

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第4689548号  
(P4689548)

(45) 発行日 平成23年5月25日(2011.5.25)

(24) 登録日 平成23年2月25日(2011.2.25)

(51) Int.Cl.

F 1

G06F 3/033 (2006.01)  
G06F 3/048 (2006.01)G06F 3/033 310Y  
G06F 3/048 656A

請求項の数 9 (全 17 頁)

(21) 出願番号 特願2006-197392 (P2006-197392)  
 (22) 出願日 平成18年7月19日 (2006.7.19)  
 (62) 分割の表示 特願2004-23013 (P2004-23013)  
 原出願日 平成16年1月30日 (2004.1.30)  
 (65) 公開番号 特開2006-294066 (P2006-294066A)  
 (43) 公開日 平成18年10月26日 (2006.10.26)  
 審査請求日 平成18年12月6日 (2006.12.6)  
 審判番号 不服2009-4811 (P2009-4811/J1)  
 審判請求日 平成21年3月5日 (2009.3.5)

(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 大場 章男  
 東京都港区南青山二丁目6番21号 株式会社ソニー・コンピュータエンタテインメント内  
 (72) 発明者 鈴木 章  
 東京都港区南青山二丁目6番21号 株式会社ソニー・コンピュータエンタテインメント内

最終頁に続く

(54) 【発明の名称】画像処理装置、画像処理方法、記録媒体、コンピュータプログラム、半導体デバイス

## (57) 【特許請求の範囲】

## 【請求項 1】

所定のイベントに関連づけされたオブジェクトについてのオブジェクト画像を生成するオブジェクト画像生成手段と、

操作者の画像をその一部に含んだ鏡面動画像に前記オブジェクト画像を合成して、所定のディスプレイ装置に表示させるための合成画像を生成する画像合成手段と、

前記鏡面動画像に含まれる前記操作者の画像から当該操作者の顔の位置及び当該操作者の顔画像の大きさを検出して、検出した顔の位置に基づく顔領域を選択する検出手段と、を備え、

前記画像合成手段は、前記検出手段により選択された前記操作者の前記顔領域に接する位置に、前記操作者の顔画像の大きさに応じて大きさが設けられる複数の提示候補領域から前記オブジェクト画像を表示する領域を選択して、選択した提示候補領域に該提示候補領域の大きさに合わせた大きさで前記オブジェクト画像が表示されるように、前記オブジェクト画像と前記鏡面動画像とを合成するように構成されている、

画像処理装置。

## 【請求項 2】

前記画像合成手段は、前記検出手段により検出された前記操作者の顔の位置を矩形で囲うマーカを表す画像を合成するように構成されており、

前記検出手段は、前記マーカを表す画像に基づいて前記顔領域を選択する、

請求項1記載の画像処理装置。

10

20

**【請求項 3】**

前記鏡面動画像のフレーム間の画像の差異を表す差分画像を、複数重ねた動き画像マップを生成する手段をさらに備えており、

前記画像合成手段は、この動き画像マップにより得られる前記操作者の画像の動きに応じて、前記複数の提示候補領域から前記オブジェクト画像を表示する領域を選択する、

請求項1記載の画像処理装置。

**【請求項 4】**

前記鏡面動画像のフレーム間で各領域の色の変化を検出する手段を備えており、

前記画像合成手段は、この色の変化でにより得られる前記操作者の画像の動きに応じて、前記複数の提示候補領域から前記オブジェクト画像を表示する領域を選択する、

請求項1記載の画像処理装置。

10

**【請求項 5】**

所定の撮影装置により撮影された前記操作者の画像をその一部に含む動画像を、この撮影装置から取り込む画像取込手段と、

取り込んだ動画像に鏡面処理を施して前記鏡面動画像を生成する画像反転手段と、をさらに備えている、

請求項1記載の画像処理装置。

**【請求項 6】**

操作者の画像をその一部に含んだ鏡面動画像と、所定のイベントに関連づけされたオブジェクトについてのオブジェクト画像とを含んだ合成画像を生成する画像合成手段と、前記鏡面動画像に含まれる前記操作者の画像から当該操作者の顔の位置及び当該操作者の顔画像の大きさを検出して、検出した顔の位置に基づく顔領域を選択する検出手段とを有し、生成された合成画像を所定のディスプレイ装置に表示させる画像処理装置により実行される方法であって、

20

前記画像合成手段が、

前記検出手段により選択された前記操作者の前記顔領域に接する位置に、前記操作者の顔画像の大きさに応じて大きさが設けられる複数の提示候補領域から前記オブジェクト画像を表示する領域を選択して、選択した提示候補領域に該提示候補領域の大きさに合わせた大きさで前記オブジェクト画像が表示されるように前記オブジェクト画像と前記鏡面動画像とを合成する、

30

画像処理方法。

**【請求項 7】**

ディスプレイ装置が接続されたコンピュータに、

所定のイベントに関連づけされたオブジェクトについてのオブジェクト画像を生成するオブジェクト画像生成手段、

操作者の画像をその一部に含んだ鏡面動画像に前記オブジェクト画像を合成して合成画像を生成する画像合成手段、

生成された合成画像を所定のディスプレイ装置に表示させる表示制御手段、

前記鏡面動画像に含まれる前記操作者の画像から当該操作者の顔の位置及び当該操作者の顔画像の大きさを検出して、検出した顔の位置に基づく顔領域を選択する検出手段、を形成させ、

40

前記画像合成手段に、

前記検出手段により選択された前記操作者の前記顔領域に接する位置に、前記操作者の顔画像の大きさに応じて大きさが設けられる複数の提示候補領域から前記オブジェクト画像を表示する領域を選択して、選択した提示候補領域に該提示候補領域の大きさに合わせた大きさで前記オブジェクト画像が表示されるように、前記オブジェクト画像と前記鏡面動画像とを合成する機能を形成させるためのコンピュータプログラム。

**【請求項 8】**

請求項7に記載されたコンピュータプログラムを記録してなる、コンピュータ読み取り可能な記録媒体。

50

**【請求項 9】**

ディスプレイ装置が接続されたコンピュータに搭載された装置に組み込まれることにより、前記コンピュータに、

所定のイベントに関連づけされたオブジェクトについてのオブジェクト画像を生成するオブジェクト画像生成手段、

操作者の画像をその一部に含んだ鏡面動画像に前記オブジェクト画像を合成して合成画像を生成する画像合成手段、

生成された合成画像を所定のディスプレイ装置に表示させる表示制御手段、

前記鏡面動画像に含まれる前記操作者の画像から当該操作者の顔の位置及び当該操作者の顔画像の大きさを検出して、検出した顔の位置に基づく顔領域を選択する検出手段、を形成させ、

前記画像合成手段に、

前記検出手段により選択された前記操作者の前記顔領域に接する位置に、前記操作者の顔画像の大きさに応じて大きさが設けられる複数の提示候補領域から前記オブジェクト画像を表示する領域を選択して、選択した提示候補領域に該提示候補領域の大きさに合わせた大きさで前記オブジェクト画像が表示されるように、前記オブジェクト画像と前記鏡面動画像とを合成する機能を形成させる半導体デバイス。

**【発明の詳細な説明】****【技術分野】****【0001】**

本発明は、ビデオカメラなどの撮影装置により撮影された撮影画像を、コマンド等の入力インターフェースとして利用するための画像処理技術に関する。

**【背景技術】****【0002】**

コンピュータ、ビデオゲーム機などによく用いられる入力装置として、キーボード、マウス、コントローラ等がある。操作者は、これらの入力装置を操作することにより所望のコマンドを入力して、入力されたコマンドに応じた処理をコンピュータ等に行わせる。そして操作者は、処理結果として得られた画像、音などを、ディスプレイ装置やスピーカにより視聴する。操作者は、入力装置に備えられる多くのボタンを操作したり、ディスプレイ装置に表示されたカーソルなどを見ながら操作することにより、コマンドの入力をすることとなる。

**【0003】**

最近は、上記のような従来からの入力装置を用いる方法の他に、ビデオカメラ等の撮影装置を用いて撮影した動画像により、コマンドの入力を可能にした技術も開発されている。本出願人も、特開2002-196855号公報（発明の名称：画像処理装置、画像処理方法、記録媒体、コンピュータプログラム、半導体デバイス）において、操作者の動画像を取り込んで、操作者の動画像と当該動画像により操作されるオブジェクト画像とを合成して所定のディスプレイ装置に表示させる技術を開示している。動画像の動きに応じてオブジェクト画像が操作されるために、動画像自体が入力インターフェースとして機能することになる。

**【0004】**

このように動画像を入力インターフェースとして用いる場合、操作者が適切な位置に表示されるように撮影装置のアングルを調整したり、画角を選択することは、正確な操作のために極めて重要である。そのために、初期設定が煩雑、複雑になりがちである。設定が不十分であったり、操作者が適切な位置にいない場合には、オブジェクト画像を操作困難な位置に操作者の動画像が表示されたり、或いは誤認識の原因となり、正確な入力動作の妨げになることもある。

**【発明の開示】****【発明が解決しようとする課題】****【0005】**

10

20

30

40

50

本発明は、上記のような問題を解決するものであり、動画像を入力インターフェースとして用いた場合の初期設定を容易に行うための画像処理技術を提供することを課題とする。

【課題を解決するための手段】

【0006】

以上の課題を解決する本発明の画像処理装置は、所定のイベントに関連づけされたオブジェクトについてのオブジェクト画像を生成するオブジェクト画像生成手段と、操作者の画像をその一部に含んだ鏡面動画像に前記オブジェクト画像を合成して合成画像を生成する画像合成手段を有し、この合成画像を所定のディスプレイ装置に表示させる画像処理装置であって、前記鏡面動画像に含まれる前記操作者の画像の位置を検出する検出手段を備え、前記画像合成手段は、前記検出手段により検出された前記操作者の画像の位置に応じて、前記操作者の手の画像が届く領域に前記オブジェクト画像が表示されるように、前記オブジェクト画像と前記鏡面動画像とを合成するように構成されている。10

操作者の画像の位置に合わせて、操作者の手の画像が届く領域にオブジェクト画像が表示されるために、操作者が操作不能な位置にオブジェクト画像が表示されることがなくなる。そのために、煩雑な初期設定を行わなくても、動画像を入力インターフェースとして容易にしようできるようになる。

【0007】

鏡面動画像は動画像を撮影する撮影装置側で生成するようにしてもよいが、例えば、本発明の画像処理装置が、所定の撮影装置により撮影された前記操作者の画像をその一部に含む動画像を、この撮影装置から取り込む画像取込手段と、取り込んだ動画像に鏡面処理を施して前記鏡面動画像を生成する画像反転手段と、をさらに備えるようにすれば、汎用の撮影装置で鏡面動画像を得ることができるようになる。20

【0008】

操作者の手の画像が届く領域とは、例えば操作者の顔の位置の周辺部分とすることができる。顔の位置の周辺部分にオブジェクト画像を表示させる場合には、例えば、前記検出手段が、前記操作者の画像から当該操作者の顔の位置を検出するように構成され、前記画像合成手段が、検出された顔の位置に応じた領域に前記オブジェクト画像が表示されるように、前記オブジェクト画像と前記鏡面動画像とを合成するように構成される。この場合、前記画像合成手段は、前記検出手段により検出された前記操作者の画像の位置に、当該操作者が検出されたことを示すマーカを表す画像を合成するように構成されていてもよい。30

【0009】

前記検出手段は、操作者の画像の位置の他に前記操作者の画像の大きさを検出するように構成されていてもよい。この場合、前記画像合成手段が、検出された前記操作者の画像の大きさに応じた大きさに前記オブジェクト画像が表示されるように、前記オブジェクト画像と前記鏡面動画像とを合成するように構成される。これにより、操作者の画像の大きさに応じた大きさでオブジェクト画像を表示できるようになり、操作者の画像とオブジェクトの画像とがアンバランスに表示されることはない。

【0010】

前記鏡面動画像のフレーム間の画像の差異を表す差分画像を、複数重ねた動き画像マップを生成する手段をさらに備える場合には、前記画像合成手段は、この動き画像マップにより得られる前記操作者の画像の動きに応じて、前記オブジェクト画像を合成する領域を決めるように構成されていてもよい。40

また、前記鏡面動画像のフレーム間で各領域の色の変化を検出する手段を備える場合には、前記画像合成手段は、この色の変化でにより得られる前記操作者の画像の動きに応じて、前記オブジェクト画像を合成する領域を決めるように構成されていてもよい。

【0011】

操作者になり得る複数の操作者候補の画像が前記鏡面動画像に含まれている場合には、前記検出手段は、複数の前記操作者候補のそれぞれについて操作者候補の画像の位置を検出するように構成され、前記画像合成手段は、前記検出手段により検出された複数の前記50

操作者候補の画像の各々に位置に応じて、前記オブジェクト画像が複数の操作者候補の手の画像が届く範囲に表示されるように、前記オブジェクト画像と前記鏡面動画像とを合成するように構成されていてもよい。

つまり各操作者候補の画像がオブジェクト画像を操作できるような位置に、オブジェクト画像が合成される。例えば、操作者候補が2人の場合にはその間にオブジェクト画像が表示されるように合成される。

#### 【0012】

複数の操作者候補から操作者を選ぶ場合には、以下のような構成となる。

例えば、前記オブジェクト画像が操作された場合に、当該オブジェクト画像を操作した操作者候補の画像を前記操作者の画像に選択する手段を備えるようにする。オブジェクトを最初に操作した操作者候補が操作者になる。10

他に、前記鏡面動画像が複数の異なる角度から同時に撮影されたステレオ画像である場合には、ステレオ画像を構成する複数の画像のそれぞれで前記オブジェクト画像に最も近い位置に表示される操作者候補の画像を前記操作者の画像に選択する手段を備えるようにする。オブジェクト画像に最も近い位置に表示される操作者候補が操作者になる。

他に、前記オブジェクト画像生成手段を、操作者候補の各々に対応して複数のオブジェクト画像を生成するように構成して、前記画像合成手段は、各操作者候補の画像に対して、対応するオブジェクト画像を当該操作者候補の手の画像が届く範囲に表示されるように、前記複数のオブジェクト画像と前記鏡面動画像とを合成するように構成してもよい。この場合、複数のオブジェクト画像のいずれか一つが操作された場合に、当該オブジェクト画像に対応する操作者候補の画像を前記操作者の画像に選択する。最初に、自分に割り当てられたオブジェクト画像を操作した者が操作者となる。20

#### 【0013】

複数の操作者候補から操作者が選ばれると、前記表示制御手段は、デジタル処理又は前記鏡面動画像を撮影するための撮影装置を制御して、選択された操作者候補の画像を、ズーム、パン、及び／又はチルトにより、画面の真ん中に大きく表示されるようにして、操作者を明確にするようにしてもよい。

#### 【0014】

本発明は、また、以下のような画像処理方法を提供する。この画像処理方法は、操作者の画像をその一部に含んだ鏡面動画像と、所定のイベントに関連づけされたオブジェクトについてのオブジェクト画像とを含んだ合成画像を生成する画像合成手段と、前記鏡面動画像に含まれる前記操作者の画像の位置を検出する検出手段とを有し、生成された合成画像を所定のディスプレイ装置に表示させる画像処理装置により実行される方法であって、前記画像合成手段が、前記検出手段により検出された前記操作者の画像の位置に応じて、前記操作者の手の画像が届く範囲に前記オブジェクト画像が表示されるように前記オブジェクト画像と前記鏡面動画像とを合成する。30

#### 【0015】

本発明は、また、以下のようなコンピュータプログラムを提供する。このコンピュータプログラムは、ディスプレイ装置が接続されたコンピュータに、所定のイベントに関連づけされたオブジェクトについてのオブジェクト画像を生成するオブジェクト画像生成手段、操作者の画像をその一部に含んだ鏡面動画像に前記オブジェクト画像を合成して合成画像を生成する画像合成手段、生成された合成画像を所定のディスプレイ装置に表示させる表示制御手段、前記鏡面動画像に含まれる前記操作者の画像の位置を検出する検出手段、を形成させ、前記画像合成手段に、前記検出手段により検出された前記操作者の画像の位置に応じて、前記操作者の手の画像が届く範囲に前記オブジェクト画像が表示されるように、前記オブジェクト画像と前記鏡面動画像とを合成する機能を形成させるためのコンピュータプログラムである。40

#### 【0016】

本発明は、また、以下のような半導体デバイスを提供する。この半導体デバイスは、ディスプレイ装置が接続されたコンピュータに搭載された装置に組み込まれることにより、50

前記コンピュータに、所定のイベントに関連づけされたオブジェクトについてのオブジェクト画像を生成するオブジェクト画像生成手段、操作者の画像をその一部に含んだ鏡面動画像に前記オブジェクト画像を合成して合成画像を生成する画像合成手段、生成された合成画像を所定のディスプレイ装置に表示させる表示制御手段、前記鏡面動画像に含まれる前記操作者の画像の位置を検出する検出手段、を形成させ、前記画像合成手段に、前記検出手段により検出された前記操作者の画像の位置に応じて、前記操作者の手の画像が届く範囲に前記オブジェクト画像が表示されるように、前記オブジェクト画像と前記鏡面動画像とを合成する機能を形成させる。

**【発明の効果】**

**【0017】**

10

以上の説明から明らかなように、本発明によれば、動画像を入力インターフェースとして用いた場合の初期設定を、容易に行うことができるようになる。

**【発明を実施するための最良の形態】**

**【0018】**

以下、本発明の実施形態を詳細に説明する。

図1は、本発明を適用した画像処理システムの構成例を示した図である。

この画像処理システムは、撮影装置の一例であるアナログ又はデジタルのビデオカメラ1、画像処理装置2、ディスプレイ装置3、及びビデオカメラ制御装置4を有して構成される。画像処理システムは、ディスプレイ装置3に相対する操作者候補をビデオカメラ1で撮影し、これにより得られた動画像を画像処理装置2に時系列的に連続に取り込んで鏡面動画像を生成するとともに、この鏡面動画像と、メニュー・カーソル等のオブジェクトについてのオブジェクト画像とを合成して合成画像（これも動画像となる）を生成し、この合成画像をディスプレイ装置3上にリアルタイムに表示させるものである。オブジェクト画像は、鏡面動画像内の操作者候補が操作可能な範囲に表示されるように合成される。

20

オブジェクトには所定の処理が対応付けられており、操作者候補から選択された操作者によりオブジェクト画像が操作されると、該当する処理（イベント）が実行されるようになっている。

**【0019】**

鏡面動画像は、ビデオカメラ1から取り込んだ動画像を画像処理装置2で鏡面処理（画像の左右反転処理）することにより生成することができるが、ビデオカメラ1の前に鏡を置き、操作者候補を映した鏡面の動画像をビデオカメラ1で撮影することによって鏡面動画像を生成するようにしてもよい。また、ビデオカメラ1が鏡面動画像を生成するような機能を予め備えていてもよい。いずれにしても、ディスプレイ装置3上には、操作者の動きに応じてその表示形態がリアルタイムに変化する鏡面動画像とオブジェクト画像との合成画像が表示されるようにする。

30

なお、ビデオカメラ1を制御するためのビデオカメラ制御装置4は、画像処理装置4からの指示により、ビデオカメラ1にズーム、パン、チルト等の動作をさせるものであるが、ビデオカメラ1或いは画像処理装置4に内蔵されていてもよい。

**【0020】**

画像処理装置2は、コンピュータプログラムにより所要の機能を形成するコンピュータにより実現される。

40

この実施形態によるコンピュータは、例えば図2にそのハードウェア構成を示すように、それぞれ固有の機能を有する複数の半導体デバイスが接続されたメインバスB1とサブバスB2の2つのバスを有している。これらのバスB1、B2は、バスインターフェースINTを介して互いに接続され又は切り離されるようになっている。

**【0021】**

メインバスB1には、主たる半導体デバイスであるメインCPU10と、RAMで構成されるメインメモリ11と、メインDMAC(Direct Memory Access Controller)12と、MPEG(Moving Picture Experts Group)デコーダ(MDEC)13と、描画用メモリとなるフレームメモリ15を内蔵する描画処理装置(Graphic Processing Unit)、以

50

下、「G P U」) 1 4 が接続される。G P U 1 4 には、フレームメモリ 1 5 に描画されたデータをディスプレイ装置 3 で表示できるようにするためのビデオ信号を生成する C R T C (CRT Controller) 1 6 が接続される。

#### 【0022】

メイン C P U 1 0 は、コンピュータの起動時にサブバス B 2 上の R O M 2 3 から、バスインターフェース I N T を介して起動プログラムを読み込み、その起動プログラムを実行してオペレーティングシステムを動作させる。また、メディアドライブ 2 7 を制御するとともに、このメディアドライブ 2 7 に装着されたメディア 2 8 からアプリケーションプログラムやデータを読み出し、これをメインメモリ 1 1 に記憶させる。さらに、メディア 2 8 から読み出した各種データ、例えば複数の基本図形(ポリゴン)で構成された 3 次元オブジェクトデータ(ポリゴンの頂点(代表点)の座標値など)に対して、オブジェクトの形状や動き等を表現するためのジオメトリ処理(座標値演算処理)を行い、そして、ジオメトリ処理によるポリゴン定義情報(使用するポリゴンの形状及びその描画位置、ポリゴンを構成する素材の種類、色調、質感等の指定)をその内容とするディスプレイリストを生成する。10

#### 【0023】

G P U 1 4 は、描画コンテクスト(ポリゴン素材を含む描画用のデータ)を保持しており、メイン C P U 1 0 から通知されるディスプレイリストに従って必要な描画コンテクストを読み出してレンダリング処理(描画処理)を行い、フレームメモリ 1 5 にポリゴンを描画する機能を有する半導体デバイスである。フレームメモリ 1 5 は、これをテクスチャメモリとしても使用できる。そのため、フレームメモリ上のピクセルイメージをテクスチャとして、描画するポリゴンに貼り付けることができる。20

#### 【0024】

メイン D M A C 1 2 は、メインバス B 1 に接続されている各回路を対象として D M A 転送制御を行うとともに、バスインターフェース I N T の状態に応じて、サブバス B 2 に接続されている各回路を対象として D M A 転送制御を行う半導体デバイスであり、M D E C 1 3 は、メイン C P U 1 0 と並列に動作し、M P E G (Moving Picture Experts Group) 方式あるいは J P E G (Joint Photographic Experts Group) 方式等で圧縮されたデータを伸張する機能を有する半導体デバイスである。30

#### 【0025】

サブバス B 2 には、マイクロプロセッサなどで構成されるサブ C P U 2 0 、R A M で構成されるサブメモリ 2 1 、サブ D M A C 2 2 、オペレーティングシステムなどの制御プログラムが記憶されている R O M 2 3 、サウンドメモリ 2 5 に蓄積された音データを読み出してオーディオ出力として出力する音声処理用半導体デバイス(S P U (Sound Processing Unit)) 2 4 、図示しないネットワークを介して外部装置と情報の送受信を行う通信制御部(A T M) 2 6 、C D - R O M や D V D - R O M などのメディア 2 8 を装着するためのメディアドライブ 2 7 及び入力部 3 1 が接続されている。

#### 【0026】

サブ C P U 2 0 は、R O M 2 3 に記憶されている制御プログラムに従って各種動作を行う。サブ D M A C 2 2 は、バスインターフェース I N T がメインバス B 1 とサブバス B 2 を切り離している状態においてのみ、サブバス B 2 に接続されている各回路を対象として D M A 転送などの制御を行う半導体デバイスである。入力部 3 1 は、操作装置 3 5 からの入力信号が入力される接続端子 3 2 、ビデオカメラ 1 からの画像信号が入力される接続端子 3 3 、及びビデオカメラ 1 からの音声信号が入力される接続端子 3 4 を備える。40

なお、本明細書では、画像についてのみ説明を行い、便宜上、音声についての説明は省略する。

#### 【0027】

このように構成されるコンピュータは、メイン C P U 1 0 、サブ C P U 2 0 、G P U 1 4 が、R O M 2 3 及びメディア 2 8 等の記録媒体から所要のコンピュータプログラムを読み込んで実行することにより、画像処理装置 2 として動作するうえで必要な機能プロック50

、すなわち図3に示す、画像入力部101、画像反転部102、オブジェクトデータ記憶部103、オブジェクトデータ取得部104、オブジェクト画像生成部105、画像合成部106、画像比較部107、表示制御部108、及び操作者検出部109を形成する。

#### 【0028】

図2に示したハードウェアとの関係では、画像入力部101は入力部31及びその動作を制御するサブCPU20により形成され、画像反転部102、オブジェクトデータ取得部104、オブジェクト画像生成部105、画像比較部107、及び操作者検出部109はメインCPU10により形成され、画像合成部106はGPU14により形成され、表示制御部108はGPU14とCRTC16との協働により形成される。オブジェクトデータ記憶部103は、メインCPU10がアクセス可能なメモリ領域、例えばメインメモリ11に形成される。10

#### 【0029】

画像入力部101は、ビデオカメラ1により撮影された撮影画像を入力部31の接続端子33を介して取り込む。入力される撮影画像がデジタル画像の場合は、そのまま取り込む。入力される撮影画像がアナログ画像の場合は、A/D変換を行ってデジタル画像に変換して取り込む。

画像反転部102は、画像入力部101により取り込んだ撮影画像を鏡面処理、すなわち左右反転処理して鏡面動画像を生成する。なお、ビデオカメラ1による撮影を、ビデオカメラ1の前に鏡を置いて行う場合、或いはビデオカメラ1に鏡面画像を生成するような機能が備わっている場合のように、画像処理装置2に入力される撮影画像が既に鏡面動画像となっている場合には、画像反転部102は不要になる。20

#### 【0030】

オブジェクトデータ記憶部103は、メニュー画像（サブメニューを含む）、カーソル画像等のオブジェクト画像を表現するためのオブジェクトデータをその識別データと共に保持する。

オブジェクトデータ取得部104は、オブジェクトデータ記憶部103からオブジェクトデータを取り込んで、オブジェクト画像生成部105へ送る。

#### 【0031】

オブジェクト画像生成部105は、オブジェクトデータ取得部104より取り込んだオブジェクトデータに基づいてオブジェクト画像を生成する。オブジェクト画像生成部105は、プログラムまたは操作者による操作に基づいてオブジェクトの表示状態を決定し、その表示状態を実現するためのオブジェクト画像を生成する。30

#### 【0032】

画像合成部106は、鏡面動画像とオブジェクト画像生成部105により生成されるオブジェクト画像とを合成した合成画像をフレームメモリ15に描画する。なお、オブジェクト画像を合成して合成画像を生成するほかに、公知のインポーズ処理により、鏡面動画像上にオブジェクト画像を表示するようにしてもよい。

#### 【0033】

画像比較部107は、鏡面動画像を1フレーム毎に比較して、前後のフレームの鏡面動画像間の差分画像を生成する。また、画像比較部107は、鏡面動画像を1フレーム毎に比較して、フレーム間で鏡面動画像の各領域の色の変化を検出する。40

#### 【0034】

差分画像は、鏡面動画像に含まれる操作者（又は操作者候補）のフレーム毎の動きの変化を表す画像である。例えば、2つの鏡面動画像間で、操作者（又は操作者候補）が動いたときの、動く前の操作者（又は操作者候補）の画像と動いた後の操作者（又は操作者候補）の画像との異なる部分からなる画像である。この差分画像を複数重ねることで、操作者（又は操作者候補）の動く範囲、その頻度がわかるようになる。この頻度に応じてオブジェクト画像をどの位置に表示させるかを決めることができる。

#### 【0035】

鏡面動画像の各領域の色の変化を検出することで、操作者（又は操作者候補）が動いた50

ことがわかる。例えば、服の前に手をかざすと、画面上、手をかざした部分の色が服の色から手の色に変わる。これにより、色の変化で手が服の前にかざされたことがわかる。

これらの操作者（又は操作者候補）の動きの頻度、色の変化が、オブジェクト画像を鏡面動画像のどの領域に合成するかを決める要因になる。

#### 【0036】

操作者検出部109は、鏡面動画像内の操作者候補の顔の位置、大きさ等を検出するものであり、例えば、既知の顔センシング技術を用いて、操作者候補の顔の位置、大きさ等を検出する。顔センシング技術には、例えば、サポートベクターマシン（SVM）、ブースティング（boosting）、ニューラルネットワーク、固有顔法（Eigen Face）などを用いることができる。また、操作者検出部109は、複数の操作者候補から一人の操作者を選択するものである。操作者の選択は、例えば操作者候補がメニューなどのオブジェクト画像を操作したときに決まる。10

#### 【0037】

表示制御部108は、画像合成部106で生成された合成画像をビデオ信号に変換してディスプレイ装置3に出力する。また、表示制御部108は、操作者検出部109により選択された操作者に対してズーム、パン、チルト等を行うものである。ズーム、パン、チルト等は、フレームメモリに描画された画像に対してデジタル処理により行ってもよいし、ビデオカメラ制御装置4により、ビデオカメラ1のズーム、パン、チルト等を行わせるようにしてもよい。

ディスプレイ装置3は、この表示制御部108からのビデオ信号により、画面上に合成画像（動画像）を表示させる。20

#### 【0038】

##### <画像処理方法>

次に、上記のような画像処理システムにより行われる画像処理方法の実施例を説明する。。

#### 【0039】

##### [実施例1]

図4は、上記の画像処理システムを用いた本発明の画像処理方法の手順を説明するためのフローチャートである。

ディスプレイ装置3には、図1に示すように、2人の操作者候補を含む鏡面動画像が映し出されているものとする。なお、説明の便宜上ディスプレイ装置3に映し出される操作者候補の数を2人としたが、これは更に多くの人数でもよく、また、1人でもよい。1人の場合には、その操作者候補が操作者となる。30

鏡面動画像は、ビデオカメラ1で撮影された動画像が、画像処理装置2に入力され、画像反転部102により画像が左右反転されることで生成される。左右反転した画像がディスプレイ装置に表示されることで、図1のような鏡面動画像が映し出される。

#### 【0040】

画像処理装置2は、操作者検出部109により鏡面動画像から顔センシング技術を用いて操作者候補の顔の位置を検出して、検出した位置を矩形のマーカ画像51、52で囲う（ステップS101）。マーカ画像51、52もオブジェクト画像の一種であり、画像合成部106は、操作者検出部109により操作者候補の顔の位置が検出されると、オブジェクト画像生成部105、オブジェクトデータ取得部104を介してオブジェクトデータ記憶部103からマーカ画像51、52のオブジェクト画像を取得し、これを鏡面動画像の該当する位置に合成する。表示制御部108は、このマーカ画像51、52が合成された鏡面動画像をディスプレイ装置3に表示させる（図5）。40

なお、マーカ画像51、52は、矩形の他に、図6に示すような帯状として、操作者候補がヘアバンドをしているように表示してもよい。

#### 【0041】

次いで操作者検出部109は、マーカ画像51、52に基づいて、鏡面動画像から操作者候補の顔領域を選択する（ステップS102）。顔領域を選択することで、図7のよう50

なメニュー画像の提示候補領域が決められる。図7の例では、メニュー画像の提示候補領域が、予め顔領域の近傍に決められるようになっている。提示候補領域内の数字は、メニュー画像が提示される優先度を表しており、顔領域の下部（胸の辺り）の候補領域は優先度“1”であり、顔領域の右側近傍の候補領域は、優先度“2”で右手で操作できる領域である。顔領域の左側近傍の候補領域は、優先度“3”で左手で操作できる領域であり、顔領域の上部の候補領域は、優先度“4”で左右どちらの手でも操作できる領域である。2つの顔領域の中間の候補領域は、優先度“5”でどちらの操作者候補も操作できる領域である。

各提示候補領域は、優先度に応じてどの操作者候補でも同じ大きさに設定するようにしてもよいが、図7では、操作候補者毎に異なる大きさで設定している。図7では、左側の操作者候補が右側の操作者候補よりも大きく表示されているので、左側の操作者候補に対する提示候補領域が、右側の操作者候補の提示候補領域よりも大きく設定されている。つまり、提示候補領域の大きさは、操作者候補の顔領域の大きさに応じて変えるようにしてある。操作者候補の大きさは、操作者検出部109により顔の大きさを検出することで知ることができる。10

#### 【0042】

画像合成部106は、これらのメニュー画像の提示候補領域の中から実際にメニュー画像を提示する領域を選択する（ステップS103）。提示領域の選択は、例えば、画像比較部107により検出される、鏡面動画像の各部分における操作者候補の動き頻度、鏡面動画像の各領域の色の変化に応じて行うことができる。提示候補領域のうち、操作者候補の動きの頻度、色の変化により操作者候補が最も操作しやすい領域を選択する。この領域がメニュー画像を提示する領域となる。20

#### 【0043】

前述のように、操作者候補の動きの頻度は差分画像を重ねることで検出することができる。図8は、差分画像を重ねて得られる動き画像マップの例示図である。このような動き画像マップにより、操作者候補の動きの頻度が密な領域と疎な領域とがわかる。図8の例では、手の頻繁に動く領域が、動きが密な領域として斜線で表されている。

色の変化は、図9に示すようなカラー領域マップを用いることで知ることができる。図9のカラー領域マップは、顔の色と同じ色が占める領域を斜線で示しており、斜線の領域以外で、顔の色と同じ色を検出することで、操作者候補が動いたことを検出できる。カラー領域マップは、例えば、顔センシング技術により操作者の顔の位置を検出する際に、顔の色を検出しておき、検出した顔の色と同系統の色の領域を鏡面動画像から検出することで生成することができる。30

#### 【0044】

メニュー画像の提示領域が選択されると、画像合成部106は、鏡面動画像の該当領域にメニュー画像を合成した合成画像を生成する。表示制御部108は、ディスプレイ装置3に生成された合成画像を表示させる（ステップS104）。図10は、メニュー画像を提示したときの、ディスプレイ装置3に表示される画像の例示図である。この例では、2人の操作者候補のそれぞれ顔の近傍、優先度“2”的領域に、メニュー画像が提示されている。40

図10では2つのメニュー画像の大きさを同じにしているが、操作者候補ごとに異なる大きさで提示するようにしてもよい。例えば、ディスプレイ装置3に映し出される操作者候補の大きさに応じてメニュー画像の大きさを変える。大きく映し出される操作者候補の提示候補領域ほど大きく設定されるので、提示候補領域から選択される提示領域も、大きく映し出される操作者候補ほど大きくなる。提示領域の大きさに合わせてメニュー画像を提示するようにすれば、大きく映し出される操作者候補ほどメニュー画像を大きく提示することができる。

#### 【0045】

2人の操作者候補の一方が手を動かし、合成画像上で当該操作者候補の手がメニュー画像に触れる、或いは操作すると、操作者検出部109により当該操作者候補が操作者とし50

て選択される（ステップS105）。これにより実際に操作を行う操作者が決まる。

実際に操作を行う操作者が決まると、表示制御部108は、図11に示すように、選択した操作者を拡大して、画面中央に表示する（ステップS106）。また、メニュー画像は操作者に操作されているので、その下位層の「select1」、「select2」、「select3」のいずれかを表すプルダウン画像が表示される。

このようにして、操作者候補の近傍にメニュー画像を表示することが可能になり、操作のための煩雑な初期設定が不要になる。また、複数の操作者候補から実際に操作を行う操作者を選択することも容易に可能になる。

#### 【0046】

##### [実施例2]

10

実施例1では、先にメニュー画像を操作した操作者候補を実際に操作を行う操作者として選択しているが、以下のようにステレオマッチングの手法を利用して実際に操作を行う操作者を選択するようにしてもよい。実施例2では、ビデオカメラ1を2台用意するなどして、2つの鏡面動画像を用意する。図12、13は2つの鏡面動画像にそれぞれメニュー画像を合成した合成画像の例示図である。

#### 【0047】

2つの鏡面動画像は、例えば右目用及び左目用のステレオ画像である。図12、13の例では、図12が右目用画像、図13が左目用画像である。操作者検出部109は、これらの2つの合成画像を比較して、メニュー画像により近い方の操作者候補を操作者として選択する。図12、13の例では、いずれも左側の操作者候補がメニュー画像に近いために、左側を実際に操作する操作者とする。

20

#### 【0048】

以上のように、いずれの実施例においてもメニュー画像を操作者（又は操作者候補）が操作しやすい領域に容易に表示することができるので、鏡面動画像を用いた入力インターフェースにおける複雑な初期設定などが不要となる。また、複数の操作者候補から適宜1人を選択することができ、複数の人物が鏡面動画像内に表示される場合でも、初期設定が困難になったり、複数の操作者が同時に入力を行うことがないようにすることができる。

#### 【0049】

また、図8のような動き画像マップや図9のようなカラー領域マップを用いて、メニュー画像以外のオブジェクト画像を表示する領域を決めることもできる。

30

前述の通りこれらのマップは、操作者の動きを定量的に表すことができる。動き画像マップは、操作者の動きが多い領域、少ない領域、まったく動きのない領域を表すことができる。表示される差分画像が多い領域が操作者の動きが多い領域、差分画像が少ない領域が操作者の動きが少ない領域、差分画像がまったくない領域が操作者の動きがまったくない領域である。

#### 【0050】

カラー領域マップは、色の変化により操作者の動きを表すことができる。例えば1フレーム前のフレームの画面上の各領域の色と、現在のフレームの画面の各領域の色を比較することで、色が変化している領域を見つけることができる。まったく異なる色に変化している場合には操作者の動きがあったと判断できる。

40

#### 【0051】

図14は、操作者の動きに応じてオブジェクト画像を表示したときの例示図である。

図14では、操作者の他にベル画像141とボタン画像142とが表示されている。操作者がベル画像141に触れるとき、例えばベルが鳴る等のイベントが発生するようになっている。ボタン画像142を触れた場合には、ボタン画像に関連づけされたイベントが発生するようになっている。

#### 【0052】

ベル画像141は、操作者の動きの頻度が少ない領域に表示される。動き画像マップにより、操作者の動きの頻度がわかるので、画像合成部106は、これに応じて頻度が少ない領域を選択して、その領域にベル画像141が表示されるように合成画像を生成する。

50

ボタン画像 142 は、色の変化が大きい領域に表示される。カラー領域マップでは色の変化を検出できるため、画像合成部 106 は、これに応じて色の変化が大きい領域を選択して、その領域にボタン画像 142 が表示されるように合成画像を生成する。

このように、オブジェクト画像を操作者の動作に応じて配置することも容易に可能になる。

【図面の簡単な説明】

【0053】

【図 1】本発明を適用した画像処理システムの全体構成図。

【図 2】本実施形態による画像処理装置の構成図。

【図 3】本実施形態の画像処理装置が具備する機能プロック図。 10

【図 4】本実施形態の処理手順を示すフローチャート。

【図 5】マーク画像が合成された鏡面動画像の例示図。

【図 6】帯状のマーク画像が合成された鏡面動画像の例示図。

【図 7】メニュー画像の提示候補領域を表す例示図。

【図 8】動き画像マップの例示図。

【図 9】カラー領域マップの例示図。

【図 10】メニュー画像を提示したときの合成画像の例示図。

【図 11】操作者を決定したときの合成画像の例示図。

【図 12】ステレオ画像の右目用画像の例示図。

【図 13】ステレオ画像の左目用画像の例示図。 20

【図 14】操作者の動きに応じてオブジェクト画像を表示したときの例示図。

【符号の説明】

【0054】

1 ビデオカメラ

2 画像処理装置

3 ディスプレイ装置

4 ビデオカメラ制御装置

101 画像入力部

102 画像反転部

103 オブジェクトデータ記憶部 30

104 オブジェクトデータ取得部

105 オブジェクト画像生成部

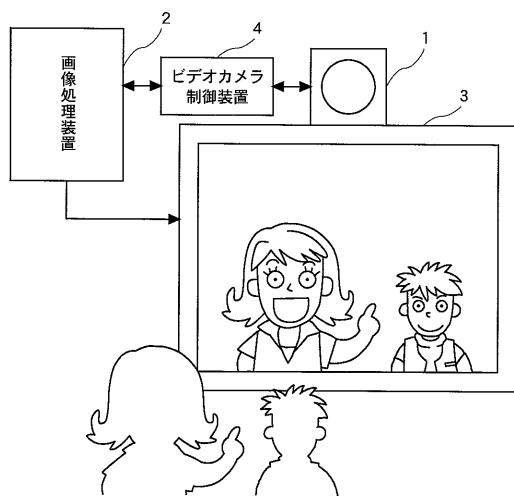
106 画像合成部

107 画像比較部

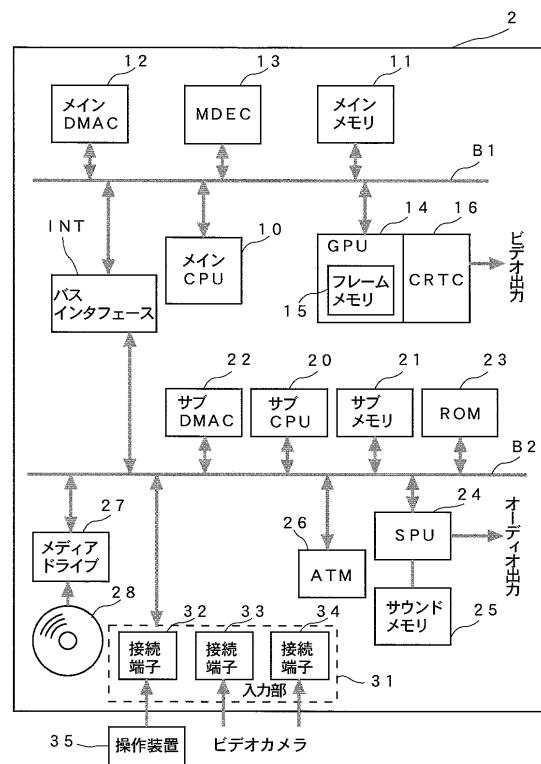
108 表示制御部

109 操作者検出部

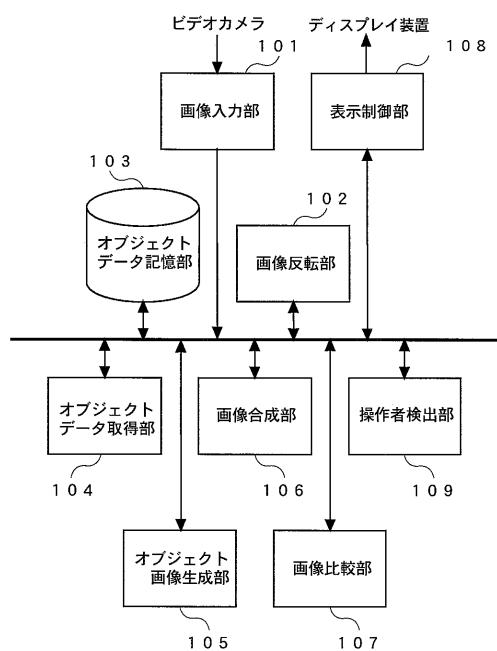
【図1】



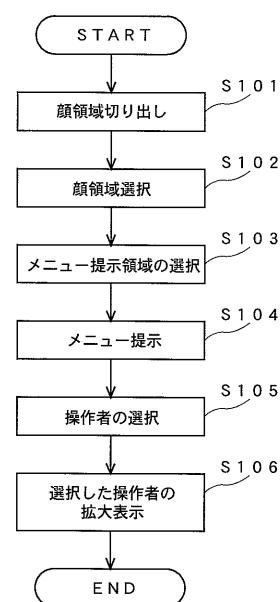
【図2】



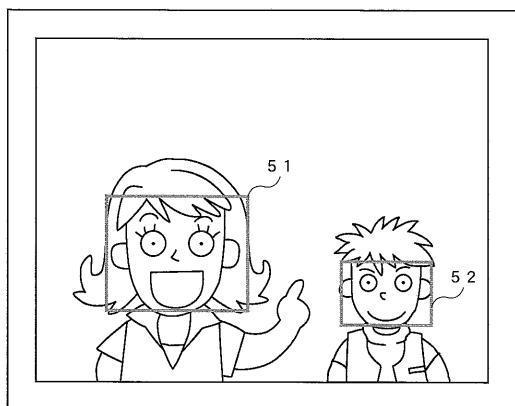
【図3】



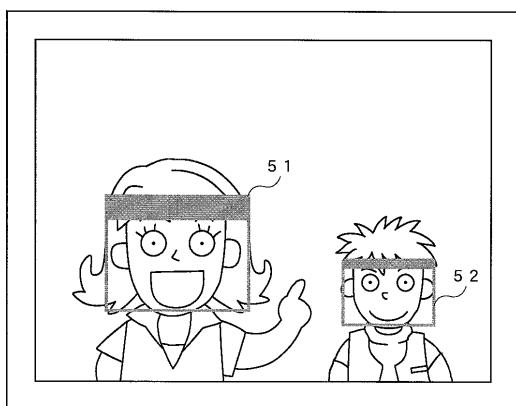
【図4】



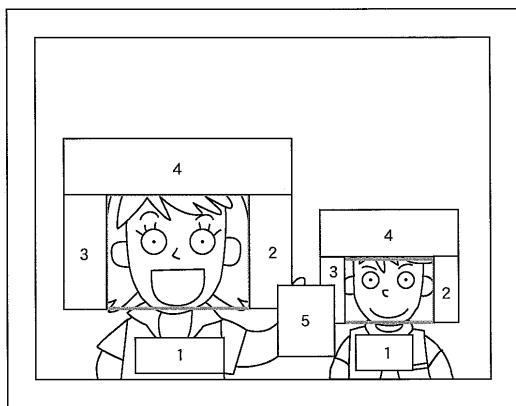
【図 5】



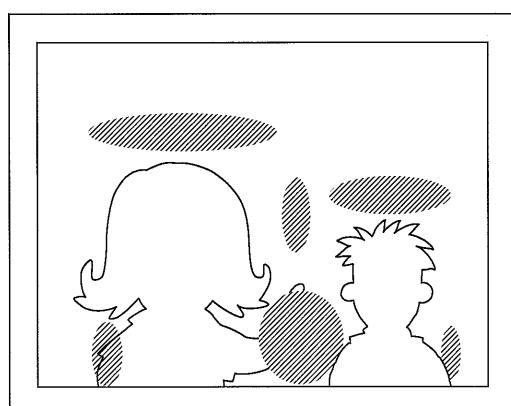
【図 6】



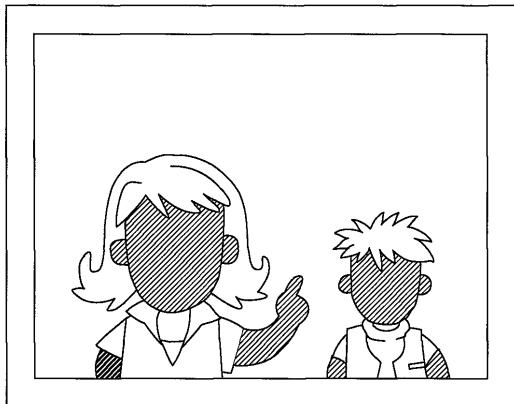
【図 7】



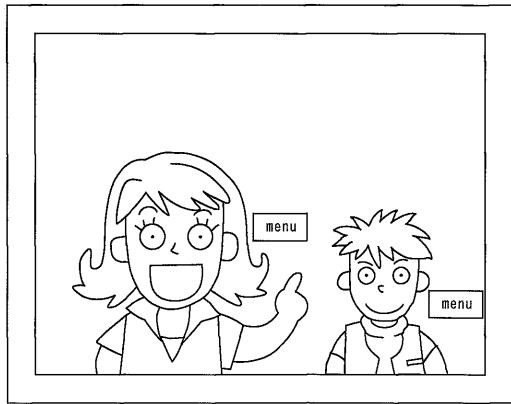
【図 8】



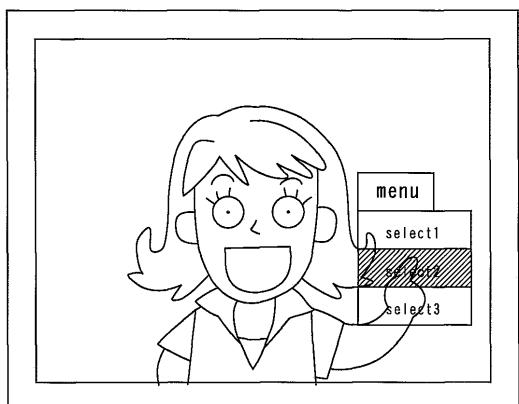
【図9】



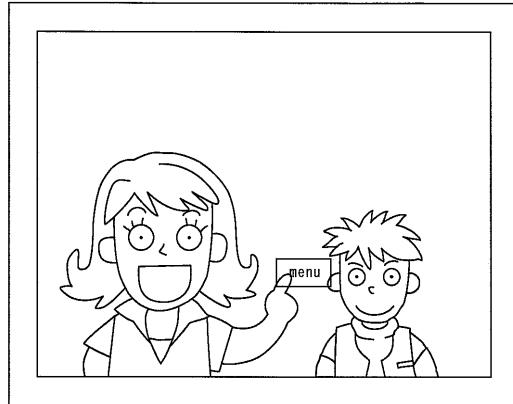
【図10】



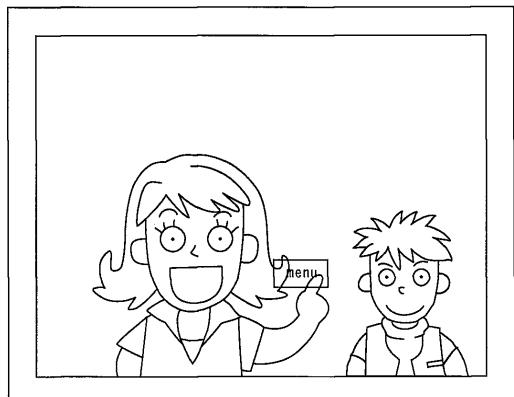
【図11】



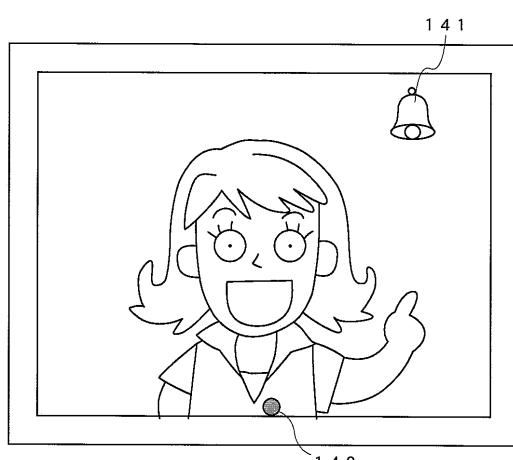
【図12】



【図13】



【図14】



---

フロントページの続き

(72)発明者 掛 智一

東京都港区南青山二丁目6番21号 株式会社ソニー・コンピュータエンタテインメント内

合議体

審判長 清田 健一

審判官 小林 義晴

審判官 須田 勝巳

(56)参考文献 特開2002-196855(JP,A)

国際公開第99/60522(WO,A1)

特開平08-315154(JP,A)

佐々木博史 ほか3名, “『てのひらめにゅう』: ウェアラブルコンピュータ用入力インタフェース”, 日本バーチャルリアリティ学会論文誌 第7巻 第3号, 日本バーチャルリアリティ学会, 2002年9月30日, p. 393-401

待井康弘 ほか2名, “動き情報の検出とヒューマンインタフェースへの応用”, ヒューマンインターフェース 39-7, 情報処理学会研究報告 Vol. 91 No. 95, 社団法人情報処理学会, 1991年11月14日, p. 1-7

佐々木博史 ほか3名, “てのひらだいやる” Wearable Computer用入力インターフェース, 電子情報通信学会技術研究報告 Vol. 100 No. 565, 社団法人電子情報通信学会, 2001年1月11日, p. 77-84

(58)調査した分野(Int.Cl., DB名)

G06F3/033, G06F3/048