

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第5807989号  
(P5807989)

(45) 発行日 平成27年11月10日(2015.11.10)

(24) 登録日 平成27年9月18日(2015.9.18)

(51) Int.Cl.

F 1

G06F 3/0346 (2013.01)

G06F 3/033

423

G06F 3/01 (2006.01)

G06F 3/01

310C

G06F 3/0481 (2013.01)

G06F 3/048

658B

請求項の数 20 (全 20 頁)

(21) 出願番号 特願2014-506442 (P2014-506442)  
 (86) (22) 出願日 平成24年4月6日 (2012.4.6)  
 (65) 公表番号 特表2014-514661 (P2014-514661A)  
 (43) 公表日 平成26年6月19日 (2014.6.19)  
 (86) 國際出願番号 PCT/US2012/032613  
 (87) 國際公開番号 WO2012/145180  
 (87) 國際公開日 平成24年10月26日 (2012.10.26)  
 審査請求日 平成27年4月1日 (2015.4.1)  
 (31) 優先権主張番号 13/092,115  
 (32) 優先日 平成23年4月21日 (2011.4.21)  
 (33) 優先権主張国 米国(US)

早期審査対象出願

(73) 特許権者 310021766  
 株式会社ソニー・コンピュータエンタテインメント  
 東京都港区港南1丁目7番1号  
 (74) 代理人 100099324  
 弁理士 鈴木 正剛  
 (72) 発明者 ジェフリー ロジャー スタッフォード  
 アメリカ合衆国、カリフォルニア州 94404、フォスター シティー、セカンド  
 フロア、イースト ヒルズデイル ブルバード 919

審査官 佐藤 匡

最終頁に続く

(54) 【発明の名称】注視支援コンピュータインターフェース

## (57) 【特許請求の範囲】

## 【請求項 1】

ユーザを、コンピューティングデバイスで実行するコンピュータプログラムとインターフェースさせるための方法であって、

ディスプレイ上の前記ユーザの注視点を識別し、

前記ディスプレイ上のカーソルの位置を移動させるための前記ユーザによる物理的行動の開始を検出し、前記カーソルは、前記コンピュータプログラムと関連付けられるフォーカス領域を画定及び検出し、

前記カーソルの現在位置と前記注視点との間の距離が閾値距離よりも大きいかどうかを判定し、

前記距離が所定の閾値距離よりも大きいときに、前記カーソルを、前記現在位置から前記注視点に近接した領域まで移動させ、

前記カーソルの移動では、前記カーソルの現在位置と前記注視点との間の前記距離に基づいて、前記カーソルを制御するマウスのカーソル移動速度を変更し、

前記カーソルは、その現在位置と前記注視点との間の距離が小さくなるにつれて移動速度が低下する、方法。

## 【請求項 2】

前記物理的行動は、マウスを移動、キーを押す、ボタンを押す、画面に触れる、タッチパッドに触れる、ジェスチャーを行う、または話す、のうちの1つである、請求項1に記載の方法。

**【請求項 3】**

前記距離の判定には、

前記現在位置と前記注視点との間の画面画素で測定される距離の値の決定が更に含まれる、請求項 1 に記載の方法。

**【請求項 4】**

前記距離の判定には、

前記現在位置および前記注視点がグラフィカルユーザインターフェース ( G U I ) の中の異なるウィンドウの中に位置しているかどうかの判定が更に含まれる、請求項 1 に記載の方法。

**【請求項 5】**

前記カーソルは、マウスカーソル、テキストカーソル、フィールド選択カーソル、またはページカーソルのうちの 1 つである、請求項 1 に記載の方法。

10

**【請求項 6】**

前記カーソルの移動には、

ウェブページ内の複数の入力フィールドを検出し、

前記ユーザがキーボードのキーを押したときに、前記複数の入力フィールドのうちから、前記注視点に最も近い入力フィールド上に前記カーソルを配置することが更に含まれる、請求項 1 に記載の方法。

**【請求項 7】**

前記ユーザの頭部の位置を追跡する、請求項 1 に記載の方法。

20

**【請求項 8】**

即時的な注視支援ナビゲーションを起動させる、前記ユーザからの入力の検出と、

前記入力の検出後に、前記カーソルを前記注視点まで移動させる、請求項 1 に記載の方法。

**【請求項 9】**

前記カーソルの移動の後に、前記カーソルを中心とする前記カーソルの周囲の領域を拡大して前記ディスプレイ上で表示し、かつ、

前記注視点が検出されている、ディスプレイ上の第 1 ウィンドウを決定し、前記第 1 ウィンドウは、前記ディスプレイ上の複数のウィンドウの表示をサポートするグラフィカルユーザインターフェース ( G U I ) に属する、請求項 1 に記載の方法。

30

**【請求項 10】**

前記第 1 ウィンドウではない G U I のウィンドウを前記ディスプレイ上で透明な表示とする、請求項 9 に記載の方法。

**【請求項 11】**

グラフィカルユーザインターフェース ( G U I ) を伴うシステムであって、

前記 G U I は 1 つ以上のディスプレイに描画されるものであり、かつ、当該 G U I を提供するコンピュータプログラムを実行するものであるプロセッサと、

入力デバイス上でユーザによって物理的行動が開始されるときに、前記 1 つ以上のディスプレイ上のカーソルの位置を移動させるように操作可能な、入力デバイスと、

前記ユーザの注視点を識別するように操作可能な注視検出モジュールと、を備え、前記コンピュータプログラムは、前記カーソルの現在位置と前記注視点との間の距離が閾値距離よりも大きいかどうかを判定し、前記コンピュータプログラムは、前記距離が所定の閾値距離よりも大きいときに、前記カーソルを、前記現在位置から前記注視点に近接した領域まで移動させ、前記カーソルの移動では、前記カーソルの現在位置と前記注視点との間の前記距離に基づいて、前記入力デバイスのカーソル移動速度を変更し、

前記カーソルは、その現在位置と前記注視点との間の距離が小さくなるにつれて移動速度が低下し、前記カーソルの移動では、ウェブページ内の複数の入力フィールドを検出し、ユーザ入力に応答して前記注視点に最も近い入力フィールドにページカーソルを移動し、前記ページカーソルは、どの入力フィールドがアクティブであるのかを示す、システム。

40

50

**【請求項 1 2】**

前記注視検出モジュールは、角膜反射の解析、またはカメラで撮影されるユーザ画像の解析のうちの1つを使用する、請求項11に記載のシステム。

**【請求項 1 3】**

前記注視検出モジュールは、前記ユーザの物理的特徴の1つ以上の場所を判定するため、複数のユーザ画像を解析する、請求項12に記載のシステム。

**【請求項 1 4】**

前記GUIは、前記注視点が前記GUIの中のオブジェクト上に配置されるときに、前記ユーザにヒントを提供する、請求項11に記載のシステム。

**【請求項 1 5】**

前記注視点に基づいて音声コマンドを実行するように構成された音声認識モジュールをさらに含む、請求項11に記載のシステム。

10

**【請求項 1 6】**

コンピュータデバイスで実行されるコンピュータプログラムとユーザとをインターフェースさせるための方法であって、

ユーザのディスプレイ上の注視点を識別し、

前記ディスプレイ上のカーソルの位置を移動させるための前記ユーザによる物理的行動の開始を検出し、前記カーソルは、前記コンピュータプログラムと関連付けられるフォーカス領域を画定するものであり、

前記カーソルの現在位置と前記注視点との間の距離が閾値距離よりも大きいかどうかを判定し、

前記距離が前記閾値距離よりも大きいときに、前記カーソルを、前記現在位置から前記注視点に近接した領域まで移動させ、

前記カーソルの現在位置と前記注視点との間の前記距離に基づいて、前記カーソルを制御するマウスのカーソル移動速度が変更され、及び、前記カーソルは、その現在位置と前記注視点との間の距離が小さくなるにつれて移動速度が低下し、かつ、

前記カーソルの移動の後に、前記カーソルを中心とする前記カーソルの周囲の領域を拡大して前記ディスプレイ上で表示する、方法。

**【請求項 1 7】**

前記注視点が検出されているディスプレイ上の第1ウィンドウを決定し、前記第1ウィンドウは、前記ディスプレイ上の複数のウィンドウの表示をサポートするグラフィカルユーザインターフェース(GUI)に属する、請求項16に記載の方法。

30

**【請求項 1 8】**

前記第1ウィンドウではないGUIのウィンドウを前記ディスプレイ上で透明な表示とする、請求項17に記載の方法。

**【請求項 1 9】**

前記GUIのウィンドウフォーカスを有するウィンドウを前記第1ウィンドウとし、

前記ウィンドウフォーカスにより、ウィンドウフォーカスされていないウィンドウとの相互作用を不可とする一方で、前記ウィンドウフォーカスに関連付けられたプログラムとユーザとの相互作用を可能とする請求項17に記載の方法。

40

**【請求項 2 0】**

開いているウィンドウを前記注視点に基づいて前記ディスプレイ上で再配置する、請求項17に記載の方法。

**【発明の詳細な説明】****【技術分野】****【0001】**

本発明は、グラフィカルユーザインターフェース(GUI)に関し、より具体的には、注視支援ナビゲーションを提供するGUIとインタフェースさせるための方法、システム、およびコンピュータプログラムに関する。

**【背景技術】**

50

**【0002】**

現在のコンピュータインターフェースにおいて、マウスで制御する画面カーソルは、手の速度およびマウスの感度によって決定される速度で、ディスプレイに描画されたデスクトップという仮想画面領域 (Virtual Real Estate:VRE) 上を進行する。ここ何年もの間に、より大きい画面サイズの出現、画面解像度の向上、マルチ画面化等のため、仮想デスクトップ画面領域のサイズが増大してきた。

**【発明の概要】****【発明が解決しようとする課題】****【0003】**

しかしながら、カーソルを制御する人間のインプットデバイス、すなわちマウスという伝達装置は、変わらないままである。例えば、 $640 \times 480$  のデスクトップ上でマウスを移動させるには、多くても数インチの手の移動だけしか必要とされず、通常は、手首の運動だけしか伴わない。しかしながら、マルチ画面を使用するときの $2560 \times 1600$  以上のデスクトップ解像度の場合、ユーザは、単に手首を使用するだけで、マウスカーソルをデスクトップの片側から反対側まで、または一方のディスプレイからもう一方のディスプレイまで移動させることができない。このような状況では、あるときには「マウスハンドシャッフル」が必要とされ、その場合、ユーザは、マウスをマウスパッドから外れるくらい遠くまで進め、次いで、マウスを持ち上げて逆戻りさせ、再度マウスを降ろし、そして再度マウスをパッド上で移動させる場合がある。これは、大きいディスプレイとのインターフェースをうんざりさせる、煩雑な操作である。

10

以下の実施形態は、このような背景においてなされたものである。

**【課題を解決するための手段】****【0004】**

本発明の実施形態は、ユーザをグラフィカルユーザインターフェース (GUI) とインターフェースさせるための方法、システム、およびコンピュータプログラムを提供する。本発明は、プロセス、装置、システム、デバイス、またはコンピュータが読み出し可能な媒体上の方法等の、数多くの方法で実現することができることを理解されたい。本発明の独創的な実施形態は、以下で説明する。

**【0005】**

一実施形態において、方法は、ユーザが見ている点、いわゆる注視点 (POG:Point Of Gaze) を識別するための操作を含む。ディスプレイ上のカーソルの位置を移動させるための、ユーザによる物理的行動の開始が検出され、ここでカーソルは、GUIを実行するコンピュータプログラムと関連付けられるフォーカス領域を画定する。さらに、本方法は、カーソルの現在位置と POGとの間の距離が閾値距離よりも大きいかどうかを判定するための操作を含む。POGの判定および物理的行動の開始の検出に応じて、カーソルを、現在位置から POGに近接した領域まで移動させる。

30

**【0006】**

別の実施形態において、GUIを伴うシステムは、プロセッサと、入力デバイスと、注視検出モジュールとを含む。プロセッサは、GUIを提供するコンピュータプログラムを実行し、GUIは、1つ以上のディスプレイの中に描画される。入力デバイスは、入力デバイス上でユーザによって物理的行動が開始されるときに、1つ以上のディスプレイ上でカーソルの位置を移動させるように操作可能である。さらに、注視検出モジュールは、ユーザの POGを識別するように操作可能であり、コンピュータプログラムは、カーソルの現在位置と POGとの間の距離が閾値距離よりも大きいかどうかを判定する。コンピュータプログラムは、距離の判定および物理的行動の検出に応じて、現在位置から POGに近接する領域までカーソルを移動させる。

40

**【0007】**

さらに別の実施形態では、コンピュータによって実行されたときに、ユーザをコンピュータプログラムとインターフェースさせるための方法を実現するプログラム命令を有する、非一時的なコンピュータが読み出し可能な媒体である。本方法は、ユーザの注視点 (P

50

OG)を識別するための操作を含む。ディスプレイ上のカーソルの位置を移動させるための、ユーザによる物理的行動の開始が検出され、カーソルは、GUIを実行するコンピュータプログラムと関連付けられるフォーカス領域を画定する。さらに、本方法は、カーソルの現在位置とPOGとの間の距離が閾値距離よりも大きいかどうかを判定するための操作を含む。POGの判定および物理的行動の開始の検出に応じて、カーソルを、現在位置からPOGに近接した領域まで移動させる。

#### 【0008】

他の形態は、添付図面と関連してなされる以下の詳細な説明から明らかになるであろう。

#### 【図面の簡単な説明】

10

#### 【0009】

本発明は、添付図面と関連してなされる以下の説明を参照することによって最良に理解され得る。

#### 【0010】

【図1】一実施形態による、長距離にわたってカーソルを移動させるために、ユーザに必要とされる手間を例示する図である。

【図2A】長距離にわたってカーソルを移動させるための手間を低減する、注視支援コンピュータシステムの実施形態を示す図である。

【図2B】長距離にわたってカーソルを移動させるための手間を低減する、注視支援コンピュータシステムの実施形態を示す図である。

20

【図3】一実施形態による、注視点(POG)とカーソルの現在位置との間の距離の推定を例示する図である。

#### 【図4A】角膜反射を使用するPOG検出の実施形態を例示する図である。

【図4B】一実施形態による、画像解析および顔追跡を使用するPOG検出のための方法を例示する図である。

【図5】一実施形態による、注視検出を使用してハンドヘルドデバイスで高速にスクロールするための方法を例示する図である。

【図6】一実施形態による、注視検出およびユーザ入力を使用する高速カーソル移動のための方法を例示する図である。

#### 【図7】一実施形態による、注視支援ゲームインターフェースを示す図である。

30

【図8】一実施形態による、ジェスチャーに対応するウィンドウグラフィカルユーザインターフェースを例示する図である。

【図9】一実施形態による、注視検出、音声認識、および1つ以上のコンピュータ周辺機器を利用する、インターフェースを例示する図である。

【図10】本発明の実施形態を実現するためのコンピュータシステムの簡単化された模式図である。

【図11】本発明の一実施形態に従って、ユーザをコンピューティングデバイスで実行するコンピュータプログラムとインターフェースさせるためのアルゴリズムを例示するフローチャートである。

#### 【発明を実施するための形態】

40

#### 【0011】

本発明の実施形態は、ディスプレイにおけるカーソルの移動を支援するために、およびより少ない手間で目標カーソル場所に到達するために、注視追跡を利用する。一実施形態において、注視検出は、ユーザとコンピューティングデバイスとの間のインターフェースを強化するために、物理的行動として表されるユーザの意図と組み合わせられる。

#### 【0012】

本発明が、これらの特定の詳細の一部または全部を伴わずに実践され得ることが明らかになるであろう。他の事例において、よく知られているプロセス操作は、本実施形態を不必要に曖昧にしないために、詳細に説明されていない。

#### 【0013】

50

図1は、一実施形態による、長距離にわたってカーソルを移動させるために、ユーザに必要とされる手間を例示する。3つのディスプレイ110a、110b、および110cは、コンピュータ108に接続される。ウィンドウズ(登録商標)環境は、3つのディスプレイの中に示され、よって、ウィンドウズ(登録商標)環境のデスクトップは、3つのディスプレイを包含し、各ディスプレイが、デスクトップの区画を表す。3つの区画は、領域が連続するデスクトップを形成することをシミュレーションするために、論理的に相互接続される。換言すれば、3つのディスプレイは、あたかも単一のディスプレイがあるかのように動作するが、この単一のディスプレイ上の描画は、3つの部分に分けられ、各部分が、ディスプレイ110a、110b、および110cのうちの1つに示される。

## 【0014】

10

ユーザ102がマウスを移動させると、マウスは、3つのディスプレイを横断して移動することができるが、1つの中にだけ存在することができる。例えば、3つのディスプレイが単一のデスクトップを表すので、ディスプレイ110aの中のマウスがディスプレイ110aの右に移動すると、マウスは、ディスプレイ110bの左に現れる。

## 【0015】

3つのディスプレイは、ユーザに大きいデスクトップを提供し、これは、複数のウィンドウを同時に表示させることを所望するユーザにとって有用な特徴である。しかしながら、大きいデスクトップの中でマウスを移動させることは、非常に時間がかかる可能性があり、マウスの感度が高い場合はなおさらである(すなわち、デスクトップ上のカーソルの同じ変位量について、高感度を有するマウスは、低感度を有するマウスよりも多くのマウスの運動を必要とする)。したがって、正確なマウスの運動およびオブジェクト選択(例えば、グラフィックの作成および編集)を必要とする活動には、高感度のマウスが有効であるが、高感度のマウスは、デスクトップ全体にわたってマウスを迅速に移動させるには不利である。例えば、ユーザ102が、ウィンドウ112の中の位置116からディスプレイ110a上の位置118までマウスカーソルを移動させたい場合は、マウス104を右から左に長い距離を移動させなければならない。マウスの運動範囲は、しばしば制限されるので、マウスカーソルの大きい変位は、しばしば、ユーザが、マウスを左に移動させ、マウスを持ち上げ、マウスを右に置き、次いで、再度マウスを左に移動させることが必要であり、カーソルを位置118まで移動させるためにこの操作106を複数回繰り返す。転送中、マウスカーソルは、軌跡120に沿って移動し、これは「マウスハンドシャッフル」による複数回の停止を含む。

20

## 【0016】

30

本明細書で使用される「カーソル」は、コンピュータプログラムのためのフォーカスの形態を示す、コンピュータプログラムと関連付けられるポインタまたは表示子を指す。カーソルは、マウスによって制御されるポインタの現在位置を示し、ディスプレイ上に示される、マウスカーソル、またはテキストがテキストフィールド(例えば、ワードプロセッサ、ウェブページの中のフィールド)に入力される場所を示す、テキストカーソル(例えば、垂直線)、あるいはページ(例えば、ウェブページ)に対して行われる操作(例えば、前ページ、次ページ、上にスクロール等)のための参照点を示す、ページカーソル等を指し得る。カーソルは、矢印、手、指、垂直バー、スクロールバー等の異なる形態で画面上に表示可能であり、または時々隠されて、カーソルを見えないようにし得るが、それでも、「次ページ」等のいくつかのインターフェース操作を行うために利用することができる。カーソルは、本明細書において、「フォーカス領域」、「マウスカーソル」、「ページカーソル」、「ウィンドウズ(登録商標)カーソル」、「テキスト挿入点」等と称され得る。実施形態は、特定のカーソル形態を使用して提示されるが、あらゆるタイプのカーソルを利用することができ、したがって、提示される実施形態は、排他的または限定的であると解釈されるべきではなく、例示的または実例であると解釈されるべきである。

40

## 【0017】

一般的に、ユーザ102は、カーソルの所望の移動先を見ており、マウスカーソルの現在の場所は見ていない。本発明の実施形態は、ユーザをコンピュータシステムとインター

50

フェースさせる方法を向上させるために、注視検出 (Gaze Detection:GD) を利用する。例えば、注視検出は、ユーザがどこまでカーソルを移動させたいのかを判定し、次いで、より速くかつより少ない手間で所望の場所までカーソルを移動させるために使用される。場合によっては、ユーザは、マウスがディスプレイ上のどこにあるのか分からなくなることさえあり、ユーザは、自分がカーソルを移動させ始めることができる前までずっと、カーソルを捜すことに時間を費やさなければならない。しかしながら、注視検出を使用すれば、カーソルがどこにあるかではなく、カーソルを進める必要がある場所を知っていることが必要であるので、マウスの場所を見つけることは必要とされない。

【0018】

いくつかの実施形態では、コンピュータシステムとインターフェースさせるために、注視だけを使用することができるが、他の実施形態では、ユーザによる関連する物理的運動を必要とすることに留意されたい。ユーザのPOGは、不安定となることがあり、よってPOGだけを使用することは、問題を引き起こし得るか、ユーザの確認を必要とし得るか、または完了するまでより長い時間がかかる操作を必要とし得る（例えば、1つの点で2秒間注視を固定することを必要とする）。注視と他の行動（カーソルがPOGの通常領域にあれば、マウスによってカーソルを移動させる、等）とを組み合わせることによって、使い勝手の良いインターフェースが提供される。

10

【0019】

図2A～2Bは、長距離にわたってカーソルを移動させるための手間を低減する、注視支援コンピュータシステムの実施形態を示す。一実施形態において、カーソルの制御は、ユーザの注視によって補助される。簡潔にする目的で、本発明の範囲の実施形態は、ディスプレイに描画されるデスクトップ用の標準的なグラフィカルユーザインターフェース（GUI）上のマウスカーソルの操作を利用して説明される。他の実施形態は、異なる形態のカーソルもしくはアバター制御、異なる入力機構（例えば、キーボード、キーパッド、タッチ画面、ペダル等）を利用可能であり、または異なる表示機構（例えば、大型ディスプレイ、マルチディスプレイ、区画に分割されるディスプレイ等）を参照し得る。したがって、例示される実施形態は、排他的または限定的であると解釈されるべきではなく、例示的または実例であると解釈されるべきである。

20

【0020】

眼球追跡または視線追跡は、注視点（POG）（すなわち、ユーザが見ている場所）、または頭部に対する眼球の運動のいずれかを測定する過程である。POGは、人の注視が向けられる空間の中の点であり、また、各眼球の網膜（窩）の最高鋭敏領域の中心で結像される空間の中の点としても定義されている。注視検出に関するより多くの詳細は、図3および図4を参照して下で提供される。

30

【0021】

図2Aを参照すると、ユーザ102は、ディスプレイ110cの中の位置204からディスプレイ110aの位置208までマウスカーソルを移動させることを所望する。コンピュータ108上で実行する注視検出システムは、GUIのデスクトップを備えるディスプレイ内のユーザのPOGを決定する。ユーザは、所望するマウスカーソルの行き先に注視210を固定しており（すなわち、カーソルの目標方向を見ており）、注視検出システムは、およその場所206を決定する。ユーザ102の注視210が目標上にある間に、ユーザ102は、カーソルを目標へ移動させることを意図して、マウス104を移動202させる。コンピュータシステム108は、マウスの移動開始を検出すると、行き先の近くにカーソルを移動させるために注視検出を使用する（すなわち、マウスカーソルが、行き先の近くの場所または行き先まで「ジャンプ」する）。注視支援を利用してカーソルを移動させるときに、ユーザが注視支援運動を行うためにマウスをドラッグすることを必要とすることなく、カーソルが第1の場所から第2の場所まで移動することに留意されたい。注視検出は、カメラ214によって撮影される画像の画像解析に基づくが、他のタイプの注視検出も可能である。

40

【0022】

50

システムは、マウスカーソルを、注視検出システムによって計算される POG である、位置 206 まで移動させ、システムは次いで、カーソルの最終的な行き先を見つけるために、単にマウスの移動だけに依存する。マウスカーソルの移動を完了するために必要とされることは、その行き先であるので、ユーザは、さらに現在マウスがどこにあるのかを知る必要はない。このように、ユーザは、カーソルが現在デスクトップの中のどこにあるのかを見つけるために、マウスカーソルを捜す必要はない。

#### 【0023】

一実施形態において、カーソルの移動速度は、可変的であり、よって、カーソルが POG に接近するにつれて、カーソルの移動速度は低下する。このように、カーソルが POG から遠く離れているときに、カーソルは速く移動するが、カーソルが行き先に接近し始めると、カーソルは減速し、マウスの感度が増加して、ユーザが行き先の近くでマウスカーソルをより繊細に制御することを可能にする。ユーザが求めていることは目標に速く到達することであり、行き先に到達するためにマウスが従う軌跡は重要ではないので、カーソルが目標から遠く離れている間、繊細な感度は必要とされない。

10

#### 【0024】

注視検出を使用することによって、システムは、必要なときに、すなわち、ユーザが関心の領域で操作しているときに、より高いレベルのマウス感度を提供することができる。他の加速方式は、ユーザがどのくらい速くマウスを移動させているかに基づくが、これら的方式は、しばしば、予期しない結果をもたらし、また、長い距離を通じてカーソルを速く移動させるために、感度と能力との間に平衡をもたらすように適切に調整することが難しい。別の実施形態において、マウスカーソルは、マグネット効果を提供し、よって、マウスカーソルは、ディスプレイ上の POG に引き付けられる。

20

#### 【0025】

一実施形態において、マウスの運動は、注視検出および POG 領域の中の表示オブジェクトのコンテキストによって支援される。このように、POG が、複数の画面オブジェクトから 1 つ以上の画面オブジェクトを含む領域の中に置かれる場合、カーソル位置は、推定される POG に最も近い複数から画面オブジェクトに変更される。例えば、ユーザの注視の目標領域が、ウィンドウズ（登録商標）デスクトップの左下のクリック起動アイコンの近くにあると思われる場合、カーソルの位置は、POG に最も近いクリック起動アイコンに変更される。

30

#### 【0026】

さらに別の実施形態において、注視検出は、1 つのウィンドウの中のフィールド間のナビゲーションを支援するために使用される。ユーザがキーボードナビゲーションを使用している場合、ページカーソルは、キーボード入力および POG の組み合わせに基づいて移動し、よって、POG は、ナビゲーションを迅速化するために使用され、別の場合であれば注視支援を伴わずにより多くのキーストロークを費やすことになる。ページカーソルは、本明細書において、ワードプロセッサでタイプするときの挿入点、または複数の使用可能な入力フィールドを含むフォーム内のアクティブなフィールド等の、ページのフォーカスを示す、ページの中のポインタを指す。したがって、マウスカーソルは、マウスの移動に基づいて移動する、マウスと関連付けられるポインタであるので、ページカーソルは、マウスカーソルと異なる。例えば、複数の入力フィールドを伴い、また、ページ上の最上部フィールドに位置するページカーソルを伴うページにおいて、ユーザは、ページの最下部のフィールドまで移動させることを所望する。注視検出を伴わない、標準的なシーケンスは、ユーザが、タブキーまたは下向き矢印キーを複数回押して（または、当然、マウスを使用して）、全ての中間フィールドを通過して、最上部のフィールドから最下部のフィールドまで進行する。しかしながら、注視検出を有する場合、ユーザは、最下部フィールドに注目し、次いで、ユーザがタブキー（またはいくつかの他のキー）を押すと、ページカーソルが最下部フィールドまで即時にジャンプし、ユーザのキーストロークを節約し、またマウスを使用する必要はない。

40

#### 【0027】

50

いくつかのウインドウズ（登録商標）に基づくG U Iは、ウインドウフォーカスの概念を有する。ある時点で、ユーザとの相互作用のために1つのウインドウだけが選択され、すなわち、デスクトップが選択され、かついかなるウインドウもフォーカスされていなければ、選択されたウインドウが「フォーカス中」になる。さらに別の実施形態において、注視検出は、デスクトップの中のフォーカス中のウインドウを変更するために使用される。ユーザがデスクトップのウインドウに注目し、マウスを移動させ始めると、ユーザが注目しているウインドウが、G U Iのフォーカスを有するウインドウになる。カーソルはまた、ウインドウの中心、ウインドウ上の最上部、最下部、P O G等の、ウインドウのいくつかの場所にも位置付けられる。

## 【0028】

10

眼球の移動が不安定になる場合もあることから、あるときにはマウスカーソルが不安定に振る舞い、ユーザがマウスカーソルに追従することが困難になり、注視だけによってマウスカーソルを制御することが困難になり得る。加えて、あるときには、ユーザは何かを読みみたいだけの場合もあり、P O Gに従ってカーソル移動させることが、気を散らせて、イライラさせることになる。こういった理由で、いくつかの実施形態は、ユーザがカーソルを移動させるという意図を示す必要があり、この意図は、マウスまたはキーを押すこと等の、注視を変更すること以外の物理的行動によって表される。

## 【0029】

一実施形態では、注視検出が起動する前に、注視が1つの目標上にある最小時間量を必要とする、閾値タイマーが定義される。タイマーは、潜在する不安定な注視の変更のフィルタリングを可能にする。例えば、ユーザは、自分がマウスを移動させ始めるときに、注視検出が行われる前に、少なくとも1秒にわたって画面のウインドウを見なければならぬ。

20

## 【0030】

図2Bは、一実施形態による、ユーザをコンピュータシステムとインターフェースさせるための環境を示す。大型ディスプレイ252は、3つの隣接するパネルに分割される。ウインドウ260は、コンピュータアプリケーションと関連付けられ、3つの異なるパネルを横断して展開する。コンピュータアプリケーションは、ウインドウ260の異なる領域に位置することができる、画面オブジェクト（例えば、ツールボックス262a）を含む。一実施形態において、画面オブジェクト262aは、マウスカーソルが位置するディスプレイのパネルの中に、またはマルチディスプレイの場合、マウスカーソルが位置するディスプレイの中に位置する。

30

## 【0031】

最初に、（位置204の）マウスカーソルおよびツールボックス262aは、ディスプレイ252の左パネルにある。ユーザ102は、注視210をP O G256に置き、マウス104を移動させ202始める。上で説明したように、システムが、マウスを移動させるという物理的行動によって表されるユーザの意図を検出すると、マウスカーソルは、P O G256上にある、またはP O G256の近傍にある、位置258までジャンプする。マウスカーソルが左パネルまでジャンプしたので、マウスカーソルの現在位置と関連付けられる画面オブジェクトも左パネル上の位置262bまで移動する。したがって、ユーザは、注視支援カーソル運動、ならびに画面オブジェクトの注視支援運動から恩恵を受ける。

40

## 【0032】

一実施形態において、システムは、注視支援カーソルナビゲーションを使用する前に、P O G256がマウスカーソルの現在位置からの閾値距離の範囲内にあるかどうかを判定する。一実施形態において、閾値距離（例えば、100画面画素）は、P O G256の周囲に円254を画定する。正方形、長方形、橢円、八角形、任意の正多角形等の、他のタイプのP O Gの周囲の領域はまた、いつ注視支援を使用するのかを決定することもできる。カーソルの現在位置とP O Gとの間の距離が閾値距離よりも大きい場合（または現在位置が、P O Gの周囲の画定領域の外側にある場合）、マウスカーソル運動は注視支援され

50

る。しかしながら、距離が閾値距離未満である（または現在のマウスカーソルが領域の内側にある）場合、ユーザは、近隣の行き先までマウスを容易に移動させることができるので、注視支援は使用されない。このように、注視支援は、より大きいカーソル移動のためにだけ使用され、ユーザが小さいカーソル移動を行っているときの混乱を回避する。

【0033】

さらに別の実施形態において、注視支援が使用されているかどうかの判定に対するトリガーは、POGが現在のマウスカーソルと同じウィンドウにあるかどうかである。この場合、図2Bで表されるシナリオは、マウスが同じウィンドウ内を移動しているので、注視支援は起動されない。

【0034】

図2Bの実施形態において、マウス運動の開始は、注視支援を起動させる、ユーザの物理的行動である。他の実施形態は、キーボードを押すこと、ボタンを押すこと、画面に触れること、話すこと、指をスナップすること、拍手すること等の、他のタイプのユーザの物理的行動をトリガーとして利用し得る。

【0035】

一実施形態において、ユーザは、画面上のマウスカーソルを移動させることを意図することなく、マウスを移動させる。例えば、ユーザは、マウスを保持しながらウェブページを読んでいる場合があり、ユーザは、僅かな手の運動のため、マウスを数ミリメートル移動させ得る。一実施形態では、ランダムで僅かなマウスの運動をフィルタリングするために、注視支援を起動させる前のマウスに対する運動の閾値が設定される。したがって、注視支援は、マウスが少なくともマウスの閾値距離移動したときにだけ開始する。一実施形態において、閾値距離は、約0.5~5センチメートルのマウスの運動であり、別の実施形態では、約0.5~4センチメートルであり、別の実施形態では、1センチメートルを超えるが、インターフェース環境に基づいて、他の閾値も可能である。別の実施形態において、閾値距離は、画面上のマウスカーソルの運動に関して測定され得る。例えば、閾値は、画面上で少なくとも2、5、10、20、25画素、または20を超える画素のマウスカーソルの移動に設定され得るが、他の値も可能である。

【0036】

さらに、初期位置204からPOG258の近くの位置までマウスカーソルを移動させることは、複数の方法で行うことができる。例えば、カーソルを、元の位置から消滅させて、行き先の位置に再び出現させる（すなわち、カーソルをジャンプさせる）ことができ、またはカーソルを、元の位置と行き先との間でデスクトップを通して迅速に移動させることができ、またはカーソルを、元の位置で短く点滅させ、次いで、行き先位置に再び出現させて点滅させてもよく、またはカーソルを、デスクトップを横断して迅速に移動させ、行き先に接近するにつれて減速させてもよく、またはカーソルを、元の位置と行き先との間の軌跡の中で、デスクトップ上の複数の地点で短時間出現させてもよい（すなわち、マウスの痕跡を残す）、等である。したがって、注視支援を使用してマウスカーソルが移動していることをユーザに示すための、多数の方法がある。

【0037】

図3は、一実施形態による、POGとカーソルの現在位置との間の距離の推定を例示する図である。異なる方法は、POGと現在のカーソルとの間の距離を測定するために、異なる距離メトリックを使用し、該距離は、注視支援によるカーソルの移動が利用されているかどうかを判定するために使用される。

【0038】

一実施形態において、距離は、画素で測定される。一般的に、ディスプレイは、画素について水平および垂直の解像度を有する。デスクトップ上でPOGの場所が推定されると、POGと対頂点の現在のカーソルとを含む、長方形の対角線の画素のサイズとして、距離312が算出される。POGおよびカーソルが水平または垂直に整列する場合、距離は、それぞれ、2つの点の間の水平または垂直の画素数に等しくなる。

【0039】

10

20

30

40

50

別の実施形態において、距離は、POGとカーソルとの間の実際の物理的距離として測定される。POGは、ディスプレイ上の第1の点を画定し、カーソルは、ディスプレイ上の第2の点を画定する。距離は、第1の点と第2の点との間の距離（例えば、インチで測定される）として算出される。

#### 【0040】

さらに別の実施形態において、距離は、ユーザ102（例えば、眼球の間、鼻の先端部等）を頂点とする角度308として測定され、第1の光線302は、該頂点からPOG256まで進み、第2の光線310は、該頂点からカーソルの現在位置204まで進む。したがって、角度308がより広くなるほど、マウスは、POGに到達するまで、より多く進行しなければならない。注視起動のための閾値は、角度で定義される（例えば、角度308が30°よりも大きいときに注視支援運動を使用するが、他の値も可能である）。別の実施形態において、注視支援は、角度308およびユーザの奥行304に基づいて起動され、ここで、奥行は、ユーザからディスプレイの中のある場所までの距離である。例えば、ディスプレイからより遠くに位置するユーザは、ユーザがディスプレイにより近い場合よりも、同じ距離を進行するためにより小さい角度を画定することになるので、角度で測定される閾値は、ユーザがディスプレイからより遠く離れているときには小さくなり得る。

#### 【0041】

図4Aは、角膜反射を使用するPOG検出の実施形態を例示する図である。眼球追跡装置は、眼球の位置および眼球の移動を測定するためのデバイスである。眼球の移動および注視方向を測定するための、いくつかの方法がある。いくつかの方法は、そこから眼球位置が抽出されるビデオ画像を使用し、他の方法は、探りコイルを使用するか、電気眼球図に基づいているか、または注視検出に赤外光を使用する。注視検出を利用することによって、デバイスは、より直観的な相互作用、および改善された体験をユーザに提供することができる。

#### 【0042】

図4Aは、ユーザに向かって方向付けられる光の角膜反射が解析される、注視検出のための一実施形態を例示する。図4Aで示されるように、赤外光源404は、赤外光を放射し、ユーザ402の角膜からの赤外光の反射が、赤外線カメラ406によって撮像される。次いで、ユーザ402のPOGを決定するために、この反射が解析される。赤外光源404はまた、カメラ406内に組み込まれ得る。別の実施形態では、赤外光の代わりに可視スペクトルの光が使用され、また、可視光および赤外光の組み合わせも可能である。

#### 【0043】

POGを計算するために、空間の中の眼球の位置および方向を推定するよう、虹彩-強膜境界（縁）の透視投影が使用される。POG推定は、瞳孔および1つ以上の角膜反射の中心位置を決定する。角膜反射は、眼球を照らす光源（例えば、赤外線）の仮想画像であり、凸面鏡として作用する角膜の前面によって作成される。角膜反射を使用する注視検出に関するさらなる詳細は、E.GuestinおよびM.Eizenmanによる「General theory of remote gaze estimation using the pupil center and corneal reflections」、IEEE Transactions on Biomedical Engineering, 2006で見られ、参照により本明細書に組み込まれる。

#### 【0044】

図4Bは、一実施形態による、画像解析および顔追跡を使用するPOG検出のための方法を例示する図である。一実施形態において、注視検出は、眼球416および418、顔、鼻422、口腔420、胴体414等のユーザの特徴の検出および追跡を含む、ユーザ領域410のビデオ画像412の画像解析を利用して判定される。図4Bの実施形態では、頭部の3次元場所を推定し、眼球の回転中心位置を導き出すために、顔の特徴が追跡される。追跡する特徴が多くなるほど、より信頼性のある注視検出となる。例えば、ユーザが自分の頭部をカメラから背けた場合、1つの眼球だけしかカメラで確認できない。鼻および口腔の位置を理解することによって、このシステムは、眼球の検出に失敗したとみな

10

20

30

40

50

すのではなく、ユーザが頭部を回したと判定する。ユーザの特徴の位置が判定されると、P O G が算出される。頭部追跡および顔追跡は、注視検出の精度を向上させるために使用することができる。一般に、頭部は、眼球と異なり、不安定に移動することもないので、頭部追跡は、注視追跡よりも正確である。

#### 【 0 0 4 5 】

ディスプレイに対するユーザの注視方向およびカメラの位置から、ディスプレイ上のどこでユーザの注視の焦点が合っているのかを決定することが可能である。カメラおよびディスプレイは、互いから離れて位置し得るので、ディスプレイの場所を決定するために、較正操作が行われ得る。ユーザは、ディスプレイに注目するように求められ、システムは、1つのディスプレイまたは複数のディスプレイの場所を推定する。別の実施形態において、較正操作は、ディスプレイの場所およびディスプレイ内の G U I オブジェクトの場所の推定をさらに向上させるために、ユーザを1つのディスプレイまたは複数のディスプレイの中の1つ以上の目標に注目させる。

#### 【 0 0 4 6 】

一実施形態では、画像解析に基づく注視検出を使用するときに、デスクトップ上的一群の注視位置が収集される。人間の視覚系の不安定な性質のため、これらの位置は、特定の領域内のランダムな一連の点のように見え得る。中心焦点位置は、統計解析を利用して P O G を決定するアルゴリズムを使用することによって、一群の位置から導き出され得る。ある期間にわたって複数の画像において得られる焦点位置は、不安定な注視移動をフィルタリングするために平均される。

10

#### 【 0 0 4 7 】

焦点位置は、「目標領域」の中心を画定する。目標領域のサイズも、一連の注視位置から、および以前の焦点位置から決定され得る。1つのアルゴリズムは、各注視位置から目標領域中心（焦点位置）までの距離を見出すことと、そのような距離の平均または合計のいずれかを使用して、目標領域の中心を画定するセントロイドを推定することとを含む。

20

#### 【 0 0 4 8 】

注視焦点の目標領域を使用して、アプリケーションは、より効果的に目標に到達するために、カーソルを移動させる方法を変更する。これは、上で説明したように、目標領域からの距離に基づいてマウスカーソル移動の速度を変更すること、カーソルを目標フォーカス領域に引き付けること、またはカーソルの位置を目標位置までジャンプさせることを含む。

30

#### 【 0 0 4 9 】

一実施形態において、追跡は、段階に分割され、頭部追跡によって定義される第1の段階と、注視検出によって定義される第2の段階とを伴う。頭部が移動しているときには、一般に注視追跡よりも頭部追跡の方が信頼性が高いので、目標領域を決定するために頭部追跡が使用される。頭部の運動が実質的に停止すると、目標領域の場所を微調整するために注視追跡が利用され、その後に、さらに微細な解像度を得て、カーソルをその行き先間で移動させるために、マウスの運動が続く。別の実施形態において、追跡の段階は、少なくとも、複数のディスプレイから1つのディスプレイを選択するために頭部追跡が使用される、第1の段階と、選択したディスプレイ内で目標を決定するために注視追跡が使用される、第2の段階とを含む。

40

#### 【 0 0 5 0 】

図5は、一実施形態による、注視検出を使用してハンドヘルドデバイスで高速にスクロールするための方法を例示する。携帯型デバイス506は、注視検出を行うために使用される、フロントフェーシングカメラを含む。しかしながら、スタンドアロンディスプレイの領域と比較すると、表示領域が小さいので、注視検出単独でディスプレイの中の小さいオブジェクトを選択することは困難である。しかしながら、本発明の実施形態は、画面の中または画面の外側の P O G を検出することによって、携帯型デバイス506を使用しながらナビゲーションを迅速化するために、ナビゲーション支援を提供する。

#### 【 0 0 5 1 】

50

ユーザ 502 は、コマンドまたはヒントをナビゲーション用のコンピュータプログラムに提供するために、携帯型デバイス 506 のディスプレイの外側の領域に注視 504 を方向付ける。例えば、ユーザ 502 が、長い文書を読んでいて、文書の一部分だけしか一度に表示されない状態である場合、ユーザは、ユーザがナビゲーションを迅速化したいことを示すために、別の入力（例えば、ボタンまたはタッチ画面）を使用しながら、画面の上、下、左、または右に注視を方向付ける。したがって、ユーザが、画面上の注視を保ちながら、タッチ画面を横断して指を下に摺動させた場合、プログラムは、表示を上にスクロールさせる。しかしながら、ユーザがディスプレイの上側に注視 504 を方向付けた場合、これは、ユーザが、文書の最上部に進みたいことを示すので、ユーザがディスプレイ上で指を摺動させ始めたときに、プログラムは、文書のいかなる部分が示されていても、文書の最上部にジャンプさせる。同様に、ユーザは、ディスプレイの下側を見て、ページの最下部へのジャンプを生じさせることができる。これは特に、電子メールの最下部に添付アイコンを示す、一部の電子メールクライアントに有用である。電子メールが長い場合、最下部のアイコンに辿り着くために、数多くの指のスクロールを要する。注視を使用することによって、最下部へのジャンプがあり、手間と時間の節約がもたらされる。10

#### 【0052】

別の実施形態において、注視支援ナビゲーションは、スマートフォンまたはタブレットデバイスで本または雑誌を読むユーザに提供される。ユーザがページをめくりたいときに、ユーザは、画面の左側または右側の向こうを見て、次いで、画面をタップしてページめくりを生じさせる。さらに別の実施形態において、ユーザは、画面をタップすることと併せて、ディスプレイの左縁部または右縁部を見るだけで、ユーザに画面の外側を見る要求することなく、ページめくりを生じさせる。タブレットデバイスは、より大きいディスプレイを有し、注視検出は、ユーザが左を見ているのか、または右を見ているのかを判定するのに十分正確であり得る。例えば、角膜解析は、瞳孔が左に向いているのか、または右に向いているのかを示す。20

#### 【0053】

図 6 は、一実施形態による、注視検出およびユーザ入力を使用する高速カーソル移動のための方法を例示する。一実施形態では、物理的行動と関連付けられる所定の入力が、補完的な入力としてコンピュータプログラムに注視を使用させる。例えば、ユーザ 602 は、タッチパッド 612、ディスプレイ 610、およびユーザフェーシングカメラを含む、ラップトップ 604 で作業している。タッチパッド上での 2 本指タップは、ユーザの注視に基づくマウスカーソルの移動を起動させる。30

#### 【0054】

図 6 の実施形態において、マウスカーソルは、位置 614 にある。ユーザが、ユーザの注視 608 を POG 620 に方向付けながら、タッチパッド 612 上で 2 本指タップを行うと、マウスカーソルは、POG 620 に近接するか、または正確に POG 620 である、位置 618 にジャンプする。マウスカーソルは、この時点でウィンドウ 616 の内側にあるので、ウィンドウ 616 が選択される。

#### 【0055】

図 6 で示される実施形態は、例示的なものであることに留意されたい。他の実施形態は、注視支援ナビゲーションを起動させるために、異なる入力（例えば、マウスのダブルクリック、タッチパッドの 2 回連続タップ、キーボード上でキーを押す等）を利用可能であり、または他の実施形態は、異なる方法（例えば、ウィンドウを選択する、ブラウザのアクティブなタブを変更する、ウィンドウをスクロールする等）でナビゲーションを支援し得る。したがって、図 6 で例示される実施形態は、排他的または限定的であると解釈されるべきではなく、例示的または実例であると解釈されるべきである。40

#### 【0056】

一実施形態において、GUI は、ユーザの注視が検出されたときに、ディスプレイ上にフィードバックを提供する。例えば、POG のほぼ中央にある領域がディスプレイの中で拡大される、GUI 上のいくつかのウィンドウが透明になる、アイコンもしくはウィンド50

ウが再配置される、またはポップアップメニューが表示される等である。

【0057】

図7は、一実施形態による、注視支援ゲームインターフェースを示す図である。ユーザ702は、注視検出に使用されるカメラ706に接続される、ゲームコンソール704上で実行するゲームを行っている。ユーザ702がコントローラ708上のボタンを押すと、ゲームは、ボタンが押されたときのユーザの注視に応じて、異なる行動で応答する。図7の実施形態では、敵キャラクタ714との戦いに関わるユーザが、ゲームキャラクタ712を制御している。ゲームは、ユーザのPOGに応じて、異なって反応する。

【0058】

ユーザのPOG716aがディスプレイの縁部またはディスプレイの外側にあるときに、コントローラ708上のXボタン710を押すと、ゲームに第1の行動（例えば、キャラクタ712が銃に再装弾する）を行わせる。ユーザのPOG716bがユーザ702によって制御されるキャラクタ712上にあるときに、Xボタン710を押すと、ゲームに第2の行動（例えば、ゲームキャラクタの健康値を増加させる）を行わせる。ユーザのPOG716cが敵のキャラクタ714上にあるときに、Xボタン710を押すと、ゲームに第3の行動（例えば、キャラクタ712がPOG716cの場所を撃つ）を行わせる。

10

【0059】

図7で示される実施形態は、例示的なものであることに留意されたい。他の実施形態は、異なるボタン、異なるタイプの入力デバイスを利用でき、またはゲームによって異なる反応（例えば、オブジェクトを選ぶ、ドアを開ける、キャラクタを移動させる等）を生じさせ得る。したがって、図7で例示される実施形態は、排他的または限定的であると解釈されるべきではなく、例示的または実例であると解釈されるべきである。

20

【0060】

図8は、一実施形態による、ジェスチャーに対応するウィンドウズ（登録商標）グラフィカルユーザインターフェースを例示する。GUIは、ユーザのGUIとの相互作用を高めるために、入力として注視およびジェスチャーを使用する。ユーザが所定のジェスチャーを行うと、ユーザの注視と関連付けられる、またはそれによって選択されるオブジェクトに対して行動が行われる。

【0061】

例えば、ユーザ814は、ディスプレイ802からディスプレイ806までウィンドウ810を移動させることを所望する。移動を行うために、ユーザは、ウィンドウ810に注目し、手818を上げ、そして、左から右に手818を移動させ816、ユーザ814が左から右にウィンドウ810を移動させることを所望することを示す。デスクトップが2つのディスプレイ上に表示されるので、左から右へのウィンドウの移動804は、ウィンドウ810をディスプレイ806の中の位置812にさせる結果となる。カメラ820からの画像は、注視検出に、およびジェスチャー検出に使用されるが、あらゆる注視検出方法またはジェスチャー検出方法を使用することもできる。この場合、ユーザの意図は、手でジェスチャーを行うという物理的行動で表される。

30

【0062】

別の実施形態では、ディスプレイ上に描画され、注視によって選択されるオブジェクトを操作するために、ユーザのジェスチャーが使用される。例えば、ゲーミング環境において、ユーザは、ゲームの中のオブジェクト（例えば、武器、財宝、コイン、はしご等）に注目し、そしてユーザが手を移動させると、ゲームは、オブジェクトに対して関連する行動を行うことによって、手の運動に応答する。例えば、ユーザは、注視によって地面のコインを選択でき、次いで、ユーザがそのコインを拾うかのように手を移動させると、結果として、アバター、すなわちユーザによって制御されるゲームキャラクタがコインを拾う。

40

【0063】

図9は、一実施形態による、注視検出、音声認識、および1つ以上のコンピュータ周辺機器を利用する、インターフェースを例示する。一実施形態では、ディスプレイの中の領

50

域またはオブジェクトを選択するために注視が使用され、ユーザ入力がオブジェクトの選択を示し（例えば、マウスを移動させる、またはキーを押す）、そして、オブジェクトに対して行動を行うために音声認識が使用される。

#### 【0064】

例えば、ユーザ912は、Windows（登録商標）環境で作業しており、デスクトップ上でフォルダアイコンとして描画されるフォルダ904を開くことを望んでいる。マウスは、フォルダ904から離れた、デスクトップの中の位置908に位置する。ユーザ912は、フォルダ904に注視を方向付けて、マウス916を移動させ914始める。マウスの運動が検出されると、システムは、ユーザの注視を確認し、POGが、フォルダアイコン904の近くに位置する、目標領域906であると判定する。システムは、次いで、フォルダ904上またはフォルダ904の近くにあるマウスカーソルを位置902まで移動またはジャンプさせる。カーソル位置がフォルダ上にない場合、ユーザは、フォルダ904の最上部にカーソルを位置付けるために、僅かなマウスの運動を行う。マウスがフォルダ904上にある時点で、ユーザは、音声認識プログラムによって検出されるコマンド（例えば、「開く」、「コピー」、「削除」、「移動」）を言う910。音声コマンドが検出されると、要求された操作が行われる（例えば、ウィンドウを開いて、フォルダ904の内容を表示する）。

#### 【0065】

図10は、本発明の実施形態を実現するためのコンピュータシステムの単純化された模式図である。本明細書で説明される方法は、従来の汎用コンピュータシステム等のデジタル処理システムで行われ得ることを認識されたい。1つの機能だけを行うように設計またはプログラムされた特殊用途コンピュータが代替的に使用され得る。コンピューティングデバイス1012は、メモリ1034に、永久記憶デバイス1058に、およびコンピューティングデバイス1012の内部の、またはそれに接続された他のモジュールに連結される、プロセッサ1032を含む。GUIコンピュータプログラム1036は、メモリ1034の中に存在するが、永久記憶デバイス1058の中にも存在し得る。

#### 【0066】

コンピューティングデバイス1012は、音声取り込みデバイス1008、画像取り込みデバイス1020、およびディスプレイ1026と通信する。一実施形態において、音声取り込みデバイス1008、画像取り込みデバイス1020、およびディスプレイ1026は、コンピューティングデバイス1012内に組み込み可能で、ビデオレコーダ等の1つ以上のデバイスに組み合わせられ得る。一実施形態において、音声取り込みデバイスはマイクロホンを含み、別の実施形態において、音声取り込みデバイスは、マイクロホンアレイを含む。

#### 【0067】

上述したように、注視目標1028は、ディスプレイ内に画定されるが、注視目標1028はまた、例えば、コンピューティングデバイス1012上、またはディスプレイ1028を囲む領域等の、ディスプレイの外側にも設定され得る。注視検出モジュール1024は、ユーザのPOGを決定し、ユーザ1002の注視1006が注視目標1028に向かって方向付けられているかどうかを判定する。音声側で、音声認識モジュール1016は、音声取り込みデバイス1008から受け取った音声1004を処理し、有効なユーザ音声コマンドが発行されたかどうかを判定するために、音声認識を行う。GUIコンピュータプログラム1036は、上で説明される本発明の実施形態を実現するために、注視検出1024、音声認識モジュール1016、および頭部追跡モジュール1038から入力を受け取る。

#### 【0068】

永久記憶デバイス1058は、ローカルまたはリモートであり得る、フロッピーディスク（登録商標）ドライブまたは固定ディスクドライブ等の永続的なデータ記憶デバイスを表す。ネットワークインターフェース1046は、他のデバイスとの通信を可能にする、ネットワーク接続性を提供する。プロセッサ1032は、汎用プロセッサ、特殊用途プロ

10

20

30

40

50

セッサ、特にプログラムされた論理デバイスで具現化され得ることを認識されたい。入力／出力（I／O）インターフェース1042は、ディスプレイ1026、キーボード1052、マウス1050、音声取り込みデバイス1008、画像取り込みデバイス1020、スピーカー1054、ボタン、センサ、タッチ画面1056等の、異なる周辺機器との通信を提供する。ユニバーサルシリアルバス（USB）モジュール1044は、USBデバイスへの接続性を提供する。

#### 【0069】

ディスプレイ1026は、本明細書で説明されるユーザインターフェースを表示するよう構成される。キーボード1052、マウス1050、および他の周辺機器は、プロセッサ1032に情報を通信するために、I／Oインターフェース1042に連結される。外部デバイスへの、およびそこからのデータは、I／Oインターフェース1042を通して通信され得ることを認識されたい。本発明はまた、有線または無線ネットワークを通してリンクされるリモート処理デバイスによってタスクが行われる、分散コンピューティング環境において実践することもできる。

#### 【0070】

図10で示される実施形態は、例示的なものであることに留意されたい。他の実施形態は、異なるモジュールを利用し得る、または1つのモジュールによって行われる複数の機能を有し得る等である。したがって、図10で例示される実施形態は、排他的または限定的であると解釈されるべきではなく、例示的または実例であると解釈されるべきである。

#### 【0071】

図11は、本発明の一実施形態に従って、ユーザをコンピューティングデバイスで実行するコンピュータプログラムとインターフェースさせるためのアルゴリズムを例示するフローチャートを示す。操作1102で、ユーザのPOGが識別され、操作1104で、ユーザがカーソルを移動させるための物理的行動を開始したかどうかを判定するための確認が行われる。ユーザが物理的行動を開始した場合、本方法は、操作1106に進み、ユーザが物理的行動を開始していなかった場合は、操作1102に戻る。

#### 【0072】

操作1106で、カーソルの現在位置とPOGとの間の距離が閾値距離よりも大きいかどうかを判定するための確認が行われる。距離が閾値距離よりも大きい場合、本方法は、操作1108に進み、そこでは、カーソルをPOGに近接した領域に移動させる。距離が閾値距離以下である場合、本方法は、操作1102へ戻る。

#### 【0073】

本発明の実施形態は、ハンドヘルドデバイス、マイクロプロセッサシステム、マイクロプロセッサに基づくまたはプログラム可能な家庭用電化製品、ミニコンピュータ、メインフレームコンピュータ等を含む、種々のコンピュータシステム構成で実践され得る。本発明はまた、ネットワークを通してリンクされるリモート処理デバイスによってタスクが行われる、分散コンピューティング環境において実践することもできる。

#### 【0074】

上記実施形態を考慮して、本発明は、コンピュータシステムに記憶されたデータを含む、種々のコンピュータにより実現される操作を利用できることを理解されたい。こうした操作は、物理量の物理的操作を必要とする操作である。本発明の一部を形成する本明細書で説明される操作はいずれも、有用な機械操作である。本発明はまた、これらの操作を行うためのデバイスまたは装置にも関する。この装置は、特殊用途コンピュータ等の、必要とされる目的のために特別に構築され得る。特殊用途コンピュータと定義されるときに、該コンピュータはまた、特殊用途の一部ではない他の処理ルーチン、プログラム実行ルーチン、またはルーチンを行うこともできるが、それでも、特殊用途のための操作を行うことができる。あるいは、操作は、コンピュータメモリもしくはキャッシュに記憶される、またはネットワークを通じて取得される1つ以上のコンピュータプログラムによって選択的に起動または構成される、汎用コンピュータによって処理され得る。データがネットワークを通じて取得されるときに、データは、ネットワーク上の他のコンピュータ

10

20

30

40

50

タ、例えばコンピューティングリソースのクラウドによって処理され得る。

【0075】

本発明の1つ以上の実施形態はまた、コンピュータが読み出し可能な媒体上のコンピュータが読み出し可能なコードとして製造することもできる。コンピュータが読み出し可能な媒体は、データを記憶することができ、その後に、コンピュータシステムによって読み出すことができる、任意のデータ記憶デバイスである。コンピュータが読み出し可能な媒体の例としては、ハードドライブ、ネットワーク接続記憶装置(NAS)、リードオンリーメモリ、ランダムアクセスメモリ、CD-ROM、CD-R、CD-RW、磁気テープ、ならびに他の光学式データ記憶デバイスおよび非光学式データ記憶デバイスが挙げられる。コンピュータが読み出し可能な媒体は、コンピュータが読み出し可能なコードが分散方式で記憶され、実行されるように、ネットワークに連結されたコンピュータシステムを通じて分散される、コンピュータが読み出し可能な有形の媒体を含むことができる。  
10

【0076】

本方法の操作を特定の順序で説明したが、オーバーレイ操作が所望の方法で行われるのであれば、他の準備操作が操作の間に行われてもよく、または操作が僅かな時間差で行われるように操作を調整してもよく、または処理と関連付けられる種々の間隔で処理操作を発生させることを可能にするシステムに操作を分散させてもよいことを理解されたい。

【0077】

上述の本発明は、理解を明確にする目的でいくらか詳細に記載されているが、添付の特許請求の範囲内で、特定の変更および修正を行うことができるようになるであろう。故に、本実施形態は、限定的なものではなく例示的なものであるとみなされるべきであり、また本発明は、本明細書に与えられる詳細に限定されるものではなく、添付の特許請求の範囲およびその均等物の範囲内で変更され得るものである。  
20

【図1】

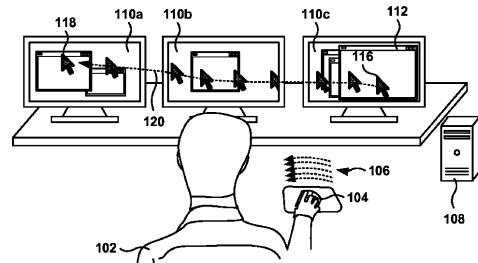


Fig. 1

【図2B】

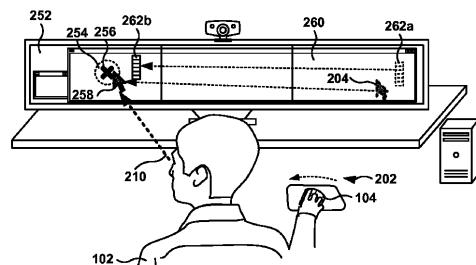


Fig. 2B

【図2A】

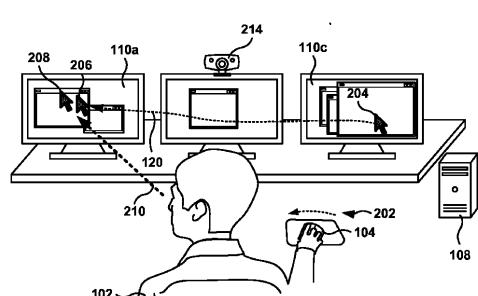
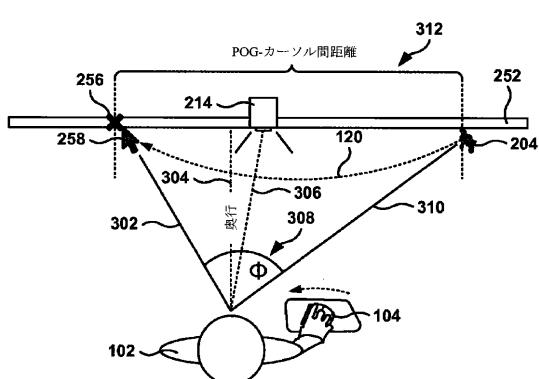
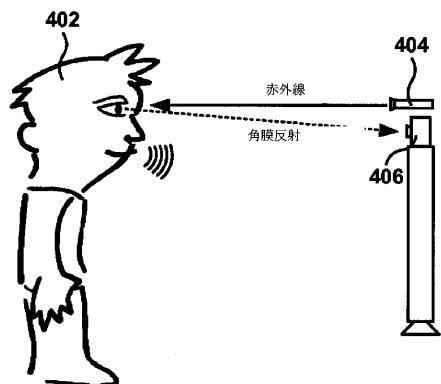


Fig. 2A

【図3】



【図4A】



【図4B】

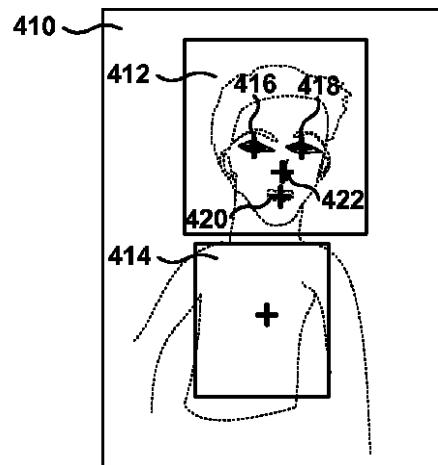


Fig. 4B

【図5】

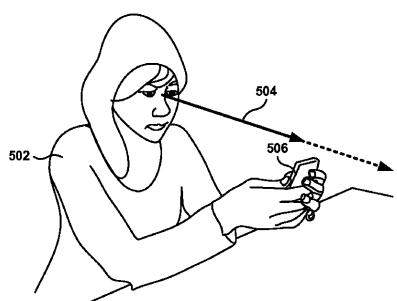


Fig. 5

【図7】

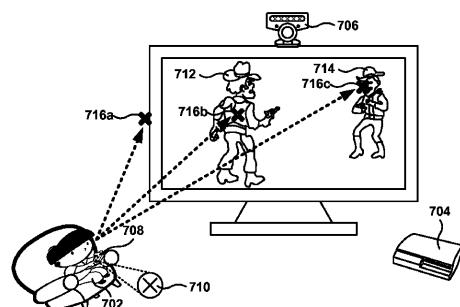


Fig. 7

【図6】

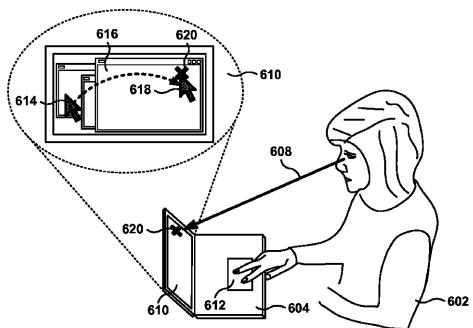


Fig. 6

【図8】

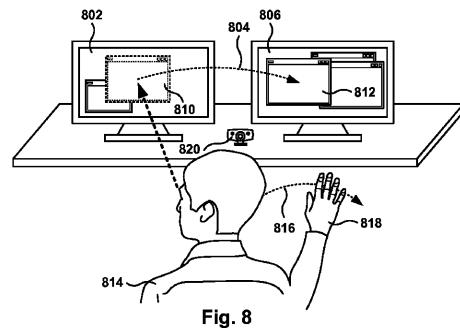
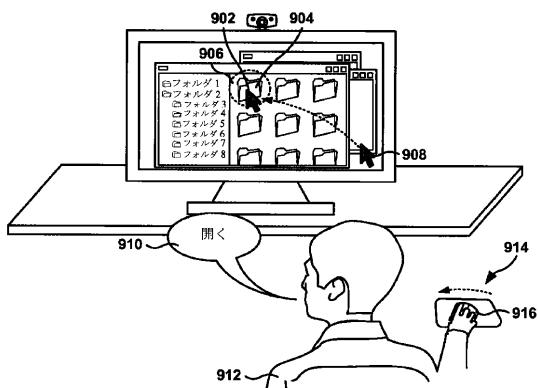
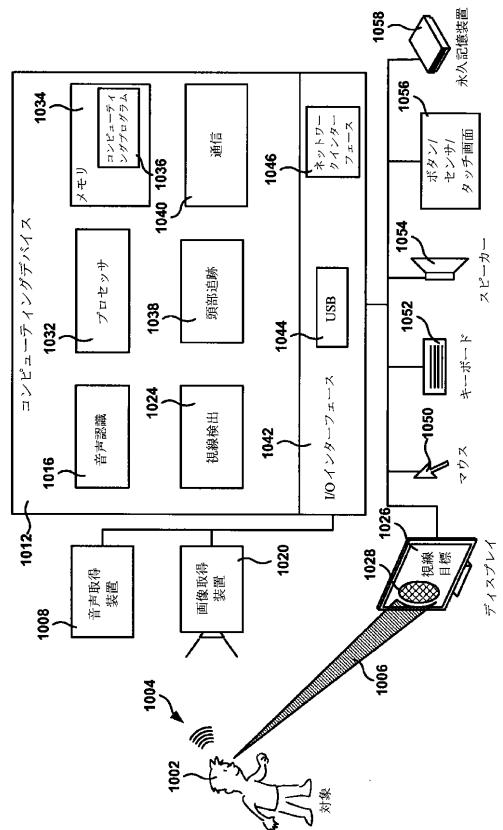


Fig. 8

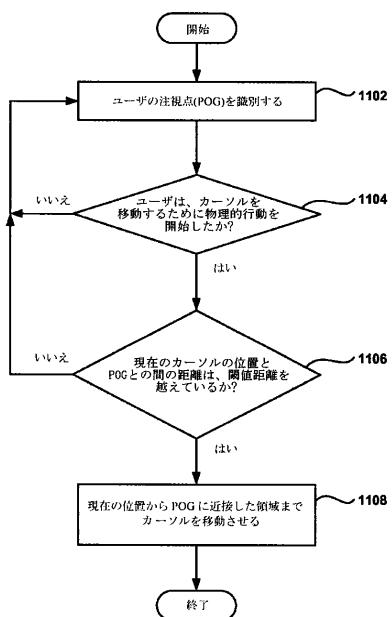
【図9】



【図10】



【図11】



---

フロントページの続き

(56)参考文献 英国特許出願公開第02440348(GB,A)  
米国特許第06204828(US,B1)  
特開2006-059147(JP,A)

(58)調査した分野(Int.Cl., DB名)

G 06 F 3 / 0346  
G 06 F 3 / 01  
G 06 F 3 / 0481