

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6207240号
(P6207240)

(45) 発行日 平成29年10月4日(2017.10.4)

(24) 登録日 平成29年9月15日(2017.9.15)

(51) Int. Cl. F 1
G 0 6 F 3 / 0 1 (2 0 0 6 . 0 1) G O 6 F 3 / 0 1 5 7 0
G 0 6 F 3 / 0 3 4 6 (2 0 1 3 . 0 1) G O 6 F 3 / 0 3 4 6 4 2 2
G 0 6 F 3 / 0 4 8 3 (2 0 1 3 . 0 1) G O 6 F 3 / 0 4 8 3

請求項の数 13 (全 25 頁)

(21) 出願番号	特願2013-118875 (P2013-118875)	(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
(22) 出願日	平成25年6月5日(2013.6.5)	(74) 代理人	100126240 弁理士 阿部 琢磨
(65) 公開番号	特開2014-235698 (P2014-235698A)	(74) 代理人	100124442 弁理士 黒岩 創吾
(43) 公開日	平成26年12月15日(2014.12.15)	(72) 発明者	新井 常一 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
審査請求日	平成28年6月2日(2016.6.2)	審査官	原 秀人

最終頁に続く

(54) 【発明の名称】 情報処理装置及びその制御方法

(57) 【特許請求の範囲】

【請求項1】

所定領域を撮像する撮像部から得られた画像に基づいて、ジェスチャ物体によるジェスチャ操作を認識する情報処理装置であって、

前記所定領域に存在する注目物体の位置を取得する取得手段と、

前記画像に基づいて、前記ジェスチャ物体の状態を特定する特定手段と、

前記取得手段によって取得された前記注目物体の位置と前記特定手段によって特定された前記ジェスチャ物体の状態とに基づいて、前記所定領域のうち一部の領域を、前記注目物体に対応する少なくとも1つのジェスチャ操作のジェスチャ認識領域として設定する設定手段と、

前記撮像部から得られた画像のうち、前記設定手段によって設定されたジェスチャ認識領域に対応する部分から得られる情報に基づいて、前記ジェスチャ物体によるジェスチャ操作を認識する認識手段と

を備えることを特徴とする情報処理装置。

【請求項2】

前記取得手段によって取得された注目物体を識別する識別手段と、

前記識別手段によって識別された前記注目物体と、前記特定手段によって特定された前記ジェスチャ物体の状態とに基づいて、前記ジェスチャ物体によって行われるジェスチャ操作を推定する推定手段を更に備え、

前記設定手段は、前記推定手段の推定結果に応じて、前記ジェスチャ認識領域を設定す

ることを特徴とする請求項 1 に記載の情報処理装置。

【請求項 3】

前記ジェスチャ物体とは、ユーザの手であって、

前記推定手段は、前記識別手段が識別した物体に対する、前記特定手段が特定したユーザの手の状態に基づいて、前記ユーザの手によって行われるジェスチャ操作を推定することを特徴とする請求項 2 に記載の情報処理装置。

【請求項 4】

前記識別手段は、少なくとも前記注目物体が紙であること、あるいは本であることを識別することを特徴とする請求項 2 又は 3 に記載の情報処理装置。

【請求項 5】

前記推定手段は、前記識別手段により前記注目物体が紙であることが識別され、かつ、前記特定手段によって、前記ユーザの手が、前記取得手段によって取得された前記注目物体を掴む姿勢であると特定された場合には、前記ユーザの手によって、前記紙を裏返すジェスチャ操作が行われると推定し、

前記設定手段は、前記ユーザの手が前記紙を裏返す際に動く可能性が高い範囲を含むように、前記ジェスチャ認識領域を再設定することを特徴とする請求項 3 に記載の情報処理装置。

【請求項 6】

前記推定手段は、前記識別手段により前記注目物体が本であることが識別され、かつ、前記特定手段によって、前記ユーザの手が、前記取得手段によって取得された前記注目物体を掴む姿勢であると特定された場合には、前記ユーザの手によって、前記本のページをめくるジェスチャ操作が行われると推定し、

前記設定手段は、前記ユーザの手が前記本のページをめくる際に動く可能性が高い範囲を含むように、前記ジェスチャ認識領域を再設定することを特徴とする請求項 3 又は 5 に記載の情報処理装置。

【請求項 7】

所定領域を撮像する撮像部から得られた画像に基づいて、ジェスチャ物体によるジェスチャ操作を認識する情報処理装置であって、

前記所定領域に存在する注目物体の位置を取得する取得手段と、

前記所定領域のうち、前記取得手段によって取得された前記注目物体の位置に応じた一部の領域を、前記注目物体に対応する少なくとも 1 つのジェスチャ操作のジェスチャ認識領域として設定する設定手段と、

前記撮像部から得られた画像のうち、前記設定手段によって設定されたジェスチャ認識領域に対応する部分から得られる情報に基づいて、前記ジェスチャ物体によるジェスチャ操作を認識する認識手段と、

を備え、

前記設定手段は、前記取得手段によって位置が取得された前記注目物体が、前記ジェスチャ物体によって前記所定領域に挿入された時に前記撮像部が撮像した画像に基づいて、前記ジェスチャ認識領域にジェスチャ操作の方向を定義するための座標系を設定することを特徴とする情報処理装置。

【請求項 8】

所定領域を撮像する撮像部から得られた画像に基づいて、ジェスチャ物体によるジェスチャ操作を認識する情報処理装置であって、

前記所定領域に存在する注目物体の位置を取得する取得手段と、

前記所定領域のうち、前記取得手段によって取得された前記注目物体の位置に応じた一部の領域を、前記注目物体に対応する少なくとも 1 つのジェスチャ操作のジェスチャ認識領域として設定する設定手段と、

前記撮像部から得られた画像のうち、前記設定手段によって設定されたジェスチャ認識領域に対応する部分から得られる情報に基づいて、前記ジェスチャ物体によるジェスチャ操作を認識する認識手段と、

10

20

30

40

50

を備え、

前記注目物体には、表示部によって前記所定の空間内に表示された仮想的な物体と、前記所定の空間に挿入された物理的な物体を含むことを特徴とする情報処理装置。

【請求項 9】

前記注目物体には、表示部によって前記所定の空間内に表示された仮想的な物体と、前記所定の空間に挿入された物理的な物体を含み、

前記識別手段により、前記取得手段によって取得された前記注目物体が、表示部によって前記所定の空間内に表示された仮想的な注目物体と識別された場合、

前記認識手段は、前記ジェスチャ物体が、前記設定手段によって前記注目物体の周囲に設定された領域内を横切る動きに応じて、前記表示部に対して、前記所定の空間内に表示する仮想的な注目物体を変更させるスワイプ操作を認識することを特徴とする請求項 2 に記載の情報処理装置。

10

【請求項 10】

所定領域を撮像する撮像部から得られた画像に基づいて、ジェスチャ物体によるジェスチャ操作を認識する情報処理装置の制御方法であって、

取得手段により、前記所定領域に存在する注目物体の位置を取得する取得工程と、

特定手段により、前記画像に基づいて、前記ジェスチャ物体の状態を特定する特定工程と、

前記取得工程で取得された前記注目物体の位置と、前記特定工程で特定された前記ジェスチャ物体の状態とに基づいて、前記所定領域のうち一部の領域を、設定手段により、前記注目物体に対応する少なくとも 1 つのジェスチャ操作のジェスチャ認識領域として設定する設定工程と、

20

認識手段により、前記撮像部から得られた画像のうち、設定された前記ジェスチャ認識領域に対応する部分から得られる情報に基づいて、前記ジェスチャ物体によるジェスチャ操作を認識する認識工程と

を備えることを特徴とする情報処理装置の制御方法。

【請求項 11】

識別手段により、前記取得手段によって取得された注目物体を識別する識別工程と、推定手段により、前記識別工程で識別された前記注目物体と、前記特定工程で特定された前記ジェスチャ物体の状態とに基づいて、前記ジェスチャ物体によって行われるジェスチャ操作を推定する推定工程を更に有し、

30

前記設定工程では、前記推定工程における推定結果に応じて、前記ジェスチャ認識領域を設定することを特徴とする請求項 10 に記載の情報処理装置の制御方法。

【請求項 12】

コンピュータに読み取らせ実行させることによって、前記コンピュータを請求項 1 に記載の情報処理装置の各手段として動作させるためのプログラム。

【請求項 13】

請求項 12 に記載のプログラムが記憶されたコンピュータが読み取り可能な記憶媒体。

【発明の詳細な説明】

【技術分野】

40

【0001】

本発明は、ユーザの動作を認識する情報処理装置に関する。

【背景技術】

【0002】

近年、可視光カメラや赤外線カメラによる撮像画像からユーザの手が写った手領域を検出し、手の動きや位置に応じたジェスチャ操作を認識する装置が広まりつつある。これらの装置では、手領域の追跡から認識されるユーザの手の動きと、予め指示内容に対応付けて登録されている様々なジェスチャパターンとの比較から、装置に対して行われたジェスチャ操作と、その指示内容を認識する。手領域の動きを検出する場合、ある指定領域内に手が入ってきた時から、手領域の追跡（ジェスチャの認識）を開始することが一般的であ

50

る。

【0003】

特許文献1では動画像から抽出された人物の顔部分の横幅、縦長さ、又は面積のうち少なくとも一つに基づいて、前記人物がジェスチャを行う際に手が存在すると推定される付近に、ジェスチャを認識するための指定領域を設定する。ジェスチャを認識するための指定領域は、ジェスチャ認識領域と呼ばれることがある。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2003-216955号公報

10

【発明の概要】

【発明が解決しようとする課題】

【0005】

ジェスチャ認識領域は、ジェスチャ操作を行う手の一連の動きが全て収まる程度に大きい必要があるが、一方で大きすぎると、次のジェスチャ操作を開始する前に手を領域から除くために大ぶりの動作が必要となってしまう。特に、仮想的あるいは物理的な物体に注目し、当該注目物体に関連したジェスチャ操作を行う場合には、ユーザが手を動かす範囲の大きさは、当該注目物体の大きさに影響されやすい。従って、初期状態を作るための大ぶりの動作が不自然な動きと感じられやすく、操作性を低減させることがあった。

【0006】

20

特許文献1では、ユーザの現在の位置において、ジェスチャを切り出す領域を、各ユーザの手が存在するであろう範囲に絞り込むことはできるが、行われるジェスチャ操作の各々に適した領域の大きさについては考慮されていなかった。

【0007】

本発明は上記の問題点に鑑みてなされたものであり、注目物体に関連したジェスチャ操作を行う場合の操作性を向上させることを主な目的とする。

【課題を解決するための手段】

【0008】

本発明は、上記課題を鑑みてなされたものであり、所定領域を撮像する撮像部から得られた画像に基づいて、ジェスチャ物体によるジェスチャ操作を認識する情報処理装置であって、前記所定領域に存在する注目物体の位置を取得する取得手段と、前記画像に基づいて、前記ジェスチャ物体の状態を特定する特定手段と、前記取得手段によって取得された前記注目物体の位置と前記特定手段によって特定された前記ジェスチャ物体の状態とに基づいて、前記所定領域のうち一部の領域を、前記注目物体に対応する少なくとも1つのジェスチャ操作のジェスチャ認識領域として設定する設定手段と、前記撮像部から得られた画像のうち、前記設定手段によって設定されたジェスチャ認識領域に対応する部分から得られる情報に基づいて、前記ジェスチャ物体によるジェスチャ操作を認識する認識手段とを備えることを特徴とする。

30

【発明の効果】

【0009】

40

本発明によれば、注目物体に関連したジェスチャ操作を行う場合の操作性を向上させることができる。

【図面の簡単な説明】

【0010】

【図1】情報処理装置のハードウェア及びソフトウェアの構成と外観の一例を表す図

【図2】情報処理装置が認識するジェスチャ操作とその関連情報の一例を表す表

【図3】情報処理装置がジェスチャ操作を認識するメイン処理の流れの一例を示すフローチャート

【図4】情報処理装置が注目物体とユーザの位置に合わせた認識領域を設定する処理の流れの一例を示すフローチャート

50

【図5】情報処理装置が手の状態を検出する処理の流れの一例を示すフローチャート

【図6】推定された操作毎に適した認識領域を設定する処理の流れの一例を示すフローチャート

【図7】「裏返し」操作の操作例を表す図

【図8】「ページめくり」操作が推定された場合の認識領域の例を表す図

【図9】「ページめくり」操作が推定された場合の認識領域の例を表す図

【図10】「ページめくり」操作の操作例を表す図

【発明を実施するための形態】

【0011】

以下、添付図面を参照して本発明に係る実施の形態を詳細に説明する。ただし、この実施の形態に記載されている構成要素はあくまでも例示であり、本発明の範囲をそれらのみ

10

に限定する趣旨のものではない。

【0012】

(第1の実施形態)

本実施形態では、ジェスチャを切り出す領域を固定にするのではなく、その都度起こり得るジェスチャ操作の各々に適した領域の大きさに設定する。

【0013】

また、注目物体に関連したジェスチャ操作を認識する際に、事前にユーザが行おうとしている操作を推定することとする。そして、注目物体に関連して次に行われる可能性が高いジェスチャ操作に適切な大きさ及び位置のジェスチャ認識領域を設定する。以下、具体

20

例を説明する。

【0014】

以下において、ジェスチャ認識領域とは、当該領域に、ジェスチャ操作に用いられる物体(例えばユーザの手)が検出され始めた時点からジェスチャが開始されたと判断するための領域である。以降では、ジェスチャ操作に用いられる物体をジェスチャ物体という。また、本実施形態では、ジェスチャ認識領域からジェスチャ物体が出たことに応じて、ジェスチャが終了したと判断する。そして、ジェスチャが開始されたと判断してから終了したと判断されるまでのジェスチャ物体の一連の動きが、予め記憶された所定の動きである場合に、ジェスチャ操作を認識する。つまり、本実施形態の情報処理装置は、ジェスチャ認識領域内でのジェスチャ物体の動きを追跡し、追跡した動きが所定の動きである場合に、当該所定の動きに関連付けられたジェスチャ操作が入力されたと認識する。ジェスチャ認識領域は、単に認識領域とも言う。なお、本実施形態において、ジェスチャ操作にはユーザの手が用いられるとするが、スタイラスやその他のツールがジェスチャ物体であってもかまわない。

30

【0015】

図1(a)は、本実施形態に係る情報処理装置100の外観の一例を示す図である。本実施形態では、ユーザがタッチすることで操作を入力するユーザインタフェース用の表示オブジェクト101(例えば、ボタンやアイコン)等を含む投影画像が、操作面102に対し上部に設置されたプロジェクタ103から投影することによって表示される。また、カメラ104も同様に上部に設置され、俯瞰撮影された映像に基づき、ユーザの動作を認識することや、背景の操作面102上に置かれた注目物体105を撮像して画像データを得ることができる。図1(a)では、注目物体の一例として本105が操作面上に載置されている。以下では、カメラ104が紙や本等の物理的な物体である注目物体の内容を画像データとして得るために行う撮像は、特に「読み取り」ということとする。また、情報処理装置100は、距離センサ105によって、ジェスチャ物体と操作面102との間の距離を検知し、操作面102の表面から所定の距離まで範囲内にジェスチャ物体が検出される場合に、操作面102がタッチされていると判断する。本実施形態において注目物体やユーザの手の位置は、操作面102に関して、図2(a)に示すように、紙面に向かって左下の頂点を原点としたx軸、y軸に基づく座標として扱う。

40

【0016】

50

本実施形態では、カメラ104が撮像した画像に基づいて、ユーザのジェスチャ操作として認識する。特に、操作面102上に載置された注目物体に対するユーザの動作を、ジェスチャ操作として認識する。例えば、本105のページをユーザがめくる手の動きを、見開きページ毎の読み取りを指示するジェスチャ操作として認識し、めくる動きが終了したことを操作終了として認識したら、そのめくられた後の見開きページの画像データを得るために読み取りする。本実施形態では、こうして読み取りされた画像データあるいはそのサムネイル画像を、プロジェクタ103により操作面102に投影することができる。そして本実施形態では、複数の画像データのうちの1つの画像データが操作面102に表示されている場合、表示されている画像データの上で手を左右いずれかの方向に横切らせる動きが、画像送りを指示するジェスチャ操作として認識される。操作が認識されたこと
10
に応じて、表示される画像データが順次変更される。以下では、このように物理的な物体（本等）仮想的な物体（表示された画像データ等）に対するジェスチャ操作を、注目物体に関連したジェスチャ操作という。なお、本実施形態において「本」とは、複数の紙が複数のページを構成し、閉じられたページを繰って見ていくことによって内容を閲覧することができる形状の物（書籍等）を差す。

【0017】

図1(b)は、本実施形態における情報処理装置100のハードウェア構成図である。同図において、CPU(中央処理装置)111は、情報処理装置の制御プログラムを実行して各種処理のための演算や論理判断等を行い、システムバス115に接続された各構成要素を制御する。ROM112は、プログラムメモリであって、後述するフローチャート
20
に示す各種処理手順を含むCPUによる制御のためのプログラムを格納する。RAM113は、CPU111のワークエリア、エラー処理時のデータの退避領域、上記制御プログラムのロード領域等を提供するために用いられる。記憶装置114は本実施形態に係るデータやプログラムを記憶しておくためのハードディスクや接続された外部記憶装置等である。表示装置103は、表示オブジェクトや画像等の電子データを表示するためのディスプレイ、プロジェクタ等である。本実施形態では、表示装置として液晶プロジェクタを用いる。カメラ104は、本実施形態では操作エリアを上から撮像するように設置され、入力画像を撮像する撮像部である。距離センサ105は、例えば赤外線センサで、ユーザが操作のために用いるジェスチャ物体（例えば、ユーザの手）と、操作面との距離に基づき、ユーザが操作面をタッチしたことを検知する。
30

【0018】

なお、本実施形態の情報処理装置100には、表示装置103及びカメラ104が内蔵されたが、外部に接続された表示装置103及びカメラ104を制御する形態であっても構わない。

【0019】

図1(c)は、第1の実施形態に係る情報処理装置100の機能ブロック図である。例えば、上記CPU111を用いたソフトウェア処理の代替としてハードウェアを構成する場合には、ここで説明する機能(処理部)に対応させて演算部や回路を構成すれば良い。

【0020】

撮像部120は、カメラ104及びCPU111等から構成され、操作面102を上から撮像して、所定領域(撮像領域)を撮像した画像を取得する。また、操作面102に載置された紙や本等の物理的な物体である注目物体を読み取る。撮像部120に撮像される空間(カメラ104のレンズから操作面102までの空間)を、以下では作業空間という。作業空間を、カメラ104の撮像素子に平行な二次元方向に見た領域を、撮像領域という。撮像部120は、作業空間を周期的に撮像し続けている。撮像された複数の画像を時系列に従って比較することで、被写体の動きが検出できる。本実施形態では、読み取りのための撮像は、周期的な撮像とは別に行われる。ただし、周期的に繰り返されている撮像のうち1回で得た画像を読み取り用の撮像画像とすることももちろんできる。また、複数回の撮像から得られた複数の画像に基づいて読み取り用の撮像画像を取得してもよい。カメラ104の撮像可能な範囲を読み取りエリアとするが、ここでは、操作面のサイズと
40
50

、読み取りエリアのサイズが等しいものとする。

【0021】

検出部121は、CPU111等から構成され、ROM112に記憶されたプログラムやデータをRAM113に読み出して実行し、撮像部120が撮像した画像に基づいて、作業空間に存在する検出対象を検出する。本実施形態での検出対象は、ユーザの手、注目物体読み取り処理の対象である物理的な物体、および表示されている仮想的な物体である。本実施形態では、検出部121は、撮像部120が撮像した画像と、予め初期状態（検出対象が存在しない状態の作業空間）が撮像された初期画像とを利用した背景差分法によって、検出対象の検出を行う。初期画像を背景とし、撮像部120が撮像した入力画像との差分を取ることによって、入力画像のうち、作業空間に出現した移動体が写る領域を検出することができる。さらに、時系列で撮像された入力画像同士の差分を取ることによって、検出された領域に写る移動体の少なくとも一部が静止したことを検出することができる。本実施形態においては、初期画像には含まれず、かつ、作業空間に出現した後も静止しない移動体が写る領域に対し、予め記憶されている人間の手の形状モデルとのマッチング処理を行う。マッチングの尤度が所定値よりも高い場合に、当該領域はユーザの手が写っている領域として検出される。一方、初期画像には含まれず、かつ、作業空間に出現した後に静止した移動体が写る領域は、注目物体読み取り処理の対象である物理的な物体、および表示されている仮想的な物体が写る領域として検出される。以上のような検出部121による処理を、以下では、「ユーザの手を検出する」あるいは「注目物体を検出する」と省略して表記する。なお、検出部121、ジェスチャ認識領域が所定の撮像領域のうちのどこに設定されているかに関わらず、ユーザの手を検出し続ける。

10

20

【0022】

識別部122は、CPU111等から構成され、ROM112に記憶されたプログラムやデータをRAM113に読み出して実行し、検出部121によって検出された注目物体の種類を識別する。具体的には、撮像画像のうち注目物体が写る領域の画像の情報と、距離センサ105が検出する注目物体の厚みの情報、予め記憶されたサンプルモデルとのマッチングにより、注目物体が紙、本、あるいは投影画像のうち何れかを識別する。本実施形態において、紙は、閉じられていない1枚の紙であって、A4、B5等規定サイズの用紙や、レシートやメモ用紙のような規定サイズ外の印刷物を含む。注目物体が紙であることの識別は、注目物体が厚みを持たないことや、画像の特徴分析により背表紙が存在しないことやステープル等による閉じ込みが無いことを判別して行う。また、注目物体が本であることの識別には、注目物体が所定以上の厚みを有することや、画像の特徴分析により背表紙が存在することを認識する等の方法を用いる。検出部121は、識別した注目物体を示す位置情報をRAM113に保持する。

30

【0023】

特定部123は、CPU111等から構成され、ROM112に記憶されたプログラムやデータをRAM113に読み出して実行する。特定部123は、検出部121によって検出されたユーザの手の、検出部121によって検出された注目物体に対する位置と姿勢を特定する。ただし、特定部123の処理は、検出部121が注目物体を検出したことに応じて、すなわち作業空間に挿入された移動体の一部が静止したことが検出されたことに応じて開始される。特定部123は、まず撮像部120による撮像画像において検出部121が検出したユーザの手が挿入されている方向に基づいて、ユーザの向きを特定し、それに基づいて方向を定義する。そして、定義された方向、またはユーザの手の指先等特定の特定部分が存在する位置（例えば、注目物体の上、輪郭上）等より、ユーザの手が、検出部121が検出した注目物体に対する位置を特定するとともに、どのような手の状態であるかを特定する。ユーザの手の状態とは、例えば注目物体を掴んでいる、押さえている、といった姿勢に対応する手のモデルを用い、検出部121が、移動体が人の手であることを検出する処理と同様のマッチング処理により、現在の状態が特定される。

40

【0024】

推定部124は、CPU111等から構成され、ROM112に記憶されたプログラム

50

やデータをRAM 113に読み出して実行し、ユーザが行おうとしている可能性が高い操作を推定する。その際、記憶部127に記憶されている操作辞書に基づき、識別部122が識別した注目物体の種類と、特定部123が特定した情報に対応付けられた操作を特定し、推定結果とする。そして、操作辞書を参照することで、推定された操作に対応づけられた領域情報を、次に行われる可能性が高いジェスチャ操作に適したジェスチャ領域であるとして、領域を示す情報を特定する。

【0025】

設定部125は、CPU 111等から構成され、ROM 112に記憶されたプログラムやデータをRAM 113に読み出して実行し、ジェスチャ操作を認識するためのジェスチャ認識領域を設定する。本実施形態では、情報処理装置の電源投入直後の初期状態においては、操作面102の大きさ全体を含む領域がジェスチャ認識領域として設定される。そして、識別部122が、ユーザの手以外の注目物体を識別したことに応じて、当該識別された注目物体の周囲に新たなジェスチャ認識領域が設定される。ここで設定される所定の認識領域は、注目物体の大きさと位置に応じた領域であって、具体的には注目物体の外接矩形を包含するように設定される。また、ユーザが行うジェスチャの方向を定義するための座標系は、ユーザが当該注目物体を作業空間に挿入した時の手(腕)の方向に基づいて定義される。ただし、仮想的な物体の場合には、投影方向に基づいて定義される等変形を加えてもかまわない。さらに、設定部125は、推定部124に推定されたジェスチャ操作に適したジェスチャ領域を示す情報に基づいて、撮像部120の所定の撮像領域のうちの少なくとも一部を、ジェスチャ認識領域と設定する。

【0026】

認識部126は、CPU 111等から構成され、ROM 112に記憶されたプログラムやデータをRAM 113に読み出して実行し、設定部125に設定されたジェスチャ認識領域内でのユーザの手の動きを追跡し、操作として認識する。すなわち、追跡した手の動きに対応する操作コマンドを特定する。具体的には、検出部121が検出するユーザの手が、ジェスチャ認識領域内において検出され始めた時点をジェスチャの開始時、ジェスチャ認識領域から出た時点をジェスチャの終了時、として一連の動きを追跡する。そして、予め記憶部127に記憶されているジェスチャモデルと、追跡したユーザの手の動きとのマッチング処理を行い、マッチング結果が条件を満たすモデルに対応付けられた操作コマンドを特定する。操作コマンドは、例えば、撮像部120に対する読み取りの実行の指示や、表示制御部128に対する表示内容の変更の指示等、様々なものがある。ただし、本実施形態において「ジェスチャ認識領域内において検出され始めた」とは、認識領域が設定された時点で既に領域内にユーザの手が存在した場合も含む。また、ユーザの手の動きを追跡する際、本実施形態では、撮像部120による撮像画像においてユーザの掌が写っている領域の重心位置を求め、その変化を、ユーザの手の位置の移動と認識する。これは、重心位置はユーザの掌が写る領域の変形に対して、比較的安定するからである。ただし、重心に限らず、ユーザの掌が写る領域全体、あるいは中心点、その他の特徴点を追跡しても構わない。

【0027】

また、本実施形態の認識部126は、ユーザが操作面をタッチして行うタッチ操作を認識する。認識部126は、検出部121が検出したユーザの手のうち、指先の位置情報と、距離センサ105の出力に基づいて、ユーザがタッチしている位置を認識する。操作面に対するタッチ操作を認識する場合には、操作面をタッチされてからリリースされるまで、撮像部120による撮像画像においてユーザの掌が写っている領域のうち、ユーザが伸ばしている指の指先の位置を追跡し、タッチ操作の動きとして認識する。

【0028】

記憶部127は、記憶装置114に相当し、本実施形態の処理で用いられる操作辞書やジェスチャモデルと操作コマンドが対応付けられた情報等が記憶されている。ジェスチャモデルは、ジェスチャ認識領域内でユーザの手が検出され始めてから検出されなくなるまでの位置の変化と、形状の変化の基準となるモデルの情報である。

【 0 0 2 9 】

表示制御部 1 2 8 は、CPU 1 1 1 等から構成され、ROM 1 1 2 に記憶されたプログラムやデータを RAM 1 1 3 に読み出して実行し、認識部 1 2 6 が認識した操作コマンドの指示に従って、表示部 1 2 9 に出力する表示画像を生成する。

【 0 0 3 0 】

表示部 1 2 9 は、表示装置 1 0 3 及び CPU 1 1 1 等から構成され、表示制御部 1 2 8 が生成した表示画像を、操作面上に表示する。本実施形態では、表示装置 1 0 3 はプロジェクタであるため、操作面に向けた投影により、表示を行う。

【 0 0 3 1 】

図 2 (a) は、本実施形態で認識可能な操作の内容の一例を挙げた一覧表である。ここに含まれる操作は、ジェスチャ操作に限らず、タッチ操作やジェスチャ操作の準備段階の操作も含まれている。項目 2 0 4 に「 」が示されている操作は、本実施形態でいう「ジェスチャ操作」であり、ジェスチャ認識領域内でのユーザの手の検出開始から検出終了までの一連の動きに基づいて、操作コマンドが特定される操作である。ジェスチャ操作に関して、手の動き 2 0 1 を示すモデルと、指示内容 2 0 2 を示すコマンドとの対応付け情報は、記憶部 1 2 7 に記憶されており、認識部 1 2 6 は、当該対応付け情報を参照することで、これらの操作を認識する。

10

【 0 0 3 2 】

例えば本実施形態では、上述の通り、注目物体として本が操作面に載置されたときは、その本のページをめくる動きをジェスチャ操作として認識し、めくられる度に、めくった後の見開きページを読み取る。以下この明細書では、このジェスチャ操作の名称を「ページめくり」と記載する。

20

【 0 0 3 3 】

また、本実施形態では、ユーザが手で紙や本、もしくは投影された画像データ等の注目物体を、移動させたり回転させたりする動きを操作として認識し、移動又は回転後の注目物体の位置と向きに合わせて認識領域をリセットさせる。これは、注目物体の位置が変化した場合には、それらの注目物体に関連したジェスチャ操作が行われる位置も変化するからである。本実施形態では、これらの操作をそれぞれ、「紙移動」「本移動」「(画像データの)ドラッグ」という。これらは次に行われるジェスチャの準備段階の操作といえる。

30

【 0 0 3 4 】

また、本実施形態では、ユーザの手が紙を裏返す動きをジェスチャ操作として認識し、裏返し完了したことに応じて、すなわち裏返した紙が静止したことに応じて、読み取りを実行する。本実施形態では、このジェスチャ操作を「裏返し」という。

【 0 0 3 5 】

また、本実施形態では、ユーザの手が、投影された画像データの上空付近で右から左(左方向)、あるいは左から右(右方向)に横切る動きをジェスチャ操作として認識し、表示されている画像データを変更させる。本実施形態では、ユーザの手が左方向に動かされたときには、記憶されている複数の画像データのうち、所定の表示順において表示されている画像データの後に当たる画像データを表示させる。手の動きが右方向であった場合には、所定の表示順において表示されている画像データの前に当たる画像データを表示させる。本実施形態では、このジェスチャ操作をそれぞれ「左スワイプ」「右スワイプ」という。

40

【 0 0 3 6 】

また、本実施形態では、投影された画像データに対する、2箇所以上のタッチによるタッチ操作(マルチタッチ操作)を、画像データの拡大・縮小、あるいは回転等画像を変形させる指示として認識する。そして、変形された後の画像データの大きさ及び位置、向きに合わせて認識領域をリセットさせる。これは、画像データが変形した場合には、当該画像データに関連したジェスチャ操作が行われる位置も変化するからである。本実施形態では、この操作をそれぞれ、「画像変形」という。これも、次に行われるジェスチャの準備

50

段階の操作といえる。

【 0 0 3 7 】

図 2 (a) で挙げた操作は一例にすぎず、本実施形態の情報処理装置 1 0 0 では、記憶部に 1 2 7 に操作辞書として記憶させることによって、他にもユーザの手の様々な動きをジェスチャ操作として認識することができる。

【 0 0 3 8 】

図 2 (b) は、推定される操作と、推定結果に応じて設定されるジェスチャ認識領域の対応関係の一例を挙げた一覧表である。本実施形態の推定部 1 2 4 は、識別された注目物体 2 1 0 と注目物体に対する手の位置 2 1 1 と手の姿勢 2 1 2 とに基づいて、対応する推定される操作の名称 2 1 3 と、認識領域 2 1 4 を特定する。

10

【 0 0 3 9 】

例えば、推定部 1 2 4 は、識別された注目物体 2 1 0 は本であって、注目物体 (本) に対する手の位置 2 1 1 が、注目物体の輪郭付近であり、手の姿勢 2 1 2 は、「注目物体を掴む」姿勢であれば、「ページめくり」が行われると推定する。なお、「注目物体を掴む」姿勢は、本実施形態では、手の一部が注目物体に隠れていて撮像画像に写っていないことによって判定する。このような状況において「ページめくり」が行われると推定するのは、ページをめくろうとするユーザは、ページを掴むために、注目物体 (本) の輪郭付近に手を移動させ、指等手の一部がページの下に隠れる可能性が高いからである。そして、「ページめくり」の操作が行われると推定された時には、設定部 1 2 5 は、めくり終わる手の動きまで検出できるように、認識領域 2 1 4 として、その本の見開きページとページを掴む手が収まる大きさの認識領域を設定する。その他、ユーザが各ジェスチャ操作を行おうとする際に手を移動させる可能性が高い位置と、その時にとる可能性が高い手の姿勢とから、各情報が図 2 (b) のように対応づけられ、記憶される。

20

【 0 0 4 0 】

図 3 は、本実施形態の情報処理装置 1 0 0 がジェスチャ操作を認識するメイン処理の流れの一例を示すフローチャートである。本実施形態では、情報処理装置 1 0 0 の電源が投入されたことに応じて図 3 のフローチャートが起動する。ただし、特定のアプリケーションが実行されたり、モードが選択されたりしたことに応じて開始される処理とすることもできる。

【 0 0 4 1 】

ステップ S 3 0 1 において、設定部 1 2 5 は、撮像部 1 2 0 の所定の撮像領域のうち、操作面 1 0 2 の大きさ全体を含む領域をジェスチャ認識領域として設定する。このとき、認識領域の座標系は、図 1 (a) に示した x y 軸に基づいて初期設定される。本実施形態では、y 軸の正方向がユーザにとっての前方向、y 軸の負方向がユーザにとっての後方方向、x 軸の正方向がユーザにとっての右方向、y 軸の負方向がユーザにとっての左方向が初期設定となる。

30

【 0 0 4 2 】

ステップ S 3 0 2 において、撮像部 1 2 0 は、作業空間の撮像を開始し、検出部 1 2 1 は、撮像部 1 2 0 の撮像画像に基づいて、ユーザの手と注目物体の検出を開始する。本実施形態では、検出部 1 2 1 が、撮像部 1 2 0 の撮像画像に基づいて、作業空間に存在する動体を検出し、その形状が予め記憶された人の手のモデルと類似する場合に、当該動体をユーザの手として検出する。従って、情報処理装置 1 0 0 はこの時点で行われる注目物体に関連しないジェスチャ操作も認識することができる。

40

【 0 0 4 3 】

ステップ S 3 0 3 では、検出部 1 2 1 が、注目物体が作業空間に挿入されたか否かを判定する。本実施形態では、作業空間に存在する動体の形状が、予め記憶された人の手のモデルと類似しない場合、ユーザによって物理的な物体が注目物体として作業空間に挿入されたと判定する。あるいは、ユーザ操作によって、操作面上に画像データを投影させる指示が入力された場合にも、画像データの投影の開始にともなって、仮想的な物体が注目物体として作業空間に挿入されたと判定する。注目物体が作業空間に挿入されたと判断され

50

た場合（ステップS303でYES）には、ステップS304に進む。注目物体が作業空間に挿入されたと判断された場合（ステップS303でYES）には、注目物体が挿入されるまで待機する。

【0044】

ステップS304では、検出部121が、注目物体が作業空間に挿入された時の撮像画像を保持する。本実施形態では、検出部121が、注目物体が作業空間に挿入されたことを判定した際の処理に用いた撮像部120の撮像画像をRAM113に保持する。例えば、ユーザの手が注目物体を操作面上に置いた時の撮像画像や、画像データを投影させる指示を入力した時の撮像画像である。

【0045】

ステップS305では、識別部122が、挿入された注目物体を識別する。本実施形態では、注目物体の挿入が画像データを投影させる指示に基づいて判定された場合には、注目物体は投影画像であると識別される。一方、注目物体が挿入された場合には、撮像部120が撮像している撮像画像に基づいて、画像の特徴解析等の方法により識別を行う。例えば、操作面上に載置された注目物体が写る領域の画像の情報と、予め記憶されたサンプルモデルとのマッチングにより、注目物体が本であるか、紙であるのか、あるいは投影画像であるのかを識別する。ここで、注目物体が本の場合には、開かれているか、閉じられているかも含めて識別できるものとする。

【0046】

ステップS306では、設定部125が、識別された注目物体と、ユーザの位置に合わせた認識領域の設定を行う。

【0047】

ここで、図4は、ステップS306において情報処理装置が実行する、注目物体とユーザの位置に合わせた認識領域を設定する処理の流れの一例を示すフローチャートである。

【0048】

ステップS401では、検出部121が、撮像部120の撮像画像に基づいて、作業空間に挿入された注目物体の大きさと位置を示す情報を取得する。本実施形態では、撮像部120が撮像している最新の撮像画像において、注目物体が撮像されている領域を検出し、特徴点から、外接矩形を示す位置情報を取得することで、注目物体に大きさと位置を示す情報を得る。本実施形態では、外接矩形の4つの頂点の座標を取得する。ただし、少なくとも1つの頂点の座標と、外接矩形の幅等を示す情報を取得することもできる。

【0049】

ステップS402では、設定部125が、外接矩形を包含する矩形領域を認識領域として設定する。本実施形態では、ユーザの手の動きを追跡する際に、掌部分が撮像された領域の重心を追跡する。注目物体を掴む場合等には、掌部分が撮像された領域の重心は、注目物体より外側に存在する。従って、ユーザの手が注目物体を掴む場合、掌の重心位置が認識領域に含まれるように、外接矩形に対して所定の余裕分を加えた矩形領域を、新たな認識領域として設定する。

【0050】

ステップS403において、検出部121が、ステップS304においてRAM113に保持した、注目物体が作業空間に挿入された時の撮像画像を読み込む。

【0051】

ステップS404において、検出部121が、読み込んだ画像に基づいて、作業空間に対するユーザの腕の挿入方向を特定する。例えば、画像からユーザの腕と掌部分を切り出し、腕と操作面の輪郭との交差部分を特定する。特定された交差部分からユーザの掌の重心（あるいは中心でもよい）に向かうベクトルの向きを、ユーザの腕の挿入方向とする。本実施形態では、ステップS304においてRAM113に保持した撮像画像は、ステップS404の処理が終了した時点で破棄するものとする。

【0052】

ステップS405において、設定部125は、特定された挿入方向に基づいてユーザの

10

20

30

40

50

向きを推定し、認識領域の座標系を設定する。本実施形態では、操作面の中心点を通り、ステップS404において特定された方向に沿ったベクトルの向きが、ユーザの視線方向に一致するとみなし、当該ベクトルの始点が、ユーザが存在する位置であるとみなす。従って、ユーザの存在する位置から操作面の中心点に向かう方向を、ユーザにとっての前方向とし、それに準じてユーザの前後方向と左右方向を定める座標系を、ステップS403において設定された認識領域に対して新たに設定する。このように設定されたジェスチャ座標系により、ジェスチャを認識する際にジェスチャ操作の方向が特定される。なお、ここで新たに設定された座標系に平行となるように、認識領域を調整してもよい。図4のフローチャートでは、認識領域(の位置と大きさ)を設定する処理の後に座標系を設定する処理を行ったが、これらの順序は逆であっても構わない。

10

【0053】

このように、本実施形態では、注目物体が作業空間に挿入されたことに応じて、注目物体の外接矩形に合わせた大きさの認識領域を、ユーザの位置に合わせた座標系とともに設定する。注目物体に合わせた大きさの認識領域が設定されることにより、ユーザは大ぶりの動作をせずに、注目物体に関連したジェスチャを行い易くなる。また、ユーザに合わせた座標系を設定することにより、情報処理装置100は、ユーザの手の動き(例えば「右から左」に動かしたのか、「左から右」に動かしたのか)を精度よく区別して認識することができる。特に、画像データに対する左右スワイプ等、注目物体の形状を変更することなく行われる処理に対しては、ステップS306において設定される認識領域が有効的に利用される。左右のスワイプは、設定されたジェスチャ領域を、ユーザの手が右から左、あるいは左から右に横切ったことに応じて認識される。

20

【0054】

認識領域と座標系が設定されたら、処理は図3のメイン処理にリターンする。

【0055】

次に、ステップS307では、識別部122が、作業空間に挿入された注目物体が、ステップS305において識別された注目物体が、紙かあるいは本が開かれた状態(見開き)であったかを判定する。注目物体が、ステップS305において識別された注目物体が、紙かあるいは本の見開きであると判定された場合(ステップS307でYES)、ステップS308に進む。注目物体が、ステップS305において識別された注目物体が、紙かあるいは本の見開きではないと判定された場合(ステップS307でNO)、ステップS309に進む。

30

【0056】

ステップS308では、撮像部120が、注目物体の読み取りを実行し、紙や本等、の内容の画像データを取得する。

【0057】

ステップS307及びステップS308の処理は、本実施形態において、紙、あるいは本の見開きページを操作面に載置することが、当該注目物体の読み取りを行う指示であるという設定に基づく。これにより、紙を裏返したり、本のページをめくったりすることによって連続して複数回の読み取りを行う場合に、ユーザが明示的な操作を入力する手間を省いて速やかな処理を実現できる。ただし、ステップS307及びステップS308の処理を省略し、読み取りはユーザが読み取りを指示する操作を行った場合にのみ実行されても構わない。

40

【0058】

ステップS309では、特定部123が、注目物体に対するユーザの手の状態を特定する。特定部123は、ユーザの手が、図2(b)の表の項目212に示されたユーザの手の状態のいずれに該当するか、あるいは該当しないかを特定する。

【0059】

ここで、図5は、ステップS309において情報処理装置が実行する、ユーザの手の状態を検出する処理の流れの一例を示すフローチャートである。

【0060】

50

ステップS501において、検出部121が、撮像部120が撮像している撮像画像に基づいて、ユーザの手の位置を検出する。本実施形態では、撮像画像においてユーザの掌が写っている領域の重心位置を求め、ユーザの手の位置とする。なお、このときユーザの手が所定時間以上継続して静止するのを待機する処理ステップを設け、ユーザの手が静止した位置を検出するようにすると、より精度良く、ユーザの手の状態を特定することができる。

【0061】

ステップS502では、特定部123が、撮像部120が撮像している撮像画像に基づいて、検出されたユーザの手の位置が、注目物体の上であるか否かを判定する。検出されたユーザの手の位置が注目物体の上であると判定された場合（ステップS502でYES）、ステップS503に進む。検出されたユーザの手の位置が、注目物体の上ではないと判定された場合（ステップS502でNO）、ステップS509に進む。

10

【0062】

ステップS503では、特定部123が、撮像部120が撮像している撮像画像に基づいて、ユーザの手の姿勢が、掌を広げた状態か否かを判定する。本実施形態では、掌が写った領域の形状の特徴と、予め記憶された人が掌を広げている状態のモデルとのマッチング処理によって判定する。他のステップでも、ユーザの手の姿勢を判断する処理は同様に行われる。ユーザの手の姿勢が、掌を広げた状態であると判定された場合（ステップS503でYES）、ステップS504に進む。ユーザの手の姿勢が、掌を広げた状態ではないと判定された場合（ステップS503でNO）、ステップS505に進む。

20

【0063】

ステップS504では、特定部123が、ユーザの手の状態は「上から押さえる」姿勢であると特定し、メイン処理にリターンする。ステップS503及びステップS504の処理は、本実施形態において、紙を移動させる場合には、手を広げた状態で上から押さえて移動させると設定していることに基づく。これにより、他の操作との区別を精度よく行うことができる。ただし、掌を広げた姿勢に限らず、他の姿勢で移動させる設定としても構わない。

【0064】

一方、ステップS505では、特定部123が、撮像部120が撮像している撮像画像に基づいて、ユーザの手の姿勢が、一部の指だけを伸ばした状態か否かを判定する。ユーザの手の姿勢が、一部の指だけを伸ばした状態であると判定された場合（ステップS505でYES）、ステップS506に進む。ユーザの手の姿勢が、一部の指だけを伸ばした状態ではないと判定された場合（ステップS505でNO）は、ユーザの手の状態を特定せずにメイン処理にリターンする。

30

【0065】

ステップS506では、特定部123が、撮像部120が撮像している撮像画像に基づいて、伸ばされている指が1本か否かを判定する。なお、伸ばされている指が1本ではないとは、ユーザが片手のうちの複数の指（例えば、親指と人差し指）を伸ばしている場合と、両手でそれぞれ1本以上の指を伸ばしている場合を含む。伸ばされている指が1本であると判定された場合（ステップS506でYES）、ステップS507に進む。伸ばされている指が1本ではない、すなわち2本以上であると判定された場合（ステップS506でNO）、ステップS508に進む。

40

【0066】

ステップS507では、特定部123が、ユーザの手の状態は「操作面上をタッチ」する姿勢であると特定し、メイン処理にリターンする。

【0067】

ステップS508では、特定部123が、ユーザの手の状態は「操作面上を2箇所以上タッチ」する姿勢であると特定し、メイン処理にリターンする。

【0068】

ステップS509では、特定部123が、撮像部120が撮像している撮像画像に基づ

50

いて、検出されたユーザの手の位置が、注目物体の輪郭付近であるか否かを判定する。検出されたユーザの手の位置が注目物体の輪郭付近であると判定された場合（ステップS509でYES）、ステップS510に進む。検出されたユーザの手の位置が、注目物体の輪郭付近ではないと判定された場合（ステップS509でNO）、ユーザの手の状態を特定せずにメイン処理にリターンする。

【0069】

ステップS510では、特定部123が、撮像部120が撮像している撮像画像に基づいて、ユーザの手の一部が注目物体に隠れているか否かを判定する。本実施形態では、掌が写った領域の形状の特徴と、予め記憶された人の手が注目物体を掴むために掌の一部が注目物体の下あるいは側面に隠れた状態のモデルとのマッチング処理によって判定する。ただし、特定の指の位置関係が所定の条件を満たす等、別の条件によってユーザの手が注目物体を掴んでいる姿勢であることを判定することもできる。ユーザの手の一部が注目物体に隠れていると判定された場合（ステップS510でYES）、ステップS511に進む。ユーザの手の姿勢が、一部の指だけを伸ばした状態ではないと判定された場合（ステップS510でNO）は、ユーザの手の状態を特定せずにメイン処理にリターンする。

10

【0070】

ステップS511では、特定部123が、一部が注目物体に隠れていると判定されたユーザの手の位置が、注目物体を挟む2箇所が存在するか否かを判定する。ユーザの手の位置が、注目物体を挟む2箇所が存在すると判定された場合（ステップS511でYES）、ステップS512に進む。ユーザの手の姿勢が、ユーザの手の位置が、注目物体を挟む2箇所が存在しないと判定された場合（ステップS511でNO）は、ステップS513に進む。

20

【0071】

ステップS512では、特定部123が、ユーザの手の状態は「注目物体を持ち上げる」姿勢であると特定し、メイン処理にリターンする。これは、本実施形態において、本を移動させる場合には、ページをめくる場合と区別する為に、両手で持ち上げて移動させると設定していることに基づく。ただし、現実的には、例えば薄い本であれば、紙と同様に、注目物体の上を押さえることによって移動させることが可能である。ユーザが上から押さえて移動させる方法を選んだ場合には、ステップS502、ステップS503、ステップS504の処理によってユーザの手は「上から押さえる」状態と特定されることとなる。

30

【0072】

ステップS513では、特定部123が、ユーザの手の状態は「注目物体を掴む」姿勢であると特定し、メイン処理にリターンする。

【0073】

図3のメイン処理のフローチャートに戻り、ステップS310では、推定部124が、次に行われようとしている操作を推定する。推定部124は、記憶部127に記憶された操作辞書を参照し、ステップS309の処理において特定されたユーザの手の状態と、ステップS305において識別された注目物体とに関連付けられた操作を特定する。操作辞書には、例えば図2(b)のような関連付け情報が記憶されている。推定部124は、ステップS309の処理において特定されたユーザの手の状態と、ステップS305において識別された注目物体との組み合わせに対応する操作が記憶されていない場合、特定の操作が推定不可能であるという情報を推定結果とする。同様に、推定部124は、ステップS309においてユーザの手の状態が特定されていない場合にも、特定の操作が推定不可能であるという情報を推定結果とする。

40

【0074】

ステップS311では、推定部124が、特定の操作を推定できたか否かを判定する。ステップS310の処理の結果として、操作辞書に記憶されたジェスチャ操作が特定されている場合には、特定のジェスチャ操作を推定できたと判定し（ステップS311でYES）、ステップS312に進む。ステップS310の処理の結果が、特定のジェスチャ操

50

作が推定不可能であるという情報であった場合には、特定のジェスチャ操作を推定できないと判定し（ステップS311でNO）として、ステップS313に進む。

【0075】

ステップS312では、設定部125が、推定された操作に適した処理で認識領域を再設定する。推定された操作毎に行われる処理の詳細は後述する。

【0076】

ステップS313では、認識部126が、設定されているジェスチャ認識領域内部でのユーザのジェスチャ操作を認識し、各機能部に通知することにより、情報処理装置100がジェスチャ操作に応答する。例えば、撮像部120が読み取りを実行したり、表示制御部128が操作面に投影させる画像の内容を変更し、表示部129に出力させたりする。ここで所定時間の間ジェスチャ操作が行われなければ、そのままステップS314に進む。

10

【0077】

ステップS314では、認識部126が、終了指示が入力されたか否かを判定する。終了指示とは、情報処理装置100の電源を切る操作や、GUI操作によるアプリケーションの終了指示を含む。終了指示が入力されたと判定された場合（ステップS314でYES）には、メイン処理を終了する。終了指示が入力されていないと判定された場合（ステップS314でNO）には、ステップS315に進む。

【0078】

ステップS315では、検出部121が、ステップS303で作業空間に挿入されたことが判定された注目物体が、作業空間から除かれたか否かを、撮像部120の撮像画像に基づいて判定する。注目物体が作業空間から除かれたと判定された場合（ステップS315でYES）には、ステップS301に戻る。注目物体が作業空間から除かれていないと判定された場合（ステップS315でNO）には、ステップS309に戻り、次に行われる操作を推定して認識領域を再設定する処理を繰り返す。

20

【0079】

ここで、図6(a)～(d)は、ステップS312において実行される、推定された操作のそれぞれに適した認識領域を設定する処理の一例を示すフローチャートである。各処理の詳細を説明する。

【0080】

<「紙移動」「本移動」「ドラッグ」操作の場合>

図6(a)は、推定された操作が「紙移動」「本移動」「ドラッグ」のいずれかであった場合に適した認識領域を再設定する処理の一例を示すフローチャートである。

30

【0081】

ステップS601では、設定部125が、現在の認識領域での認識部126によるジェスチャ認識を抑制する。図2(b)の認識領域214に示したように、推定されたジェスチャが「紙移動」「本移動」「ドラッグ」であった場合、情報処理装置100では、移動が終了するまでジェスチャ操作の認識を行わない。これは、ユーザが注目物体を動かす手が認識領域内に含まれる場合に、誤ってジェスチャ操作の動きであると認識しないためである。

40

【0082】

ステップS602では、検出部121が、撮像部120による撮像画像に基づいて、作業空間内の注目物体の移動が終了したか否かを判定する。本実施形態では、撮像画像の注目物体が写っている領域が所定時間以上静止した場合に、移動が終了したと判定する。なお、ジェスチャ操作が「ドラッグ操作」である場合には、タッチ位置の移動に応じて画像データの表示位置を変更する処理が並行して行われる。移動が終了したと判定された場合（ステップS602でYES）には、ステップS603に進む。移動が終了していないと判定された場合（ステップS602でNO）には、ステップS602を繰り返して終了を待機する。

【0083】

50

ステップS 6 0 3では、検出部1 2 1が、撮像部1 2 0の撮像画像に基づいて、静止した注目物体の外接矩形を示す位置情報を取得する。

【0 0 8 4】

ステップS 6 0 4では、設定部1 2 5が、ステップS 6 0 3で取得された位置情報に基づいて、外接矩形を包含する矩形領域を認識領域として設定する。ステップS 6 0 3、ステップS 6 0 4では、ステップS 4 0 1及びステップS 4 0 2と同様の処理が行われる。

【0 0 8 5】

以上のようにして、次に行われると推定される操作が「紙移動」「本移動」「ドラッグ」のいずれかであった場合には、移動中はジェスチャの誤認識を抑制した上で、移動後の注目物体に合わせてジェスチャ操作に適した認識領域を再設定する。

10

【0 0 8 6】

<「裏返し」操作の場合>

図6(b)は、推定された操作が、ジェスチャ操作である「裏返し」であった場合に適した認識領域を再設定する処理の一例を示すフローチャートである。

【0 0 8 7】

ステップS 6 1 1では、検出部1 2 1が、紙のどの部分が掴まれているかを特定する。例えば、矩形の紙であれば、ユーザにとっての前後左右いずれの辺を掴んでいるか、辺の中央付近か、端に寄っているか等を、操作面上の座標系に基づいて特定する。

【0 0 8 8】

ステップS 6 1 2では、設定部1 2 5が、特定された部分に基づいて、裏返し動作が行われる範囲を推定する。これは、ユーザが紙を裏返してから再び操作面上の元の位置に置こうとする場合、ユーザが紙のどの部分を掴んでいるかによって、裏返す動作において手が動く範囲の広さあるいは向きが違ふことが想定されるからである。なお、本実施形態において裏返し動作が行われる範囲とは、撮像部1 2 0の撮像画像において、裏返し動作を行うユーザの、掌が写った領域の重心が移動する範囲である。例えば、1枚の紙を裏返す場合、辺の中央を持てば、手をx y平面に沿う方向にはそれほど動かさずに裏返すことが可能な一方、紙の端をもって裏返す場合は、ある程度手がx y平面に沿う方向に移動する可能性が高いと考える事が出来る。一例として、ユーザが紙の手前側で左端寄りの位置を掴んでいた場合、裏返す際に手は右方向に移動する可能性が高い。また、ユーザが紙の右側の辺の後ろ寄りの位置を掴んでいた場合、裏返す際に手は前方向に移動する可能性が高い。このように、ユーザが紙の端に寄った位置を掴んでいた場合には、紙の辺の中央付近を掴んでいた場合より、裏返し動作が行われる範囲は広いと推定される。もちろん、変則的な動きで裏返し動作が行われる場合や、紙が矩形ではない場合も対応できるように考慮した上で、余裕を持った範囲を推定する。本実施形態では、ユーザの手の動きを追跡する際に、掌部分が撮像された領域の重心を追跡する。注目物体を掴む場合、掌部分が撮像された領域の重心は、注目物体より外側に存在する。従って、ユーザの手が注目物体を掴む場合、掌の重心位置が認識領域に含まれるように、外接矩形に対して所定の余裕分を加えた矩形領域を、新たな認識領域として設定する。標準的な操作に関してはこの様に処理を行うが、特定のユーザにおいて特定の動きをする場合は、この裏返し時の紙を掴む位置と操作領域の関係をユーザが登録できるようにしても良い。

20

30

40

【0 0 8 9】

ステップS 6 1 3では、設定部1 2 5が、ステップS 6 1 2で推定された範囲を包含する認識領域を再設定する。本実施形態では、ステップ3 0 6において設定された認識領域に、ステップS 6 1 2で推定された範囲を付加することで拡張するように再設定する。

【0 0 9 0】

以上のようにして、次に「裏返し」操作が行われると推定される場合には、裏返しを行う過程においてユーザの手が動くことが予想される範囲を含むように拡張したジェスチャ操作適した認識領域を再設定する。

【0 0 9 1】

<「ページめくり」操作の場合>

50

図6(c)は、推定された操作が、ジェスチャである「ページめくり」であった場合に適した認識領域を再設定する処理の一例を示すフローチャートである。

【0092】

ステップS621では、検出部121が、本のどの部分が掴まれているかを特定する。ここでは、本の表紙あるいはページについて、ユーザに掴まれているのは、ユーザにとっての前後左右いずれの辺か、いずれの端に寄っているか等を、操作面上の座標系に基づいて特定する。

【0093】

ステップS622では、設定部125が、本の見開き方向を特定する。本実施形態では、撮像部120による撮像画像に対する画像解析によって、本の背表紙がジェスチャ座標系に対してどのような位置に存在するかという情報や、文章が横書きか、あるいは縦書きかというような情報を取得する。また本実施形態では、ステップS305において注目物体を識別する際、本が閉じられているか開かれているかを含めて識別が可能である。従って、本が開かれている場合には、画像解析からページの境界線を検出し、その直交方向を本が開く方向と特定することが可能である。本が閉じられている場合で、背表紙が画像の特徴から判別しにくい場合(例えばパンフレットのような薄い本等)の場合には、ステップS621において特定したユーザの手の位置の情報からも、ユーザが本を開こうとする方向が推定可能である。

【0094】

ステップS623では、設定部125が、ページのめくり動作が行われる範囲を推定する。なお、本実施形態においてページのめくり動作が行われる範囲とは、撮像部120の撮像画像において、ページのめくり動作を行うユーザの、掌が写った領域の重心が移動する範囲である。この際、ステップS622で推定された方向には、当該方向に沿った1ページの幅の少なくとも2倍の幅を包含するように、ユーザの手が動くことが予想される範囲が推定される。例えば、ユーザの座標系に対して左右開きの本の場合、1ページの横幅(y軸方向の幅)を $d \times [dot]$ とすると、めくり動作を行うユーザの手が動く範囲は、少なくとも $2d \times [dot]$ の幅を有するように推定される。

【0095】

ステップS624では、設定部125が、ステップS623で推定された範囲を包含する認識領域を再設定する。本実施形態では、ステップ306において設定された認識領域に、ステップS623で推定された範囲を付加することで拡張するように再設定する。

【0096】

以上のようにして、次に「ページめくり」操作が行われると推定される場合には、ページめくりを行う過程においてユーザの手が動くことが予想される範囲を含むように拡張したジェスチャ操作適した認識領域を再設定する。

【0097】

<「画像変形」操作の場合>

図6(d)は、推定された操作が「画像変形」であった場合に適した認識領域を再設定する処理の一例を示すフローチャートである。

【0098】

ステップS631では、設定部125が、現在の認識領域での認識部126によるジェスチャ認識を抑制する。これは、ユーザがタッチ操作によって画像データの変形を指示するために動かす手が、認識領域内に含まれる場合に、誤ってジェスチャ操作の動きであると認識しないためである。

【0099】

ステップS632では、認識部126が、タッチ操作を認識し、表示制御部128が、認識された操作に応じて画像データを変形させて表示部129に出力する。

【0100】

ステップS633では、認識部126が、タッチ操作が終了したか否かを判定する。本実施形態では、距離センサ105の出力に基づいて、操作面の表面から所定の距離の範囲

10

20

30

40

50

内でユーザの手が検出されなくなった場合に、タッチ操作が終了したと判定する。タッチ操作が終了したと判定された場合（ステップS 6 3 3でYES）には、ステップS 6 3 4に進む。タッチ操作が終了していないと判定された場合（ステップS 6 3 3でNO）には、ステップS 6 3 3を繰り返して終了を待機する。

【0101】

ステップS 6 3 4では、検出部1 2 1が、撮像部1 2 0の撮像画像に基づいて、注目物体（ここでは、変形された画像データ）の外接矩形を示す位置情報を取得する。

【0102】

ステップS 6 3 5では、設定部1 2 5が、ステップS 6 3 4で取得された位置情報に基づいて、外接矩形を包含する矩形領域を認識領域として設定する。

10

【0103】

以上のようにして、次に行われると推定される操作が「画像変形」であった場合には、タッチ操作中はジェスチャの誤認識を抑制した上で、移動後の注目物体に合わせてジェスチャ操作に適した認識領域を再設定する。

【0104】

以上説明した通り、本実施形態では、ユーザが行おうとする操作を推定することで、次に行われるジェスチャ操作に適したジェスチャ認識領域を設定することができる。

【0105】

次に、ユーザが行うジェスチャ操作と、当該ジェスチャ操作に適切なジェスチャ認識領域の具体例を説明する。

20

【0106】

<「裏返し」の操作例>

図7は、「裏返し」操作の操作例を表す図であり、操作面を上から見た状態を示す。従って、紙面に向かって上下方向にy軸、左右方向にx軸が設定されているものとする。

【0107】

図7(a)は、作業空間に紙700が挿入された状態を示す。領域701aは、ステップS 306において設定部1 2 5に設定されるジェスチャ認識領域である。このとき、ユーザの手702の位置は、注目物体の輪郭付近にあり（ステップS 509でYES）、一部が注目物体に隠れており（ステップS 510でYES）、かつ、1箇所にはかない（ステップS 511でNO）。従って、特定部1 2 3により、ユーザの手の状態は「注目物体を掴む」姿勢であると特定され（ステップS 513）、注目物体が紙であることとの組み合わせにより、推定部1 2 4は、ユーザが「裏返し」を行おうとしていると推定する（ステップS 310）。

30

【0108】

図7(b)は、「裏返し」操作が推定されたことにより、ジェスチャ認識領域が再設定された様子を表す。注目物体を掴むユーザの手702は、紙の右側の辺の下寄りの位置を掴んでいると特定されることから（ステップS 611）、裏返し動作において、ユーザの手はy軸正方向に大きく移動する可能性が高い。従って、再設定される認識領域は、領域701bのように、上下方向（y軸方向）に大きく拡張され、かつ、左右方向（x軸方向）にも余裕を有するように、設定される（ステップS 613）。

40

【0109】

図7(c)は、実際にユーザが裏返し操作を開始した様子を示す。手の位置は、認識領域701b内部で、図7(a)及び(b)に比較して上（y軸の正方向）に移動している。また、紙700は、作業空間から取り除かれることなく内容が変化しているため、裏返されたことがわかる。

【0110】

図7(d)は、裏返し動作が終わり、手702が静止している状態を示す。この後、手702が認識領域701dから除かれる（領域701dから出る）ことによって、ジェスチャ操作の終了が判断される。認識部1 2 6では、紙700と共にユーザの手702も裏を向いたことや、手の位置の移動量をジェスチャモデルと比較して、「裏返し」操作を認

50

識する。例えば、本実施形態では、「裏返し」の完了が判断されたと同時に、裏返された紙の読み取りが実行される（ステップS313）。

【0111】

<「ページめくり」の操作例>

図8、図9は、「ページめくり」操作が推定された場合の認識領域の例を表す図である。図7と同様、操作面を上から見た状態を示すため、紙面に向かって上下方向にy軸、左右方向にx軸が設定されているものとする。

【0112】

図8(a)は、作業空間に本800が挿入された状態を示す。領域801aは、ステップS306において設定部125に設定されるジェスチャ認識領域である。このとき、ユーザの手802の位置は、注目物体の輪郭付近にあり（ステップS509でYES）、一部が注目物体に隠れており（ステップS510でYES）、かつ、1箇所しかない（ステップS511でNO）。従って、特定部123により、ユーザの手の状態は「注目物体を掴む」姿勢であると特定される（ステップS513）。従って、注目物体が本であることとの組み合わせにより、推定部124は、ユーザが「ページめくり」を行おうとしていると推定する（ステップS310）。

10

【0113】

図8(b)は、「ページめくり」操作が推定されたことにより、ジェスチャ認識領域が再設定された様子を表す。注目物体を掴むユーザの手802は、紙の右側の辺の下寄りの位置を掴んでいる（ステップS621）。そして、本800は、背表紙が注目物体の左側に存在している。また、文字は横書きである。これらの情報により、本800は左方向に開くと特定される（ステップS622）。従って、ページめくり動作において、ユーザの手はx軸負方向に大きく移動する可能性が高い。従って、再設定される認識領域は、領域801bのように、左方向（x軸負方向）に、少なくとも見開きページの幅が収まるように拡張され、かつ、上下方向（y軸方向）にも余裕を有するように、設定される（ステップS623）。

20

【0114】

一方、図8(c)は、作業空間に本803が挿入された状態を示す。領域801cは、ステップS306において設定部125に設定されるジェスチャ認識領域である。図8(c)と図8(a)との相違点は、本803が右方向に開く本であることである。従って、再設定されるジェスチャ認識領域は、図8(d)の801dのように、右方向（x軸正方向）に、少なくとも見開きページの幅が収まるように拡張され、かつ、上下方向（y軸方向）にも余裕を有するように設定される。

30

【0115】

同様に、図9(a)は、作業空間に本900が挿入された状態を示す。領域901aは、ステップS306において設定部125に設定されるジェスチャ認識領域である。ここで本900は、背表紙に位置やユーザの手902の位置から、上方向に開かれると特定される。従って、再設定されるジェスチャ認識領域は、図9(b)の901bのように、上方向（y軸正方向）に、少なくとも見開きページの幅が収まるように拡張され、かつ、左右方向（x軸方向）にも余裕を有するように設定される。

40

【0116】

同様に、図9(c)は、作業空間に本903挿入された状態を示す。領域901aは、ステップS306において設定部125に設定されるジェスチャ認識領域である。ここで本903は、背表紙に位置やユーザの手902の位置から、下方向に開かれると特定される。従って、再設定されるジェスチャ認識領域は、図9(d)の901dのように、下方向（y軸負方向）に、少なくとも見開きページの幅が収まるように拡張され、かつ、左右方向（x軸方向）にも余裕を有するように設定される。

【0117】

次に、図10は、「ページめくり」操作の操作例を表す図であり、操作面を上から見た状態を示す。従って、紙面に向かって上下方向にy軸、左右方向にx軸が設定されている

50

ものとする。

【0118】

図10(a)は、図8(b)を再掲するものである。領域801bは、左方向に開く本800に対するページめくり動作を認識する為に設定されたジェスチャ認識領域である。認識領域801bが設定された時点から、ユーザの手802が検出され始め、ジェスチャ操作としての動きの追跡が開始される。

【0119】

図10(b)は、実際にユーザがページめくりを行っている様子を示す。手802の位置は、認識領域801b内部で、図10(a)に比較して左(x軸の負方向)に移動している。また、本800は、作業空間から取り除かれることなく、左右方向に拡大され、内容が変化しているため、開かれたことがわかる。

10

【0120】

図10(c)は、ページめくり動作を終え、手802が静止した状態を示す。認識部126では、本800の見開きページの内容が変化したことや、手の位置の移動量をジェスチャモデルと比較して、「ページめくり」操作が開始されたことを認識する。例えば、本実施形態では、「ページめくり」操作中に手、および注目物体(本)が所定時間以上静止したことに応じて、新たに読み取り可能となった見開きページの読み取りが繰り返し実行される。そして、手802が認識領域801bから除かれる(領域801bから出る)ことによって、「ページめくり」操作の終了が判断される(ステップS313)。

【0121】

20

このように、本実施形態では、様々な注目物体に関して、注目物体に関連した複数種類のジェスチャ操作を認識する。その際、注目物体と周りの状況に応じて次に行われるジェスチャ操作を推定し、各操作に適したジェスチャ認識領域を設定することにより、大ぶりの動作が必要なくなるので、注目物体に関連したジェスチャ操作の操作性を向上させることができる。

【0122】

なお、本実施形態において、「紙移動」「本移動」「ドラッグ」「画像変形」はジェスチャ操作の準備段階の操作であるといえる。従って、準備段階の操作により、認識領域の再設定が行われた後、図3のフローチャートのステップS309からの処理を繰り返すことによって、次にジェスチャ操作行われる場合もある。例えば、「ドラッグ」操作によって画像データを移動させた後で、スワイプ操作により表示される画像データを変形するような場合には、再設定された認識領域が有効に用いられる。また、「紙移動」操作によって紙を移動させた後で、「裏返し」操作が行われる場合、ユーザが紙を掴んだことが検出されることに応じて再び「裏返し」に適した認識領域を再設定する処理が実行される。このように、様々な状況に対応して、次に行われるジェスチャ操作のために、適切な認識領域を設定することができるので、ユーザは不自然に思える大ぶりの動作をせずに、注目物体に関連したジェスチャ操作を行うことができる。

30

【0123】

<変形例>

上述した第1の実施形態において、注目物体は、紙・本・投影された画像データのいずれかとしたが、これに限らず例えば箱や瓶等の立体物であっても構わない。例えば、お菓子の箱のパッケージデザインを読み取る等の応用例も考えられる。

40

【0124】

また、第1の実施形態では、推定部124は、予め辞書に登録された所定のジェスチャの中から、ユーザが行おうとしている可能性が高いジェスチャ操作を特定することで、具体的なジェスチャ操作の推定を行った。ただし、より簡易的な推定処理でも構わない。例えば、ユーザが注目物体の左側を掴んでいる場合には、左手を使った操作が行われると推定し、設定部125には、撮像領域のうち、注目物体を掴むユーザの位置より左側をより多く含む一部領域をジェスチャ認識領域として再設定させる。あるいは逆に、ユーザが注目物体の右側を掴んでいる場合には、右手を使った操作が行われると推定し、設定部12

50

5には、撮像領域のうち、注目物体を掴むユーザの位置より右側をより多く含む一部領域をジェスチャ認識領域として再設定させるといった処理も可能である。

【0125】

また、第1の実施形態では、ユーザが注目物体を挿入したときの腕の状態に合わせて座標系を設定した(ステップS306)。さらに、注目物体に対して回転を含む移動がなされた場合には、回転が終了したときにステップS306と同様に処理を行うことにより、ユーザの腕の向きに基づく座標系を再設定することもできる。このような処理により、例えばユーザが立ち位置を変更するために注目物体を移動させたり、回転させたりした場合でも、ユーザの手の動きを精度よく区別して認識することができる。

【0126】

また、第1の実施形態では、1つの注目物体に関連した操作を1人のユーザが行う例を挙げて説明したが、複数の注目物体に対して複数のユーザがそれぞれ行うことも可能である。その場合には、注目物体が挿入されたか否かを判定する(ステップS303)際に、注目物体の個数を確認し、以降の処理では、注目物体の個数と同じ数の認識領域を扱う。また、注目物体とユーザの位置に合わせた認識領域を設定する処理(ステップS306)では、ユーザの腕の方向を区別することで、ユーザの人数を確認し、ユーザ毎に座標系を設定する。これにより、複数のユーザが同時に様々な方向からアクセスすることが可能な装置であっても、ユーザ毎に最適なジェスチャ認識領域を設定し、それぞれの注目物体に関連したジェスチャ操作の操作性を向上させることができる。

【0127】

<その他の実施形態>

また、本発明は、以下の処理を実行することによっても実現される。即ち、上述した実施形態の機能を実現するソフトウェア(プログラム)を、ネットワーク又は各種記憶媒体を介してシステム或いは装置に供給し、そのシステム或いは装置のコンピュータ(またはCPUやMPU等)がプログラムを読み出して実行する処理である。

【符号の説明】

【0128】

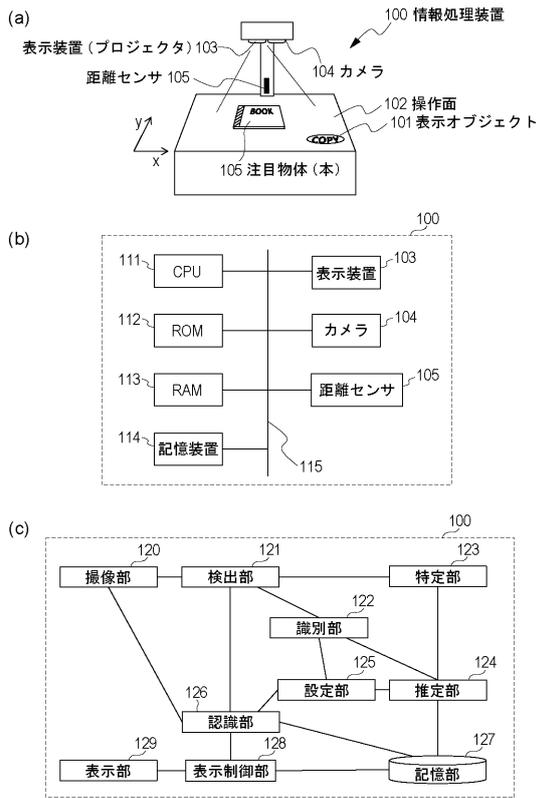
- 100 情報処理装置
- 120 撮像部
- 121 検出部
- 122 識別部
- 123 特定部
- 124 推定部
- 125 設定部
- 126 認識部

10

20

30

【図 1】



【図 2】

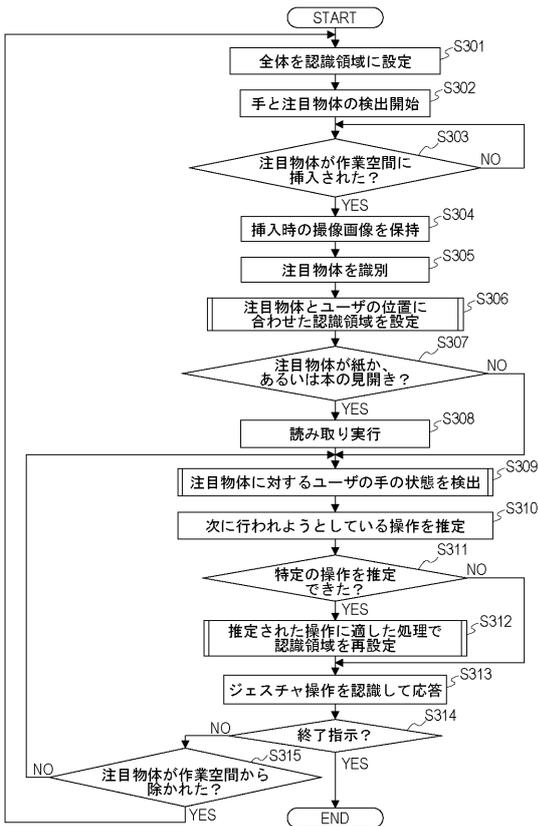
(a)

識別された注目物体	手の動き	指示内容	操作名称	ジェスチャ
紙	紙に手を載せて移動/回転	認識領域のリセット	紙移動	
紙	紙を裏返す	裏返し完了後に読み取り実行	裏返し	○
本	ページをめくる	めくり完了後に読み取り実行	ページめくり	○
本	本を握んで移動/回転	認識領域のリセット	本移動	
画像データ	左方向に動かす	次の画像データを表示	左スワイプ	○
画像データ	右方向に動かす	前の画像データを表示	右スワイプ	○
画像データ	画像データをタッチして移動	前の画像データを表示	ドラッグ	
画像データ	画像データを2箇所以上タッチしてピンチまたは回転	画像データの拡大・縮小・回転	画像変形	

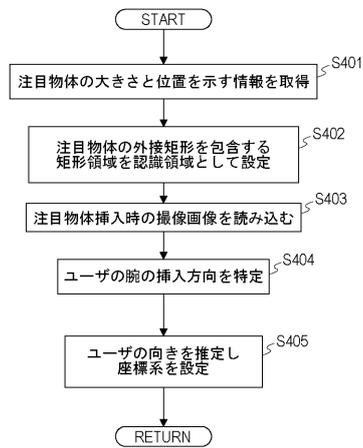
(b)

識別された注目物体	注目物体に対する手の位置	手の状態	推定される操作の名称	認識領域
紙、本	注目物体の上	上から押さえる	紙移動	(移動終了まで認識停止) 移動終了後の位置での注目物体の位置と向きに合わせて更新
紙	注目物体の輪郭付近	注目物体を握む	裏返し	所定量の余白を追加
本	注目物体の輪郭付近	注目物体を握む	ページめくり	見開きページが収まる大きさに合わせて更新
本	注目物体を挟む二か所	注目物体を持ち上げる	本移動	(移動終了まで認識停止) 移動終了後の位置での注目物体の位置と向きに合わせて更新
画像データ	注目物体の上	操作面上をタッチ	ドラッグ	(移動終了まで認識停止) 移動終了後の位置での注目物体の位置と向きに合わせて更新
画像データ	注目物体の上	操作面上を2箇所以上タッチ	画像変形	2箇所タッチを認識した時点で、最大の大きさに設定。 タッチ解除後、操作後の画像データの大きさと向きに合わせて更新

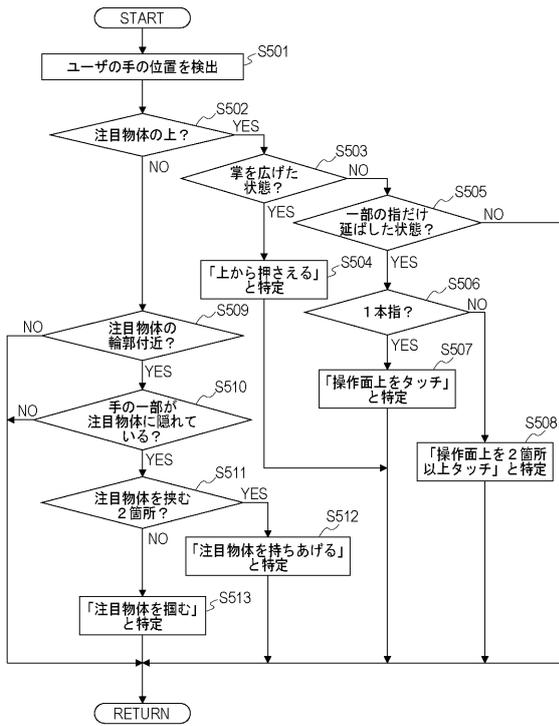
【図 3】



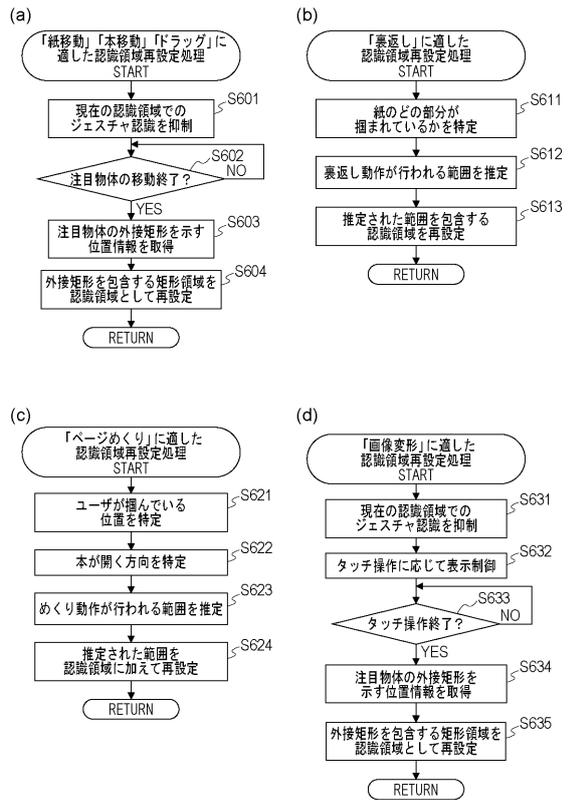
【図 4】



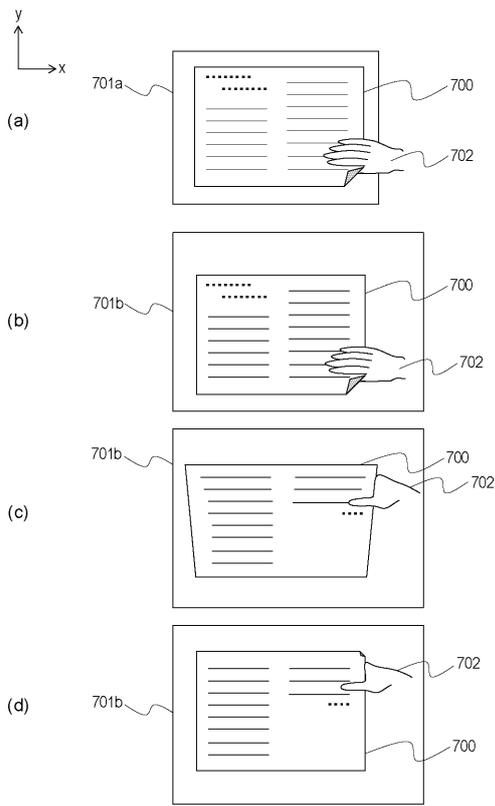
【図5】



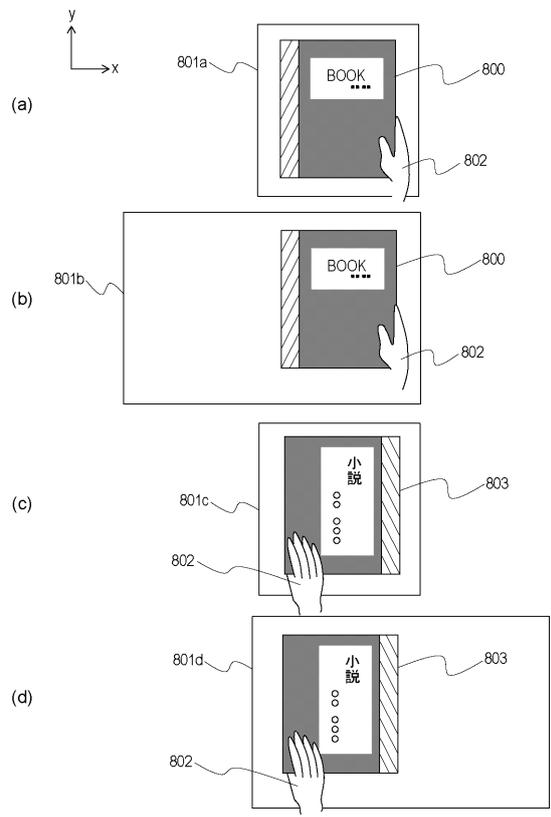
【図6】



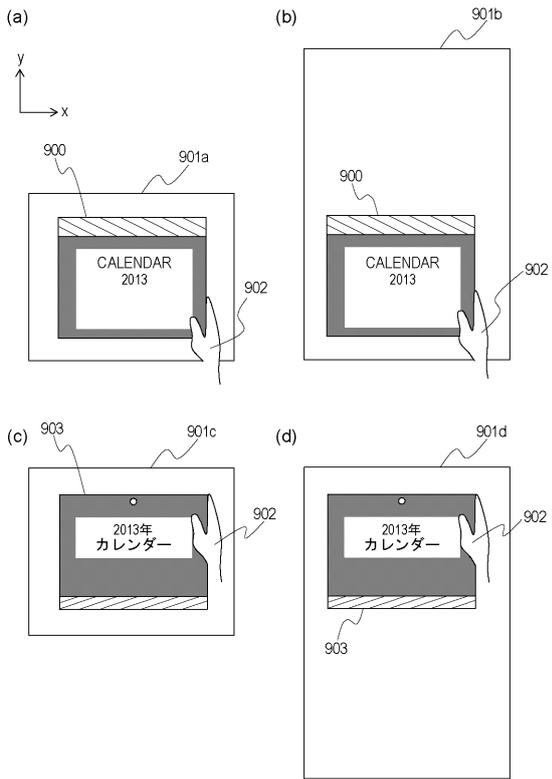
【図7】



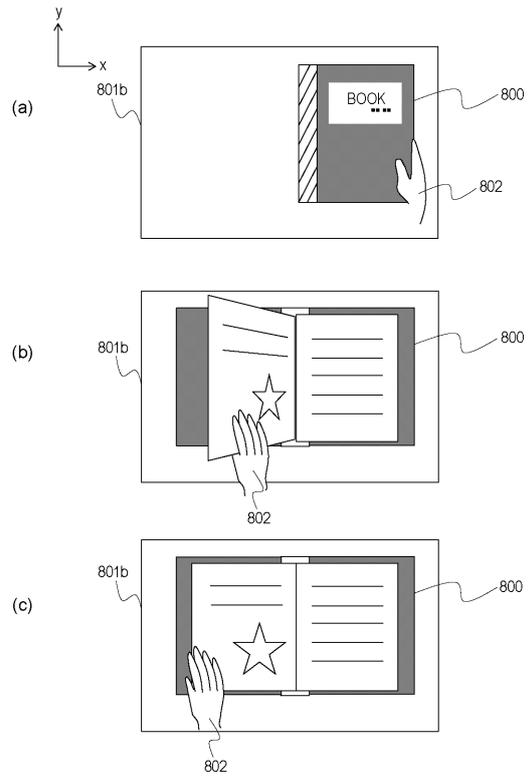
【図8】



【図9】



【図10】



フロントページの続き

- (56)参考文献 特開平07-175587(JP,A)
米国特許第05732227(US,A)
特開2011-039844(JP,A)
米国特許出願公開第2012/0119988(US,A1)
特開2012-027532(JP,A)
特開2001-056861(JP,A)
国際公開第2013/040048(WO,A2)
特開2013-016060(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/01
G06F 3/0346
G06F 3/048