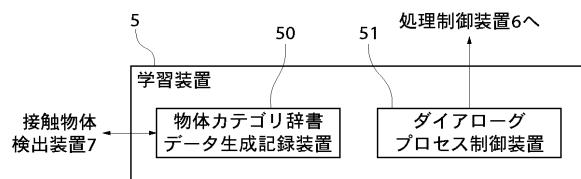
【図4】



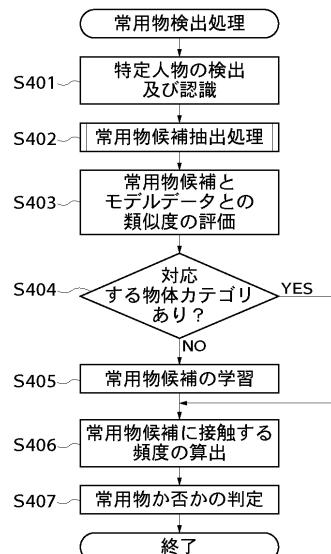
【図5】



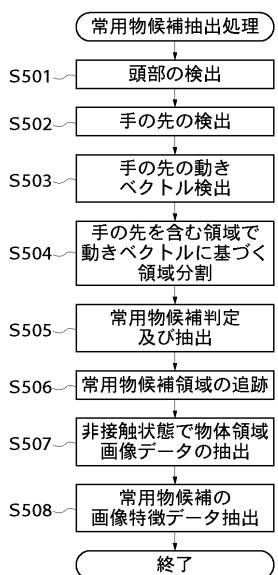
【図6】



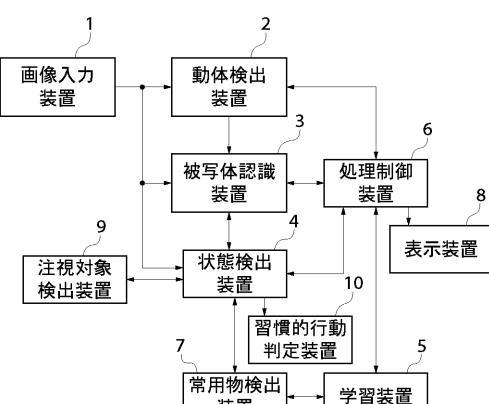
【図7】



【図8】

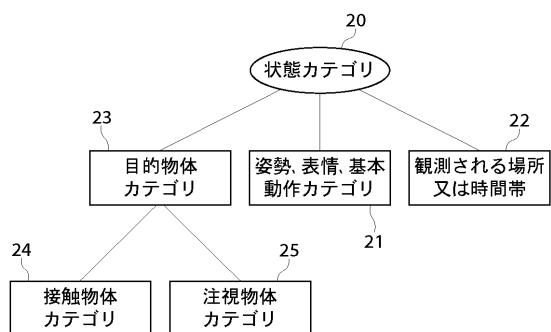


【図9】

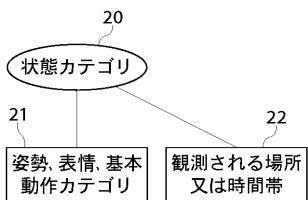


【図10】

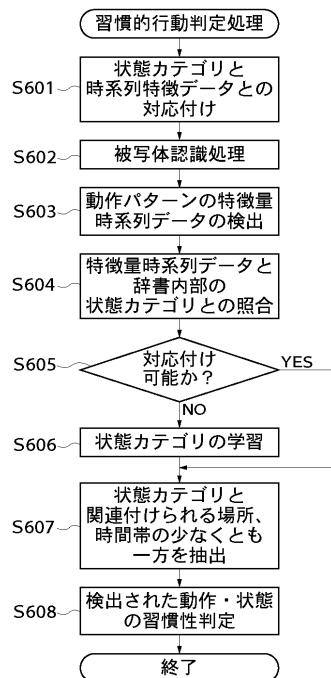
(a)



(b)

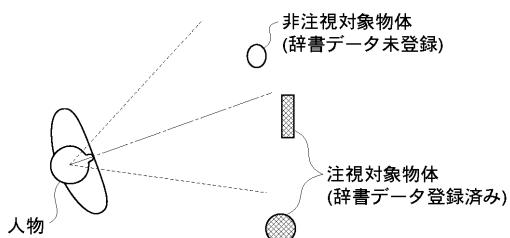


【図11】

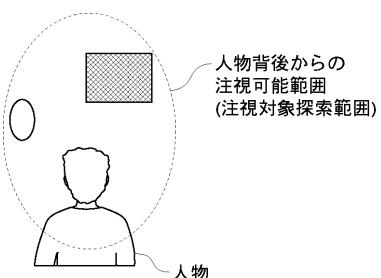


【図12】

(a)



(b)



フロントページの続き

(72)発明者 御手洗 裕輔
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 松尾 淳一

(56)参考文献 特開平08-257017(JP, A)
特開平09-081309(JP, A)
特開平09-161102(JP, A)
特開平11-249773(JP, A)
特開平10-149207(JP, A)
特開平10-162151(JP, A)
特開平11-259643(JP, A)
特開平11-338614(JP, A)
特開2000-101898(JP, A)
特開2000-222361(JP, A)
特開2000-293685(JP, A)
特開2001-034764(JP, A)
特開2001-056861(JP, A)
特開2001-202524(JP, A)
特開2002-251235(JP, A)
特開2002-259978(JP, A)
特開2003-533801(JP, A)
特許第2801362(JP, B2)
特許第3140911(JP, B2)
国際公開第03/025859(WO, A1)
米国特許第6111580(US, A)
米国特許出願公開第2003/91215(US, A1)

(58)調査した分野(Int.Cl. , DB名)

| | |
|--------|-------------------|
| G 06 F | 3 / 0 1 |
| G 06 F | 3 / 0 4 8 |
| G 06 T | 1 / 0 0 - 1 / 4 0 |
| G 06 T | 3 / 0 0 - 3 / 6 0 |
| G 06 T | 5 / 0 0 - 5 / 5 0 |
| G 06 T | 7 / 0 0 - 7 / 6 0 |
| G 06 T | 9 / 0 0 - 9 / 4 0 |