

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5636888号
(P5636888)

(45) 発行日 平成26年12月10日(2014.12.10)

(24) 登録日 平成26年10月31日(2014.10.31)

(51) Int.Cl.		F I	
G06F	3/01	(2006.01)	G06F 3/01
G06F	3/16	(2006.01)	G06F 3/01 310C
H04N	5/00	(2011.01)	G06F 3/16 320H
			G06F 3/16 320A
			H04N 5/00 A

請求項の数 17 (全 37 頁)

(21) 出願番号 特願2010-250713 (P2010-250713)
 (22) 出願日 平成22年11月9日(2010.11.9)
 (65) 公開番号 特開2012-103840 (P2012-103840A)
 (43) 公開日 平成24年5月31日(2012.5.31)
 審査請求日 平成25年10月29日(2013.10.29)

(73) 特許権者 000002185
 ソニー株式会社
 東京都港区港南1丁目7番1号
 (74) 代理人 100095957
 弁理士 亀谷 美明
 (74) 代理人 100096389
 弁理士 金本 哲男
 (74) 代理人 100101557
 弁理士 萩原 康司
 (74) 代理人 100128587
 弁理士 松本 一騎
 (72) 発明者 前田 幸徳
 東京都港区港南1丁目7番1号 ソニー株式会社内

最終頁に続く

(54) 【発明の名称】 情報処理装置、プログラムおよびコマンド生成方法

(57) 【特許請求の範囲】

【請求項1】

2種類以上の異なる入力操作により取得した2種類以上の入力情報から、該入力情報の種類毎に各入力操作の意味する意味情報を認識する入力認識部と、

2種類以上の前記意味情報を組合せることにより、対象装置に所定の動作を実行させるためのコマンドを生成する動作処理部と、

予め登録されたユーザIDの中から、前記入力操作を行うユーザのユーザIDを特定する個人判別部と、
 を備え、

前記入力認識部は、前記特定されたユーザIDに応じて、前記入力情報の種類毎に前記入力操作を行うユーザの特性に合わせた意味情報を認識し、

前記入力認識部は、前記特定されたユーザIDに応じて、前記入力情報の種類毎に、予め入力操作をパターン化した入力パターンに対してユーザID毎に予め定められた優先度に基づいて、前記各入力情報に対応する前記入力パターンを特定し、該入力パターンに関連付けられた前記意味情報を抽出する、
 情報処理装置。

【請求項2】

前記入力認識部は、前記特定されたユーザIDに応じて、前記入力情報の種類毎に、ユーザID毎に予め入力操作をパターン化した入力パターンのうち、前記各入力情報に対応する前記入力パターンを特定し、該入力パターンに関連付けられた前記意味情報を抽出す

る、請求項 1 に記載の情報処理装置。

【請求項 3】

2 種類以上の異なる入力操作により取得した 2 種類以上の入力情報から、該入力情報の種類毎に各入力操作の意味する意味情報を認識する入力認識部と、

2 種類以上の前記意味情報を組合せることにより、対象装置に所定の動作を実行させるためのコマンドを生成する動作処理部と、

を備え、

前記動作処理部は、前記コマンドが入力操作の一部を省略することが可能な省略対象に指定されている場合に、前記コマンドを生成するための 2 種類以上の前記意味情報のうち 1 種類以上の前記意味情報が認識されると、前記コマンドを生成する、
情報処理装置。

10

【請求項 4】

前記動作処理部は、前記コマンドの生成頻度、最新に生成されたコマンドから過去に向かって連続的に生成された所定個数のコマンドのうち少なくとも 1 つ以上のコマンド、前記動作処理部により繰り返し実行されるコマンド生成処理のうち最新のコマンド生成処理の開始時点から過去に向かって所定時間内に生成されたコマンド、およびユーザにより指定された省略対象に関する情報のうち少なくとも 1 つ以上に基いて、前記コマンドを前記省略対象に指定する、請求項 3 に記載の情報処理装置。

【請求項 5】

前記動作処理部は、前記コマンドにより所定の動作が実行される前に、該所定の動作の実行の可否をユーザに確認させるための確認表示を、前記対象装置またはその他の装置の表示画面に表示するように制御する、請求項 3 に記載の情報処理装置。

20

【請求項 6】

前記入力認識部は、前記 2 種類以上の入力情報から少なくとも前記所定の動作の対象を示す意味情報と前記所定の動作の内容を示す意味情報とを認識する、請求項 1 又は 3 に記載の情報処理装置。

【請求項 7】

前記 2 種類以上の入力情報は、音声による入力操作により取得した音声入力情報、および体の一部若しくは全部の動きまたは状態による入力操作により取得した 1 種類以上のジェスチャ入力情報のうちの、少なくとも 2 種類以上の入力情報を含む、請求項 1 又は 3 に記載の情報処理装置。

30

【請求項 8】

前記情報処理装置は、前記 2 種類以上の入力情報および別の入力操作により取得した別の入力情報のうち、いずれかの入力情報から、前記所定の動作の実行量を示す実行量情報を認識する変化量変換部をさらに備え、

前記動作処理部は、前記 2 種類以上の前記意味情報および前記実行量情報を組合せることにより、対象装置に所定量の動作を実行させるためのコマンドを生成する、

請求項 1 又は 3 に記載の情報処理装置。

【請求項 9】

前記いずれかの入力情報は、音声による入力操作により取得した音声入力情報、または体の一部若しくは全部の動きまたは状態による入力操作により取得したジェスチャ入力情報であり、

40

前記変化量変換部は、前記いずれかの入力情報が前記音声入力情報である場合には、前記音声入力情報から取得される音量から前記実行量情報を認識し、前記いずれかの入力情報が前記ジェスチャ入力情報である場合には、前記ジェスチャ入力情報から取得される体の一部若しくは全部の動きまたは状態の変化量から前記実行量情報を認識する、

請求項 8 に記載の情報処理装置。

【請求項 10】

前記動作処理部は、前記コマンドの生成のために必要な前記意味情報のうち 1 種類以上の前記意味情報が認識されていない場合に、前記認識されていない前記意味情報の候補を

50

特定し、該候補の意味情報を意味する入力操作を、前記対象装置またはその他の装置の表示画面に表示するように制御する、請求項 1 又は 3 に記載の情報処理装置。

【請求項 1 1】

前記動作処理部は、前記コマンドの生成のために必要な前記意味情報のうち 1 種類以上の前記意味情報が認識されていない場合に、前記認識されていない前記意味情報の候補を特定し、該候補の意味情報と既に認識された意味情報により生成される前記コマンドを特定し、該コマンドにより所定の動作が実行される前の、前記対象装置に関する該動作の対象の状態を、前記対象装置またはその他の装置の表示画面に表示するように制御する、請求項 1 又は 3 に記載の情報処理装置。

【請求項 1 2】

前記対象装置は、前記情報処理装置および前記情報処理装置に接続された機器のうち少なくとも 1 つ以上であり、

前記動作処理部は、生成された前記コマンドに従って前記情報処理装置および前記機器の少なくとも 1 つ以上に所定の動作を実行させる、

請求項 1 又は 3 に記載の情報処理装置。

【請求項 1 3】

前記動作処理部は、生成された前記コマンドに従って所定の動作が実行された結果を示す結果情報を、前記対象装置またはその他の装置の表示画面に表示するように制御する、請求項 1 2 に記載の情報処理装置。

【請求項 1 4】

コンピュータを、

2 種類以上の異なる入力操作により取得した 2 種類以上の入力情報から、該入力情報の種類毎に各入力操作の意味する意味情報を認識する入力認識部と、

2 種類以上の前記意味情報を組合せることにより、対象装置に所定の動作を実行させるためのコマンドを生成する動作処理部と、

予め登録されたユーザ ID の中から、前記入力操作を行うユーザのユーザ ID を特定する個人判別部と、

として機能させるためのプログラムであり、

前記入力認識部は、前記特定されたユーザ ID に応じて、前記入力情報の種類毎に前記入力操作を行うユーザの特性に合わせた意味情報を認識し、

前記入力認識部は、前記特定されたユーザ ID に応じて、前記入力情報の種類毎に、予め入力操作をパターン化した入力パターンに対してユーザ ID 毎に予め定められた優先度に基づいて、前記各入力情報に対応する前記入力パターンを特定し、該入力パターンに関連付けられた前記意味情報を抽出する、

プログラム。

【請求項 1 5】

コンピュータを、

2 種類以上の異なる入力操作により取得した 2 種類以上の入力情報から、該入力情報の種類毎に各入力操作の意味する意味情報を認識する入力認識部と、

2 種類以上の前記意味情報を組合せることにより、対象装置に所定の動作を実行させるためのコマンドを生成する動作処理部と、

として機能させるためのプログラムであり、

前記動作処理部は、前記コマンドが入力操作の一部を省略することが可能な省略対象に指定されている場合に、前記コマンドを生成するための 2 種類以上の前記意味情報のうち 1 種類以上の前記意味情報が認識されると、前記コマンドを生成する、

プログラム。

【請求項 1 6】

2 種類以上の異なる入力操作により取得した 2 種類以上の入力情報から、該入力情報の種類毎に各入力操作の意味する意味情報を認識する認識ステップと、

2 種類以上の前記意味情報を組合せることにより、対象装置に所定の動作を実行させる

10

20

30

40

50

ためのコマンドを生成する生成ステップと、

予め登録されたユーザIDの中から、前記入力操作を行うユーザのユーザIDを特定する特定ステップと、

を含み、

前記認識ステップは、前記特定されたユーザIDに応じて、前記入力情報の種類毎に前記入力操作を行うユーザの特性に合わせた意味情報を認識することを含み、

前記特性に合わせた前記意味情報を認識することは、前記特定されたユーザIDに応じて、前記入力情報の種類毎に、予め入力操作をパターン化した入力パターンに対してユーザID毎に予め定められた優先度に基づいて、前記各入力情報に対応する前記入力パターンを特定し、該入力パターンに関連付けられた前記意味情報を抽出すること、を含む、
コマンド生成方法。

10

【請求項17】

2種類以上の異なる入力操作により取得した2種類以上の入力情報から、該入力情報の種類毎に各入力操作の意味する意味情報を認識する認識ステップと、

2種類以上の前記意味情報を組合せることにより、対象装置に所定の動作を実行させるためのコマンドを生成する生成ステップと、

を含み、

前記生成ステップは、前記コマンドが入力操作の一部を省略することが可能な省略対象に指定されている場合に、前記コマンドを生成するための2種類以上の前記意味情報のうち1種類以上の前記意味情報が認識されると、前記コマンドを生成すること、を含む、
コマンド生成方法。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、情報処理装置、プログラムおよびコマンド生成方法に関する。

【背景技術】

【0002】

従来から、各種装置を動作させるために、例えば、入力装置として、キーボードやマウス、テレビ等の家庭用電気機械器具用のリモートコントローラ等が使用されている。しかしながら、対象装置を動作させるために上記のような従来型の入力装置を使用することは、必ずしもユーザにとって直感的で分かりやすいものではない場合がある。また、ユーザは、入力装置を紛失した場合、対象装置を動作できなくなるおそれがある。

30

【0003】

そこで、直感的で分かりやすい音声、ジェスチャ等による入力操作により対象装置を動作させることを可能とする、ユーザインターフェースに関する技術が提案されている。例えば、下記特許文献1では、ユーザの入力操作を撮影した動画像からジェスチャを認識し、認識結果に基づいて制御命令を生成する技術が開示されている。また、下記特許文献2には、音声、ジェスチャ等のうち2種類以上の入力操作を用いて、一方の入力操作により取得した入力情報に基づき処理を実行し、他方の入力操作により取得した入力情報に基づき処理の実行に対する制御（開始、中断、等）を行う技術が開示されている。

40

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2003-334389号公報

【特許文献2】特開2004-192653号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

50

しかしながら、音声、ジェスチャ等による入力操作の場合、ユーザは、対象装置に与えられる命令コマンドと各音声、各ジェスチャ等との対応関係を覚えなければならない。特に、上記特許文献2のように2種類以上の入力操作を使用する場合には、各コマンドと入力操作との対応関係を覚えておくことは非常に困難である。

【0006】

そこで、本発明は、このような問題に鑑みてなされたもので、その目的は、2種類以上の入力操作を用いて対象装置に所望の動作を実行させるための入力操作を簡易にすることが可能な新規かつ改良された情報処理装置、プログラムおよびコマンド生成方法を提供しようとするものである。

【課題を解決するための手段】

【0007】

上記課題を解決するために、本発明のある観点によれば、2種類以上の異なる入力操作により取得した2種類以上の入力情報から、当該入力情報の種類毎に各入力操作の意味する意味情報を認識する入力認識部と、2種類以上の上記意味情報を組合せることにより、対象装置に所定の動作を実行させるためのコマンドを生成する動作処理部と、を備える情報処理装置が提供される。

【0008】

また、上記入力認識部は、上記2種類以上の入力情報から少なくとも上記所定の動作の対象を示す意味情報と上記所定の動作の内容を示す意味情報とを認識してもよい。

【0009】

また、上記2種類以上の入力情報は、音声による入力操作により取得した音声入力情報、および体の一部若しくは全部の動きまたは状態による入力操作により取得した1種類以上のジェスチャ入力情報のうちの、少なくとも2種類以上の入力情報を含んでもよい。

【0010】

また、上記情報処理装置は、上記2種類以上の入力情報および別の入力操作により取得した別の入力情報のうち、いずれかの入力情報から、上記所定の動作の実行量を示す実行量情報を認識する変化量変換部をさらに備え、上記動作処理部は、上記2種類以上の上記意味情報および上記実行量情報を組合せることにより、対象装置に所定量の動作を実行させるためのコマンドを生成してもよい。

【0011】

また、上記いずれかの入力情報は、音声による入力操作により取得した音声入力情報、または体の一部若しくは全部の動きまたは状態による入力操作により取得したジェスチャ入力情報であり、上記変化量変換部は、上記いずれかの入力情報が上記音声入力情報である場合には、上記音声入力情報から取得される音量から上記実行量情報を認識し、上記いずれかの入力情報が上記ジェスチャ入力情報である場合には、上記ジェスチャ入力情報から取得される体の一部若しくは全部の動きまたは状態の変化量から上記実行量情報を認識してもよい。

【0012】

また、上記情報処理装置は、予め登録されたユーザIDの中から、上記入力操作を行うユーザのユーザIDを特定する個人判別部をさらに備え、上記入力識別部は、上記特定されたユーザIDに応じて、上記入力情報の種類毎に上記入力操作を行うユーザの特性に合わせた意味情報を認識してもよい。

【0013】

また、上記入力識別部は、上記特定されたユーザIDに応じて、上記入力情報の種類毎に、ユーザID毎に予め入力操作をパターン化した入力パターンのうち、上記各入力情報に対応する上記入力パターンを特定し、当該入力パターンに関連付けられた上記意味情報を抽出してもよい。

【0014】

また、上記入力識別部は、上記特定されたユーザIDに応じて、上記入力情報の種類毎に、予め入力操作をパターン化した入力パターンに対してユーザID毎に予め定められた

10

20

30

40

50

優先度に基づいて、上記各入力情報に対応する上記入力パターンを特定し、当該入力パターンに関連付けられた上記意味情報を抽出してもよい。

【0015】

また、上記動作処理部は、上記コマンドが入力操作の一部を省略することが可能な省略対象に指定されている場合に、上記コマンドを生成するための2種類以上の上記意味情報のうち1種類以上の上記意味情報が認識されると、上記コマンドを生成してもよい。

【0016】

また、上記動作処理部は、上記コマンドの生成頻度、最新に生成されたコマンドから過去に向かって連続的に生成された所定個数のコマンドのうち少なくとも1つ以上のコマンド、上記動作処理部により繰り返し実行されるコマンド生成処理のうち最新のコマンド生成処理の開始時点から過去に向かって所定時間内に生成されたコマンド、およびユーザにより指定された省略対象に関する情報のうち少なくとも1つ以上に基づいて、上記コマンドを上記省略対象に指定してもよい。

10

【0017】

また、上記動作処理部は、上記コマンドにより所定の動作が実行される前に、当該所定の動作の実行の可否をユーザに確認させるための確認表示を、上記対象装置またはその他の装置の表示画面に表示するように制御してもよい。

【0018】

また、上記動作処理部は、上記コマンドの生成のために必要な上記意味情報のうち1種類以上の上記意味情報が認識されていない場合に、上記認識されていない上記意味情報の候補を特定し、当該候補の意味情報を意味する入力操作を、上記対象装置またはその他の装置の表示画面に表示するように制御してもよい。

20

【0019】

また、上記動作処理部は、上記コマンドの生成のために必要な上記意味情報のうち1種類以上の上記意味情報が認識されていない場合に、上記認識されていない上記意味情報の候補を特定し、当該候補の意味情報と既に認識された意味情報により生成される上記コマンドを特定し、当該コマンドにより所定の動作が実行される前の、上記対象装置に関する当該動作の対象の状態を、上記対象装置またはその他の装置の表示画面に表示するように制御してもよい。

【0020】

また、上記対象装置は、上記情報処理装置および上記情報処理装置に接続された機器のうち少なくとも1つ以上であり、上記動作処理部は、生成された上記コマンドに従って上記情報処理装置および上記機器の少なくとも1つ以上に所定の動作を実行させてもよい。

30

【0021】

また、上記動作処理部は、生成された上記コマンドに従って所定の動作が実行された結果を示す結果情報を、上記対象装置またはその他の装置の表示画面に表示するように制御してもよい。

【0022】

また、上記課題を解決するために、本発明の別の観点によれば、コンピュータを、2種類以上の異なる入力操作により取得した2種類以上の入力情報から、当該入力情報の種類毎に各入力操作の意味する意味情報を認識する入力認識部と、2種類以上の上記意味情報を組合せることにより、対象装置に所定の動作を実行させるためのコマンドを生成する動作処理部と、として機能させるためのプログラムが提供される。

40

【0023】

さらに、上記課題を解決するために、本発明の別の観点によれば、2種類以上の異なる入力操作により取得した2種類以上の入力情報から、当該入力情報の種類毎に各入力操作の意味する意味情報を認識するステップと、2種類以上の上記意味情報を組合せることにより、対象装置に所定の動作を実行させるためのコマンドを生成するステップと、を含むコマンド生成方法が提供される。

【発明の効果】

50

【 0 0 2 4 】

以上説明したように、本発明に係る情報処理装置、プログラムおよびコマンド生成方法によれば、2種類以上の入力操作を用いて対象装置に所望の動作を実行させるための入力操作を簡易にすることができる。

【 図面の簡単な説明 】

【 0 0 2 5 】

【 図 1 】本発明の第 1 の実施形態に係る情報処理装置の機能構成を示すブロック図である。

【 図 2 】音声記憶部が記憶する音声認識辞書の一例を示す図である。

【 図 3 】ジェスチャ記憶部が記憶するジェスチャ認識辞書の一例を示す第 1 の図である。 10

【 図 4 】ジェスチャ記憶部が記憶するジェスチャ認識辞書の一例を示す第 2 の図である。

【 図 5 】コマンド記憶部が記憶するコマンド辞書の一例を示す第 1 の図である。

【 図 6 】コマンドによる動作の実行結果の一例を示す第 1 の図である。

【 図 7 】コマンドによる動作の実行結果の一例を示す第 2 の図である。

【 図 8 】入力情報と意味情報との関係の一例を示す図である。

【 図 9 】第 1 の実施形態に係るコマンド生成処理を示すフローチャートである。

【 図 1 0 】本発明の第 2 の実施形態に係る情報処理装置の機能構成を示すブロック図である。

【 図 1 1 】変化量記憶部が記憶する変化量変換辞書の一例を示す第 1 の図である。

【 図 1 2 】変化量記憶部が記憶する変化量変換辞書の一例を示す第 2 の図である。 20

【 図 1 3 】コマンド記憶部が記憶するコマンド辞書の一例を示す第 2 の図である。

【 図 1 4 】第 2 の実施形態に係るコマンド生成処理を示すフローチャートである。

【 図 1 5 】本発明の第 3 の実施形態に係る情報処理装置の機能構成を示すブロック図である。

【 図 1 6 】ユーザ I D 毎の音声認識辞書およびジェスチャ認識辞書の一例を示す第 1 の図である。

【 図 1 7 】ユーザ I D 毎の音声認識辞書およびジェスチャ認識辞書の一例を示す第 2 の図である。

【 図 1 8 】第 3 の実施形態に係るコマンド生成処理を示すフローチャートである。

【 図 1 9 】本発明の第 4 の実施形態に係る情報処理装置の機能構成を示すブロック図である。 30

【 図 2 0 】動作内容記憶部が記憶する情報の一例を示す図である。

【 図 2 1 】頻度情報記憶部が記憶する情報の一例を示す図である。

【 図 2 2 】コマンド記憶部が記憶するコマンド辞書の一例を示す第 3 の図である。

【 図 2 3 】省略対象のコマンドの候補を表示する表示画面の一例を示す図である。

【 図 2 4 】コマンドの実行可否の確認表示を表示する表示画面の一例を示す図である

【 図 2 5 】第 4 の実施形態に係るコマンド生成処理を示すフローチャートである。

【 図 2 6 】本発明の第 5 の実施形態に係る情報処理装置の機能構成を示すブロック図である。

【 図 2 7 】入力操作の候補を表示する表示画面の一例を示す第 1 の図である。 40

【 図 2 8 】入力操作の候補を表示する表示画面の一例を示す第 2 の図である。

【 図 2 9 】対象装置に関する動作の対象の状態を表示する表示画面の一例を示す第 1 の図である。

【 図 3 0 】対象装置に関する動作の対象の状態を表示する表示画面の一例を示す第 2 の図である。

【 図 3 1 】第 5 の実施形態に係るコマンド生成処理を示すフローチャートである。

【 図 3 2 】本発明の各実施形態に係る情報処理装置のハードウェア構成の一例を示すブロック図である。

【 発明を実施するための形態 】

【 0 0 2 6 】

以下に添付図面を参照しながら、本発明の好適な実施の形態について詳細に説明する。なお、本明細書および図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付すことにより重複説明を省略する。

【0027】

また、以下の順序にしたがって当該「発明を実施するための形態」を説明する。

1. 第1の実施形態

1-1. 情報処理装置の構成

1-2. 処理の流れ

2. 第2の実施形態

2-1. 情報処理装置の構成

2-2. 処理の流れ

3. 第3の実施形態

3-1. 情報処理装置の構成

3-2. 処理の流れ

4. 第4の実施形態

4-1. 情報処理装置の構成

4-2. 処理の流れ

5. 第5の実施形態

5-1. 情報処理装置の構成

5-2. 処理の流れ

6. 発明の各実施形態に係る情報処理装置のハードウェア構成

7. まとめ

【0028】

以下に説明される各実施形態では、ユーザが操作したい対象装置に対して行う入力操作として、2種類以上の入力操作が行われる。また、2種類以上の入力操作により取得した2種類以上の入力情報として、音声による入力操作により取得した音声入力情報と、体の一部若しくは全体の動きまたは状態による入力操作により取得したジェスチャ入力情報が用いられる。なお、音声入力情報およびジェスチャ認識情報は、ユーザの入力操作により取得した2種類以上の入力操作により取得した入力情報の一例である。

【0029】

また、各実施形態に係る情報処理装置は、上記入力情報から対象装置を動作させるコマンドを生成する。情報処理装置は、例えば、テレビ、プロジェクタ、DVDレコーダ、ブルーレイレコーダ、音楽プレイヤー、ゲーム機器、エアコンディショナ、洗濯機、冷蔵庫等の家電機器、PC (Personal Computer)、プリンタ、スキャナ、スマートフォン、携帯情報端末 (Personal Digital Assistant) 等の情報処理機器、または照明機器、湯沸かし器等のその他の機器であってよい。また、情報処理装置は、これら装置と接続される周辺機器であってよい。

【0030】

< 1. 第1の実施形態 >

[1-1. 情報処理装置の構成]

以下では、図1~8を用いて、本発明の第1の実施形態に係る情報処理装置の構成について説明する。

【0031】

図1は、本発明の第1の実施形態に係る情報処理装置100の機能構成を示すブロック図である。図1を参照すると、情報処理装置100は、音声入力情報取得部110、ジェスチャ入力情報取得部120、音声認識部130、音声記憶部132、ジェスチャ認識部140、ジェスチャ記憶部142、動作処理部150およびコマンド記憶部152を有する。なお、入力認識部は音声認識部130とジェスチャ認識部140との組合せであるものとして説明される。

【0032】

10

20

30

40

50

音声入力情報取得部 110 は、ユーザが行う音声による入力操作により音声入力情報を取得する。例えば、ユーザが音声による入力操作を行うと、音声入力情報取得部 110 は、収録された音声から音声波形信号を抽出し、音声波形信号をアナログ/デジタル変換することにより、デジタル化された音声情報を音声入力情報として取得する。また、音声入力情報取得部 110 は、デジタル化された音声情報から音声に関する特徴量をさらに抽出し、当該特徴量を音声入力情報として取得してもよい。その後、音声入力情報取得部 110 は、取得された音声入力情報を音声認識部 130 へ出力する。なお、情報処理装置 100 と接続された外部の装置が、収録された音声から音声入力情報を取得し、音声入力情報取得部 110 は、当該外部の装置から音声自体、デジタル化された音声情報または上記特徴量のいずれかの情報の形で音声入力情報を受信してもよい。

10

【0033】

ジェスチャ入力情報取得部 120 は、ユーザが行う体の一部若しくは全体の動きまたは状態による入力操作によりジェスチャ入力情報を取得する。例えば、ユーザが手の動きによる入力操作を行うと、ジェスチャ入力情報取得部 120 は、情報処理装置 100 に取付けられたカメラを用いてユーザの手の動きを撮影することにより、デジタル化された動画画像情報をジェスチャ入力情報として取得する。また、ジェスチャ入力情報取得部 120 は、デジタル化された動画画像情報から抽出された手の動きに関する特徴量をジェスチャ入力情報として取得してもよい。その後、ジェスチャ入力情報取得部 120 は、取得されたジェスチャ入力情報をジェスチャ認識部 140 へ出力する。なお、入力操作は、手の動きに限られず、体全体、頭、指、顔（表情）または目（視線）等の体の別の部分の動きによるものであってもよい。また、入力操作は、体の一部若しくは全体の動的な動きによるものに限られず、体の一部若しくは全体の静的な状態によるものであってもよい。また、ジェスチャ入力情報は、動画画像情報に限られず、静止画像情報でもよく、またセンサ等によって得られるその他の信号情報であってもよい。また、情報処理装置 100 と接続された外部の装置が、上記ジェスチャ入力情報を取得し、ジェスチャ入力情報取得部 120 は、当該外部の装置からデジタル化された動画画像、上記抽出された特徴量等の形式でジェスチャ入力情報を受信してもよい。

20

【0034】

音声記憶部 132 は、予め定められた入力パターンと、当該入力パターンに関連付けられた意味情報とを、音声認識辞書として記憶する。ここで、入力パターンとは、例えば、音声による入力操作を予めパターン化した情報を指す。また、意味情報とは、入力操作の意味する情報を指す。図 2 は、音声記憶部 132 が記憶する音声認識辞書の一例を示す。図 2 を参照すると、音声認識辞書には、例えば「ちゃんねる」、「おんりょう」等が入力パターンとして記憶されている。上記入力パターンは、例えば、デジタル化された音声情報や音声に関する特徴量等の、音声入力情報との比較が可能な形式で、記憶されている。また、音声認識辞書には、例えば「ちゃんねる」という入力パターンに関連付けられた「動作の対象はチャンネル」という意味情報、「おんりょう」という入力パターンに関連付けられた「動作の対象は音量」という意味情報が、意味情報として記憶されている。

30

【0035】

音声認識部 130 は、音声による入力操作により取得した音声入力情報から、音声による入力操作の意味する意味情報を認識する。例えば、音声認識部 130 は、入力パターンのうち、音声入力情報に対応する入力パターンを特定し、当該入力パターンに関連付けられた意味情報を抽出する。

40

【0036】

音声認識部 130 は、音声入力情報取得部 110 により音声入力情報を入力されると、音声記憶部 132 から入力パターンを取得する。次に、音声認識部 130 は、例えば、音声入力情報と各入力パターンとの間の適合の程度を示すスコアを算出し、スコアが最も大きい入力パターンを特定する。音声入力情報と各入力パターンとの比較による上記スコアの算出は、現在までの公知の音声認識に関する従来技術を用いることにより実行され得る。次に、音声認識部 130 は、特定された入力パターンに関連付けられた意味情報を音声

50

記憶部 1 3 2 から抽出する。このようにして、音声認識部 1 3 0 は、入力された音声入力情報から音声による入力操作の意味する意味情報を認識する。最後に、音声認識部 1 3 0 は、認識された意味情報を動作処理部 1 5 0 へ出力する。

【 0 0 3 7 】

例えば、「おんりょう」という音声により取得された音声入力情報が、音声認識部 1 3 0 に入力される。図 2 を参照すると、例えば、音声認識部 1 3 0 は、音声入力情報と各入力パターンとの間の上記スコア（図示せず）を算出した結果、最もスコアが大きかった入力パターンである「おんりょう」を特定する。すると、音声認識部 1 3 0 は、「おんりょう」と関連付けられた意味情報である「動作の対象は音量」を意味情報として抽出する。

【 0 0 3 8 】

ジェスチャ記憶部 1 4 2 は、体の一部若しくは全部の動きまたは状態による入力操作を予めパターン化した入力パターンと、当該入力パターンに関連付けられた意味情報とを、ジェスチャ認識辞書として記憶する。図 3 は、ジェスチャ記憶部 1 4 2 が記憶するジェスチャ認識辞書の一例を示す。図 3 を参照すると、ジェスチャ認識辞書には、例えば「手を上に挙げる」、「手を下にさげる」等が入力パターンとして記憶されている。上記入力パターンは、例えば、手の動きに関する動画像や手の動きに関する特徴量等の、ジェスチャ入力情報との比較が可能な形式で、記憶されている。また、ジェスチャ認識辞書には、例えば「手を上に挙げる」という入力パターンに関連付けられた「パラメータを上げる」という意味情報、「手を下にさげる」という入力パターンに関連付けられた「パラメータを下げる」という意味情報等が、記憶されている。

【 0 0 3 9 】

図 4 は、ジェスチャ記憶部 1 4 2 が記憶するジェスチャ認識辞書の他の一例を示す。手の動きまたは状態による入力操作ではなく、体の別の部分の動きまたは状態による入力操作が行われる場合、ジェスチャ記憶部 1 4 2 は、図 3 に例示される入力パターンの代わりに、図 4 に例示される入力パターンを記憶してもよい。例えば、ジェスチャ認識辞書には、「全ての指を広げて伸ばす」、「全ての指を閉じる」等が入力パターンとして記憶されてもよい。

【 0 0 4 0 】

ジェスチャ認識部 1 4 0 は、体の一部若しくは全体の動きまたは状態による入力操作により取得したジェスチャ入力情報から、体の一部若しくは全体の動きまたは状態による入力操作の意味する意味情報を認識する。例えば、ジェスチャ認識部 1 4 0 は、入力パターンのうち、ジェスチャ入力情報に対応する入力パターンを特定し、当該入力パターンに関連付けられた意味情報を抽出する。

【 0 0 4 1 】

ジェスチャ認識部 1 4 0 は、ジェスチャ入力情報取得部 1 2 0 によりジェスチャ入力情報を入力されると、ジェスチャ記憶部 1 4 2 から入力パターンを取得する。次に、ジェスチャ認識部 1 4 0 は、例えば、ジェスチャ入力情報と各入力パターンとの間の適合の程度を示すスコアを算出し、スコアが最も大きい入力パターンを特定する。ジェスチャ入力情報と各入力パターンとの比較による上記スコアの算出は、現在までの公知のジェスチャ認識に関する従来技術を用いることにより実行され得る。次に、ジェスチャ認識部 1 4 0 は、特定された入力パターンに関連付けられた意味情報をジェスチャ記憶部 1 4 2 から抽出する。このようにして、ジェスチャ認識部 1 4 0 は、入力されたジェスチャ入力情報から体の一部若しくは全体の動きまたは状態による入力操作の意味する意味情報を認識する。最後に、ジェスチャ認識部 1 4 0 は、認識された意味情報を動作処理部 1 5 0 へ出力する。

【 0 0 4 2 】

例えば、手を上に挙げる動作により取得されたジェスチャ入力情報が、ジェスチャ認識部 1 4 0 に入力される。図 3 を参照すると、例えば、ジェスチャ認識部 1 4 0 は、ジェスチャ入力情報と各入力パターンとの間の上記スコアを算出した結果、最もスコアが大きかった入力パターンである「手を上に挙げる」を特定する。すると、ジェスチャ認識部 1 4

10

20

30

40

50

0 は、「手を上に挙げる」と関連付けられた意味情報である「パラメータを上げる」を意味情報として抽出する。

【0043】

コマンド記憶部152は、ユーザが入力操作を行う対象装置に所定の動作を実行させるためのコマンドと、コマンドに対応する2種類以上の意味情報の組合せとを、コマンド辞書として記憶する。図5は、コマンド記憶部152が記憶するコマンド辞書の一例を示す。図5を参照すると、コマンド辞書には、例えば「チャンネルを昇順に変える」、「音量を上げる」等のコマンドが、記憶されている。上記コマンドは、例えば、対象装置が読み取り可能なデータ形式で記憶されている。また、コマンド辞書には、例えば「チャンネルを昇順に変える」というコマンドに対応する、「パラメータを上げる」、「動作の対象はチャンネル」等が、意味情報の組合せとして記憶されている。

10

【0044】

動作処理部150は、2種類以上の意味情報を組合せることにより、対象装置に所定の動作を実行させるためのコマンドを生成する。ここでの意味情報は、音声認識部130により認識された意味情報、ジェスチャ認識部140により認識された意味情報という2種類の意味情報である。動作処理部150は、音声認識部130およびジェスチャ認識部140から意味情報の入力を受けると、コマンド記憶部152からこれらの意味情報の組合せに対応するコマンドを抽出する。抽出されたコマンドは、対象装置に所定の動作を実行させるためのコマンドである。このようにして、動作処理部150は、対象装置に所定の動作を実行させるためのコマンドを生成する。

20

【0045】

動作処理部150は、生成されたコマンドに従って対象装置に所定の動作を実行させる。また、動作処理部150は、生成されたコマンドに従って所定の動作が実行された結果を示す結果情報を、対象装置またはその他の装置の表示画面に表示するように制御する。ここで、その他の装置とは、例えば対象装置と直接的にまたは間接的に接続された装置である。

【0046】

例えば、動作処理部150は、音声認識部130により「動作の対象は音量」という意味情報を入力され、ジェスチャ認識部140により「パラメータを上げる」という意味情報を入力される。図5を参照すると、動作処理部150は、「動作の対象は音量」という意味情報と「パラメータを上げる」という意味情報の組合せと対応する「音量を上げる」というコマンドを生成する。すると、動作処理部150は、生成された「音量を上げる」というコマンドに従って対象装置に「音量を上げる」動作を実行させる。図6は、コマンドによる動作の実行結果の一例を示す。上記のように「音量を上げる」動作が実行されると、動作処理部150は、図6に示されるように、対象装置またはその他の装置の表示画面の例えば右下に上げられた音量を結果情報として表示するように制御する。

30

【0047】

また、例えば、動作処理部150は、音声認識部130により「動作の対象はチャンネル」という意味情報を入力され、ジェスチャ認識部140により「パラメータを上げる」という意味情報を入力される。図5を参照すると、動作処理部150は、「動作の対象はチャンネル」という意味情報と「パラメータを上げる」という意味情報の組合せと対応する「チャンネルを昇順に変える」というコマンドを生成する。すると、動作処理部150は、生成された「チャンネルを昇順に変える」というコマンドに従って対象装置に「チャンネルを昇順に変える」動作を実行させる。図7は、コマンドによる動作の実行結果の一例を示す。上記のように「チャンネルを昇順に変える」動作が実行されると、動作処理部150は、図7に示されるように、対象装置またはその他の装置の表示画面の例えば右下に昇順に変えられたチャンネルを結果情報として表示するように制御する。

40

【0048】

なお、動作処理部150が動作を実行させる対象装置は、情報処理装置100および情報処理装置に接続された機器のうち少なくとも1つ以上でもよい。例えば、対象装置がテ

50

テレビであり、当該テレビ自体が情報処理装置 100 であってもよい。また、例えば、対象装置がエアコンディショナであり、情報処理装置 100 は当該エアコンディショナと接続された周辺機器であってもよい。また、例えば、対象装置が PC、プリンタ、スキャナであり、情報処理装置 100 はこれらに接続された周辺機器であってもよい。

【0049】

以上、情報処理装置 100 が有する音声入力情報取得部 110、ジェスチャ入力情報取得部 120、音声認識部 130、音声記憶部 132、ジェスチャ認識部 140、ジェスチャ記憶部 142、動作処理部 150 およびコマンド記憶部 152 の各々について説明した。ここでは、さらに、音声認識部 130 とジェスチャ認識部 140 に共通する事項を説明し、その後で音声記憶部 132 およびジェスチャ記憶部 142 に共通する事項を説明する。

10

【0050】

また、本実施形態では、音声認識部 130 は、音声入力情報から所定の動作の対象を示す意味情報を認識し、ジェスチャ認識部 140 は、ジェスチャ入力情報から所定の動作の内容を示す意味情報を認識する。入力情報に対応する入力パターンと意味情報との関係の一例を示した図 8 を参照して、その関係について説明する。図 8 のとおり、例えば、音声入力情報から「おんりょう」という入力パターンが特定された場合に、「動作の対象は音量」という意味情報が認識される。また、音声入力情報から「ちゃんねる」という入力パターンが特定された場合に、「動作の対象はチャンネル」という意味情報が認識される。このようにして、音声入力情報から動作の対象を示す意味情報が認識される。また、例えば、ジェスチャ入力情報から「手を上に挙げる」という入力パターンが特定された場合に、「パラメータを上げる」という意味情報が認識される。また、例えば、ジェスチャ入力情報から「手を下にさげる」という入力パターンが特定された場合に、「パラメータを下げる」という意味情報が認識される。このようにして、各入力情報から、それぞれ無作為に定められた意味情報が認識されるのではなく、それぞれ動作の内容を示す意味情報および動作の対象を示す意味情報が認識される。これにより、ユーザは、各入力操作が意味し得る意味情報を想定しやすいため、入力操作をより容易に覚え得る。

20

【0051】

音声記憶部 132 およびジェスチャ記憶部 142 では、図 2 および図 3 のとおり、複数の入力パターンに同一の意味情報が関連付けられ得る。図 2 を参照すると、例えば、「ちゃんねる」、「ばんぐみ」という 2 つの入力パターンに、「対象はチャンネル」という同一の意味情報が関連付けられている。また、図 3 を参照すると、例えば、「手を上に挙げる」、「手を押し出す」という 2 つの入力パターンに、「パラメータを上げる」という同一の意味情報が関連付けられている。この場合、ユーザは、特定の意味情報を装置に認識させるために細かく入力操作を覚える必要はない。ユーザは、特定の意味情報を示すいずれかの入力操作のうち覚えやすいものを覚えればよい。または、ユーザは、特定の意味情報を示すいくつかの入力操作を知り、入力操作を行う際に思い出せるものを使えばよい。その結果、ユーザが必ず覚えなければならない入力操作は減り得る。なお、入力パターンと意味情報を 1 対 1 に対応付けてもよい。

30

【0052】

[1 - 2 . 処理の流れ]

以下では、図 9 を用いて、本発明の第 1 の実施形態に係るコマンド生成処理について説明する。図 9 は、第 1 の実施形態に係るコマンド生成処理を示すフローチャートである。

40

【0053】

図 9 を参照すると、まずステップ S 310 において、音声入力情報取得部 110 が、ユーザの音声による入力操作から音声入力情報を取得する。また、ジェスチャ入力情報取得部 120 が、ユーザの体の一部若しくは全体の動きまたは状態による入力操作からジェスチャ入力情報を取得する。

【0054】

次に、ステップ S 320 において、音声認識部 130 は、音声入力情報から音声による

50

入力操作の意味する意味情報を認識する。また、ジェスチャ認識部140は、ジェスチャ入力情報から体の一部若しくは全体の動きまたは状態による入力操作の意味する意味情報を認識する。

【0055】

ステップS330において、動作処理部150は、音声認識部130およびジェスチャ認識部140から、コマンドの生成に必要な意味情報が全て認識され、入力されているかを判定する。具体的には、例えば、動作処理部150は、所定の時間内に必要な意味情報が全て入力されていなければ、本処理を終了する。一方で、動作処理部150は、コマンドの生成に必要な意味情報が全て入力されていれば、コマンドの生成に必要な意味情報が全て認識されていると判定し、ステップS340へ進む。また、例えば、動作処理部150は、所定の時間毎に意味情報の有無を確認し、一方の意味情報の入力のみがあれば、所定の時間後に再度他方の意味情報の入力の有無を確認してもよい。その結果、動作処理部150は、他方の意味情報の入力がなければ、コマンドの生成に必要ないずれかの意味情報が認識されていないと判定し、本処理を終了する。他方の意味情報の入力があれば、コマンドの生成に必要な意味情報が全て認識されていると判定し、ステップS340へ進む。

10

【0056】

次に、ステップS340において、動作処理部150は、2種類以上の意味情報を組合せることにより対象装置に所定の動作を実行させるためのコマンドを生成する。本実施形態では、動作処理部150は、認識された意味情報を組合せることにより生成できるコマンドがある場合にはコマンドを生成し、認識された意味情報を組合せることにより生成できるコマンドがない場合にはコマンドを生成しない。

20

【0057】

ステップ350において、動作処理部150は、コマンドが生成されているかを判定する。ここで、コマンドが生成されている場合には、処理はステップS360へ進む。一方、コマンドが生成されていない場合には、本処理を終了する。

【0058】

最後に、ステップ360において、動作処理部150は、生成されたコマンドに従って対象装置に所定の動作を実行させる。また、動作処理部150は、生成されたコマンドに従って所定の動作が実行された結果を示す結果情報を、対象装置またはその他の装置の表示画面に表示するように制御する。

30

【0059】

以上が本発明の第1の実施形態に係るコマンド生成処理の流れである。なお、当該コマンド生成処理は、情報処理装置の起動時に実行され、その後コマンド生成処理の終了時に繰り返し実行されてもよい。または、当該コマンド生成処理は、例えば、所定の時間間隔で繰り返し実行されてもよい。

【0060】

< 2. 第2の実施形態 >

本発明の第2の実施形態に係る情報処理装置は、本発明の第1の実施形態に係る情報処理装置が有する機能に対して、対象装置に実行させる動作の実行量を入力操作に基づいて変える機能がさらに追加されたものである。

40

【0061】

[2 - 1 . 情報処理装置の構成]

以下では、図10～13を用いて、本発明の第2の実施形態に係る情報処理装置の構成について説明する。

【0062】

図10は、本発明の第2の実施形態に係る情報処理装置100の機能構成を示すブロック図である。図10を参照すると、情報処理装置100は、音声入力情報取得部110、ジェスチャ入力情報取得部120、音声認識部130、音声記憶部132、ジェスチャ認識部140、ジェスチャ記憶部142、動作処理部150、コマンド記憶部152、変化

50

量変換部 160 および変化量記憶部 162 を有する。

【0063】

このうち、音声認識部 130、音声記憶部 132、ジェスチャ認識部 140、ジェスチャ記憶部 142 は、第 1 の実施形態として [1 - 1 . 情報処理装置の構成] において説明したとおりである。よって、新たに追加されている変化量変換部 160 および変化量記憶部 162、並びに第 1 の実施形態と一部の機能が相違する音声入力情報取得部 110、ジェスチャ入力情報取得部 120、動作処理部 150 およびコマンド記憶部 152 の相違箇所を中心に説明する。

【0064】

音声入力情報取得部 110 は、変化量変換部 160 に音声入力情報を出し、変化量変換部 160 は音声入力情報から所定の動作の実行量を示す実行量情報を認識する。

10

【0065】

ジェスチャ入力情報取得部 120 は、変化量変換部 160 にジェスチャ入力情報を出し、変化量変換部 160 がジェスチャ入力情報から所定の動作の実行量を示す実行量情報を認識する。本実施形態では、変化量変換部 160 は、音声入力情報またはジェスチャ入力情報の少なくともいずれかから実行量情報を認識する。

【0066】

変化量記憶部 162 は、所定の動作の実行量を示す実行量情報と、音声入力情報またはジェスチャ入力情報から実行量情報を認識するための判断基準とを、変化量変換辞書として記憶する。

20

【0067】

図 11 は、変化量記憶部 162 が記憶する変化量変換辞書の一例を示す。図 11 は、ジェスチャ入力情報から取得される手の動きの変化量から実行量情報を認識する場合の変化量変換辞書の例を示している。この場合、変化量変換辞書には、例えば「手の動きの変化量は X 未満」、「手の動きの変化量は X 以上 Y 未満」、「手の動きの変化量は Y 以上」の場合に動作の実行量を「小」、「中」、「大」とする判断基準が記憶されている。なお、動作の実行量は、数値として表されてもよい。

【0068】

図 12 は、変化量記憶部 162 が記憶する変化量変換辞書の一例を示す。図 12 は、ジェスチャ入力情報の他の一例である目の動きにより取得した手の動きによるジェスチャ入力情報とは別の入力情報から、実行量情報を認識する場合の、変化量変換辞書の例を示している。この場合、変化量変換辞書には、例えば「目を細めている」、「画面を凝視している」場合に、動作の実行量を「画面の輝度を上げる場合：動作の実行 大、その他の場合：動作の実行 小」、「音量を上げる / 下げる場合：動作の実行 大、その他の場合：動作の実行 小」とする判断基準が、記憶されている。

30

【0069】

変化量変換部 160 は、例えば、変化量変換部 160 は、入力情報が音声入力情報である場合には、音声入力情報から取得される音量から実行量情報を認識し、入力情報がジェスチャ入力情報である場合には、ジェスチャ入力情報から取得される体の一部若しくは全体の動きまたは状態の変化量から実行量情報を認識する。

40

【0070】

変化量変換部 160 は、音量から実行量情報を認識する場合、音声入力情報から音声の音量を取得する。または、変化量変換部 160 は、体の一部若しくは全体の動きまたは状態の変化量から実行量情報を認識する場合、ジェスチャ入力情報から体の一部若しくは全体の動きまたは状態の変化量を取得する。ここで、体の一部若しくは全体の動きの変化量とは、例えば、動きの開始時点と動きの完了時点との間で体の一部若しくは全体が変化した程度であってもよい。また、体の一部若しくは全体の状態の変化量とは、例えば、撮影等された体の一部若しくは全体の状態と基準となる体の一部若しくは全体の状態との間で、体の一部若しくは全体が相違する程度であってもよい。体の一部若しくは全体の動きまたは状態の変化量の取得は、現在までの公知のジェスチャ認識に関する従来技術を用いる

50

ことにより実行され得る。次に、変化量変換部 160 は、判断基準によると音量または変化量に対応する動作の実行量を、変化量記憶部 162 から取得する。このようにして、変化量変換部 160 は、動作の実行量を示す実行量情報を認識する。最後に、変化量変換部 160 は、認識された実行量情報を動作処理部 150 へ出力する。

【0071】

例えば、手を大きく上に挙げる動作により取得されたジェスチャ入力情報が、変化量変換部 160 に入力される。そして、変化量変換部 160 は、ジェスチャ入力情報から手の動きの変化量 A3 を取得する。図 11 を参照すると、例えば、測定された変化量 A3 は Y 以上であるため、動作の実行量が「大」という実行量情報を変化量記憶部 162 から取得する。このようにして、変化量変換部 160 は、動作の実行量が「大」という実行量情報を認識する。

10

【0072】

なお、変化量変換部 160 は、意味情報を認識するために用いた音声入力情報およびジェスチャ入力情報とは異なる、別の入力操作により取得した別の入力情報から、所定の動作の実行量を示す実行量情報を認識してもよい。変化量変換部 160 は、上記別の入力情報が入力されると、例えば変化量記憶部 162 から上記別の入力情報から実行量情報を認識するための判断基準を取得する。次に、変化量変換部 160 は、例えば、上記別の入力情報と各判断基準との間の適合の程度を示すスコアを算出し、スコアが最も大きい判断基準を特定する。次に、変化量変換部 160 は、特定された判断基準に対応する実行量情報を変化量記憶部 162 から抽出する。例えばこのようにして、変化量変換部 160 は、別の入力操作により取得した上記別の入力情報から実行量情報を認識してもよい。

20

【0073】

別の入力操作が目の動きによる入力操作である場合の例を説明する。例えば、目を細める動作により取得された上記別の入力情報が、変化量変換部 160 に入力される。図 12 を参照すると、例えば、変化量変換部 160 は、上記別の入力情報と各判断基準との間のスコアを算出した結果、最もスコアが大きかった判断基準である「目を細めている」を特定する。すると、変化量変換部 160 は、「目を細めている」という判断基準と対応する動作の実行量である「画面の輝度を下げる場合：動作の実行量 大、その他の場合：動作の実行量 小」を実行量情報として抽出する。

【0074】

30

コマンド記憶部 152 は、対象装置に所定量の動作を実行させるためのコマンドと、コマンドに対応する意味情報および実行量情報との組合せとをコマンド辞書として記憶する。図 13 は、コマンド記憶部 152 が記憶するコマンド辞書の他の例を示す。図 13 を参照すると、コマンド辞書には、例えば「音量を 1 ポイント上げる」、「音量を 3 ポイント上げる」等のコマンドが、記憶されている。また、コマンド辞書には、「パラメータを上げる」、「動作の対象は音量」等の意味情報と「小」、「大」等の実行量情報との組合せが記憶されている。

【0075】

動作処理部 150 は、2 種類以上の意味情報および実行量情報を組合せることにより、対象装置に所定量の動作を実行させるためのコマンドを生成する。ここでの意味情報は、音声認識部 130 により認識された意味情報、ジェスチャ認識部 140 により認識された意味情報という 2 種類の意味情報である。動作処理部 150 は、意味情報を入力されるだけでなく、変化量変換部 160 により実行量情報を入力されると、意味情報および実行量情報の組合せと対応するコマンドをコマンド記憶部 152 から取得する。

40

【0076】

[2 - 2 . 処理の流れ]

以下では、図 14 を用いて、本発明の第 2 の実施形態に係るコマンド生成処理について説明する。図 14 は、第 2 の実施形態に係るコマンド生成処理を示すフローチャートである。この中で、ステップ S310、ステップ S320、ステップ S330、ステップ S350 およびステップ S360 は、[1 - 2 . 処理の流れ] にて説明した第 1 の実施形態に

50

係るコマンド生成処理と同じである。よって、ここでは新たに追加されているステップ S 3 2 2、および第 1 の実施形態と一部の処理が相違するステップ S 3 4 0 の相違箇所を中心に説明する。

【 0 0 7 7 】

ステップ 3 2 2 において、変化量変換部 1 6 0 は、意味情報を認識するための音声入力情報またはジェスチャ入力情報のうちいずれかの入力情報から、所定の動作の実行量を示す実行量情報を認識する。

【 0 0 7 8 】

また、ステップ 3 4 0 において、動作処理部 1 5 0 は、2 種類以上の意味情報および実行量情報を組合せることにより、対象装置に所定量の動作を実行させるためのコマンドを生成する。

【 0 0 7 9 】

< 3 . 第 3 の実施形態 >

本発明の第 3 の実施形態に係る情報処理装置は、本発明の第 1 の実施形態に係る情報処理装置が有する機能に対して、各ユーザの特性に合わせた意味情報の認識を行う機能がさらに追加されたものである。

【 0 0 8 0 】

[3 - 1 . 情報処理装置の構成]

以下では、図 1 5 ~ 1 7 を用いて、本発明の第 3 の実施形態に係る情報処理装置の構成について説明する。

【 0 0 8 1 】

図 1 5 は、本発明の第 3 の実施形態に係る情報処理装置 1 0 0 の機能構成を示すブロック図である。図 1 5 を参照すると、情報処理装置 1 0 0 は、音声入力情報取得部 1 1 0、ジェスチャ入力情報取得部 1 2 0、音声認識部 1 3 0、音声記憶部 1 3 2、ジェスチャ認識部 1 4 0、ジェスチャ記憶部 1 4 2、動作処理部 1 5 0、コマンド記憶部 1 5 2 および個人判別部 1 7 0 を有する。

【 0 0 8 2 】

このうち、動作処理部 1 5 0 およびコマンド記憶部 1 5 2 は、第 1 の実施形態として [1 - 1 . 情報処理装置の構成] において説明したとおりである。よって、新たに追加されている個人判別部 1 7 0、並びに第 1 の実施形態と一部の機能が相違する音声入力情報取得部 1 1 0、ジェスチャ入力情報取得部 1 2 0、音声認識部 1 3 0、音声記憶部 1 3 2、ジェスチャ認識部 1 4 0 およびジェスチャ記憶部 1 4 2 の相違箇所を中心に説明する。

【 0 0 8 3 】

音声入力情報取得部 1 1 0 は、個人判別部 1 7 0 が音声入力情報から入力操作を行うユーザのユーザ ID を特定する場合には、個人判別部 1 7 0 に音声入力情報を出力する。

【 0 0 8 4 】

ジェスチャ入力情報取得部 1 2 0 は、個人判別部 1 7 0 がジェスチャ入力情報から入力操作を行うユーザのユーザ ID を特定する場合には、個人判別部 1 7 0 にジェスチャ入力情報を出力する。

【 0 0 8 5 】

個人判別部 1 7 0 は、予め登録されたユーザ ID の中から、入力操作を行うユーザのユーザ ID を特定する。個人判別部 1 7 0 は、例えば、ユーザによる入力操作により取得した音声入力情報またはジェスチャ入力情報から、予め登録されたユーザ ID を特定する。例えば、音声入力情報からユーザ ID を特定する場合、個人判別部 1 7 0 は、音声入力情報が入力されると、音声入力情報の中の音声情報と予め登録された各ユーザの音声の特徴量とを比較する。個人判別部 1 7 0 は、比較の結果に基づき、例えば最も適合する特徴量を特定することにより、ユーザ ID を特定する。また、ジェスチャに入力情報からユーザ ID を特定する場合、個人判別部 1 7 0 は、ジェスチャ力情報が入力されると、例えばジェスチャ力情報の中のユーザの顔の画像と予め登録された各ユーザの顔の特徴量とを比較する。個人判別部 1 7 0 は、比較の結果に基づき、例えば最も適合する特徴量を特定する

10

20

30

40

50

ことにより、ユーザIDを特定する。最後に、個人判別部170は、特定されたユーザIDを音声認識部130およびジェスチャ認識部140へ出力する。なお、個人判別部170は、ユーザIDの特定のために、意味情報を認識するための入力情報を使用せずに、別の情報を使用してもよい。例えば、ユーザIDカードからの読込情報、リモコン、マウス、キーボード等の入力装置により入力されたユーザID情報等の、意味情報を認識するための入力情報以外の別の入力情報が、使用されてもよい。

【0086】

音声記憶部132はおよびジェスチャ記憶部142は、それぞれユーザID毎の音声認識辞書およびジェスチャ認識辞書を記憶する。

【0087】

図16は、ユーザID毎の音声認識辞書およびジェスチャ認識辞書の一例を示す。図16には、ユーザID毎に予め定められた入力パターンが記憶された、ユーザID毎の音声認識辞書およびジェスチャ認識辞書の例を示している。図16を参照すると、ユーザAの音声認識辞書には、例えば「ちゃんねる」、「おんりょう」等の入力パターンが記憶されている。一方で、ユーザBの音声認識辞書には、「ばんぐみ」、「ぼりゅうむ」等の入力パターンが記憶されている。また、ユーザAのジェスチャ認識辞書には、例えば「手を上に挙げる」、「手を下に下げる」等の入力パターンが記憶されている。一方で、ユーザBのジェスチャ認識辞書には、「手を押し出す」、「手を引き戻す」等の入力パターンが記憶されている。なお、入力パターンに関連付けられた意味情報も記憶されている。

【0088】

また、図17は、ユーザID毎の音声認識辞書およびジェスチャ認識辞書の他の例を示す。図17には、入力パターンに対してユーザID毎に予め定められた優先度が記憶された、ユーザID毎の音声認識辞書およびジェスチャ認識辞書の例を示している。図17を参照すると、ユーザAの音声認識辞書には、例えば「ちゃんねる」という入力パターンに対して、優先度として例えば「+0.5」というスコア付加値が記憶されている。一方で、ユーザBの音声認識辞書には、例えば「ちゃんねる」という入力パターンに対して、優先度として例えば「+0」というスコア付加値が記憶されている。また、ユーザAのジェスチャ認識辞書には、例えば「手を押し戻す」という入力パターンに対して、優先度として例えば「+0」というスコア付加値が記憶されている。一方で、ユーザBのジェスチャ認識辞書には、例えば「手を押し戻す」という入力パターンに対して、優先度として例えば「+0.5」というスコア付加値が記憶されている。なお、図17には表示されていないものの、入力パターンに関連付けられた意味情報も記憶されている。

【0089】

音声認識部130およびジェスチャ認識部140は、特定されたユーザIDに応じて、入力操作を行うユーザの特性に合わせた意味情報を認識する。例えば、音声認識部130およびジェスチャ認識部140は、特定されたユーザIDに応じて、ユーザID毎の入力パターンのうち、入力情報に対応する入力パターンを特定し、当該入力パターンに関連付けられた意味情報を抽出する。

【0090】

音声認識部130およびジェスチャ認識部140は、ここでは同様の処理を行うため、音声認識部130を例にとって説明する。音声認識部130は、音声入力情報取得部110により音声入力情報を入力され、また個人判別部170により特定されたユーザIDが入力される。音声認識部130は、特定されたユーザIDの音声認識辞書に記憶されている、特定されたユーザIDに対して予め定められた入力パターンを取得する。次に、音声認識部130は、例えば、音声入力情報と各入力パターンとの間の適合の程度を示すスコアを算出し、スコアが最も大きい入力パターンを特定する。次に、音声認識部130は、特定されたユーザIDの音声認識辞書において、特定された入力パターンに関連付けられた意味情報を、音声記憶部132から抽出する。このようにして、音声認識部130は、例えばユーザID毎に予め定められた入力パターンを用いて、ユーザの特性に合わせた意味情報を認識する。

10

20

30

40

50

【 0 0 9 1 】

例えば、ユーザ A による「おんりょう」という音声により取得された音声入力情報が、音声認識部 1 3 0 に入力される。図 1 6 を参照すると、例えば、音声認識部 1 3 0 は、ユーザ A の音声認識辞書に記憶された入力パターンである「おんりょう」を特定する。すると、音声認識部 1 3 0 は、「おんりょう」と関連付けられた意味情報である「動作の対象は音量」を意味情報として抽出する。

【 0 0 9 2 】

なお、音声認識部 1 3 0 およびジェスチャ認識部 1 4 0 は、特定されたユーザ ID に応じて、入力パターンに対してユーザ ID 毎に予め定められた優先度に基づいて、入力情報に対応する入力パターンを特定し、当該入力パターンに関連付けられた意味情報を抽出してもよい。例えば、音声認識部 1 3 0 は、音声入力情報取得部 1 1 0 により音声入力情報を入力され、また個人判別部 1 7 0 により特定されたユーザ ID が入力される。音声認識部 1 3 0 は、特定されたユーザ ID の音声認識辞書に記憶されている、入力パターン、および当該入力パターンに対して予め定められた優先度である例えばスコア付加値を取得する。次に、音声認識部 1 3 0 は、例えば、音声入力情報と各入力パターンとの間の適合の程度を示すスコアを算出し、当該スコアと各入力パターンのスコア付加値の和を算出する。音声認識部 1 3 0 は、例えば、当該和が最も大きい入力パターンを特定する。次に、音声認識部 1 3 0 は、特定されたユーザ ID の音声認識辞書において、特定された入力パターンに関連付けられた意味情報を、音声記憶部 1 3 2 から抽出する。このようにして、音声認識部 1 3 0 は、例えばユーザ ID 毎に予め定められた優先度を用いて、ユーザの特性に合わせた意味情報を認識する。

【 0 0 9 3 】

上記のように、入力操作をユーザの特性に合わせた意味情報を認識する手法の具体的な例として、ユーザ ID 毎に予め定められた入力パターンを用いる場合、ユーザ ID 毎に予め定められた優先度を用いる場合を説明した。しかしながら、入力操作をユーザの特性に合わせた意味情報を認識する手法は、これらの具体的な例に限られず、他の具体的な手法により実行されてもよい。

【 0 0 9 4 】

[3 - 2 . 処理の流れ]

以下では、図 1 8 を用いて、本発明の第 3 の実施形態に係るコマンド生成処理について説明する。図 1 8 は、第 3 の実施形態に係るコマンド生成処理を示すフローチャートである。この中で、ステップ S 3 1 0、ステップ S 3 3 0、ステップ S 3 4 0、ステップ S 3 5 0 およびステップ S 3 6 0 は、[1 - 2 . 処理の流れ] にて説明した第 1 の実施形態に係るコマンド生成処理と同じである。よって、ここでは新たに追加されているステップ S 3 1 2、ステップ S 3 1 4、ステップ S 3 1 6 およびステップ S 3 1 8、並びに第 1 の実施形態と一部の処理が相違するステップ S 3 2 0 の相違箇所を中心に説明する。

【 0 0 9 5 】

ステップ 3 1 2 において、個人判別部 1 7 0 は、音声入力情報またはジェスチャ入力情報から、予め登録されたユーザ ID の中から、入力操作を行うユーザのユーザ ID を特定する。

【 0 0 9 6 】

ステップ 3 1 4 において、個人判別部 1 7 0 は、ユーザ ID は既に登録されているかを判定する。ここで、ユーザ ID が登録されていない、つまりユーザ ID が特定されなかった場合には、個人判別部 1 7 0 は、ユーザ ID が特定できなかったこと示す通知を音声認識部 1 3 0 およびジェスチャ認識部 1 4 0 に出力する。その後、処理はステップ S 3 1 6 へ進む。一方、ユーザ ID が登録されている、つまりユーザ ID が特定された場合には、個人判別部 1 7 0 は、ユーザ ID を音声認識部 1 3 0 およびジェスチャ認識部 1 4 0 に出力する。その後、処理はステップ S 3 1 8 へ進む。

【 0 0 9 7 】

ステップ 3 1 6 において、音声認識部 1 3 0 およびジェスチャ認識部 1 4 0 は、それぞ

10

20

30

40

50

れ汎用の音声認識辞書および汎用のジェスチャ認識辞書を使用することを決定する。

【0098】

ステップ318において、音声認識部130およびジェスチャ認識部140は、それぞれユーザID毎の音声認識辞書およびユーザID毎のジェスチャ認識辞書は使用することを決定する。

【0099】

また、ステップ320において、音声認識部130およびジェスチャ認識部140は、それぞれ使用することを決定した音声認識辞書およびジェスチャ認識辞書を用いて、意味情報を認識する。特にユーザID毎の音声認識辞書およびジェスチャ認識辞書を使用する場合には、音声認識部130およびジェスチャ認識部140は、特定されたユーザIDに
10 応じて、入力操作を行うユーザの特性に合わせた意味情報を認識する。例えば、音声認識部130およびジェスチャ認識部140は、特定されたユーザIDに
10 応じて、ユーザID毎の入力パターンのうち、入力情報に対応する入力パターンを特定し、当該入力パターンに関連付けられた意味情報を抽出する。

【0100】

<4.第4の実施形態>

本発明の第4の実施形態に係る情報処理装置は、本発明の第1の実施形態に係る情報処理装置が有する機能に対して、コマンドを生成するために入力操作の一部を省略することを可能とする機能がさらに追加されたものである。

【0101】

[4-1.情報処理装置の構成]

以下では、図19~24を用いて、本発明の第4の実施形態に係る情報処理装置の構成について説明する。

【0102】

図19は、本発明の第4の実施形態に係る情報処理装置100の機能構成を示すブロック図である。図19を参照すると、情報処理装置100は、音声入力情報取得部110、ジェスチャ入力情報取得部120、音声認識部130、音声記憶部132、ジェスチャ認識部140、ジェスチャ記憶部142、動作処理部150、コマンド記憶部152、動作内容記憶部154および頻度情報記憶部156を有する。

【0103】

このうち、音声入力情報取得部110、ジェスチャ入力情報取得部120、音声認識部130、音声記憶部132、ジェスチャ認識部140およびジェスチャ記憶部142は、第1の実施形態として[1-1.情報処理装置の構成]において説明したとおりである。よって、新たに追加されている動作内容記憶部154および頻度情報記憶部156、並びに第1の実施形態と一部の機能が相違する動作処理部150およびコマンド記憶部152の相違箇所を中心に説明する。

【0104】

動作内容記憶部154は、最新に生成されたコマンドから過去に向かって連続的に生成された所定個数のコマンドを記憶する。例えば図9に示したコマンド生成処理を繰り返す毎に1つのコマンドを生成する動作内容記憶部154は、動作処理部150がコマンドを
40 生成する度に、動作処理部150から当該コマンドを取得する。そして、動作内容記憶部154は、当該コマンドに基づき記憶しているコマンドを更新する。なお、動作内容記憶部154は、動作処理部150により繰り返し実行されるコマンド生成処理のうち最新の
40 コマンド生成処理の開始時点から過去に向かって所定時間内に生成されたコマンドを記憶してもよい。

【0105】

図20は、動作内容記憶部154が記憶する情報の一例を示す。図20を参照すると、動作内容記憶部154は、例えば最新に生成されたコマンドから過去に向かって連続的に生成されたN個のコマンドを記憶する。例えば、「音量を上げる」というコマンドが最新の
50 コマンドとして記憶されている。また、例えば、「音量を上げる」というコマンドに対

応する「パラメータを上げる」、「動作の対象は音量」という意味情報も記憶されている。

【0106】

頻度情報記憶部156は、各コマンドの生成頻度を記憶する。頻度情報記憶部156は、例えば、動作内容記憶部154が新たなコマンドを取得する度に、動作内容記憶部154から当該コマンドを取得する。そして、頻度情報記憶部156は、当該コマンドに基づき記憶している各コマンドの生成頻度を更新する。なお、コマンドの生成頻度は、例えば、所定の期間内にコマンドが生成された回数を示す。

【0107】

図21は、頻度情報記憶部156が記憶する情報の一例を示す。図21を参照すると、例えば、「チャンネルを昇順に変える」というコマンドに対して、「8回」というコマンドの生成頻度が記憶されている。また、「チャンネルを昇順に変える」というコマンドに対して、「パラメータを上げる」、「動作の対象はチャンネル」という意味情報も記憶されている。

10

【0108】

コマンド記憶部152は、各コマンドおよびそれに対応する意味情報の組合せに加えて、省略対象に指定されているコマンドを示す省略対象情報を記憶する。例えば、コマンド記憶部152は、各コマンドに対して、省略対象であるか否かを示す省略対象識別情報を記憶する。

【0109】

図22は、コマンド記憶部152が記憶するコマンド辞書の一例を示す。図22を参照すると、例えば、コマンドの右側に省略対象であるか否かを示す省略対象識別情報が設けられ、ここでは「音量を上げる」というコマンドが省略対象として指定されている。

20

【0110】

動作処理部150は、コマンドが入力操作の一部を省略することが可能な省略対象に指定されている場合に、コマンドを生成するための2種類以上の意味情報のうち1種類以上の意味情報が認識されると、コマンドを生成する。ここでの意味情報は、音声認識部130により認識された意味情報、ジェスチャ認識部140により認識された意味情報という2種類の意味情報である。例えば所定の時間内に、音声認識部130またはジェスチャ認識部140のいずれか一方のみから意味情報を入力された場合に、動作処理部150は、コマンド記憶部152内で、入力された意味情報から生成され得るコマンドであって省略対象に指定されているものを検索する。省略対象に指定されているコマンドがあれば、動作処理部150は、コマンド記憶部152から当該コマンドを取得する。省略対象に指定されているコマンドが存在する場合、動作処理部150は、当該コマンドを対象装置に所定の動作を実行させるためのコマンドとして決定する。このようにして、動作処理部150は、省略対象に指定されているコマンドを生成する。

30

【0111】

例えば、動作処理部150は、ジェスチャ認識部140により「パラメータを上げる」という意味情報を入力され、音声認識部130により意味情報が入力されない。図22を参照すると、「音量を上げる」というコマンドが省略対象に指定されているため、動作処理部150は、「パラメータを上げる」という意味情報から、「音量を上げる」というコマンドをコマンド記憶部152から取得する。そして、動作処理部150は、「音量を上げる」というコマンドを対象装置に所定の動作を実行させるためのコマンドとして決定する。

40

【0112】

また、動作処理部150は、特定のコマンドを省略対象として指定する。例えば、動作処理部150は、コマンドの生成頻度に基づいて、特定のコマンドを省略対象として指定する。例えば、動作処理部150は、頻度情報記憶部156に記憶されているコマンドのうち生成頻度の最も高いコマンドを省略対象に指定する。図21を参照すると、例えば、動作処理部150は、生成頻度が「15回」である「音量を上げる」というコマンドを省

50

略対象に指定する。

【0113】

例えば、動作処理部150は、最新に生成されたコマンドから過去に向かって連続的に生成された所定個数のコマンドのうち少なくとも1つ以上のコマンドに基づいて、特定のコマンドを省略対象として指定する。例えば、動作処理部150は、動作内容記憶部154に記憶されているコマンドのうち最新に生成されたコマンドを省略対象に指定する。図20を参照すると、例えば、動作処理部150は、最新に生成されたコマンドである「音量を上げる」というコマンドを省略対象に指定する。なお、動作処理部150は、動作処理部150により繰り返し実行されるコマンド生成処理のうち最新のコマンド生成処理の開始時点から過去に向かって所定時間内に生成されたコマンドに基づいて、特定のコマンドを省略対象として指定してもよい。

10

【0114】

例えば、動作処理部150は、ユーザにより指定された省略対象に関する情報に基づいて、特定のコマンドを省略対象として指定する。例えば、動作処理部150は、所定の表示画面にコマンドの一覧を表示するように制御し、ユーザの入力操作により選択されたコマンドを省略対象として指定する。図23は、省略対象のコマンドの候補を表示する表示画面の一例を示す。図23を参照すると、動作処理部150は、例えば、ユーザの入力操作により選択された「音量を上げる」というコマンドを省略対象に指定する。

【0115】

なお、動作処理部150は、コマンドにより所定の動作が実行される前に、当該所定の動作の実行の可否をユーザに確認させるための確認表示を、対象装置またはその他の装置の表示画面に表示するように制御してもよい。図24は、コマンドの実行可否の確認表示を表示する表示画面の一例を示す。図24を参照すると、例えば、「音量を上げる」という省略対象に指定されているコマンドが生成された場合に、動作処理部150は、「音量を上げますか?」という確認表示を、対象装置またはその他の装置の表示画面に表示するように制御する。

20

【0116】

[4-2. 処理の流れ]

以下では、図25を用いて、本発明の第4の実施形態に係るコマンド生成処理について説明する。図25は、第4の実施形態に係るコマンド生成処理を示すフローチャートである。この中で、ステップS310、ステップS320、ステップS330、ステップS340、ステップS350およびステップS360は、[1-2. 処理の流れ]にて説明した第1の実施形態に係るコマンド生成処理と同じである。よって、ここでは新たに追加されているステップS410、ステップS420、ステップS430およびステップS440を中心に説明する。

30

【0117】

ステップS410において、動作処理部150は、コマンドを生成するための2種類の意味情報のうち一方の意味情報が認識されているかを判定する。ここで、一方の意味情報が認識されていれば、処理はステップS420へ進む。一方、いずれの意味情報も認識されていない場合には、本処理を終了する。

40

【0118】

次に、ステップS420において、動作処理部150は、入力された一方の意味情報から生成され得るコマンドであって省略対象に指定されているものがあるかを判定する。例えば、動作処理部150は、入力された一方の意味情報に基づいてコマンド記憶部152から当該コマンドを取得する。ここで、上記コマンドがあれば、処理はステップS430へ進む。一方、上記コマンドがなければ、本処理を終了する。

【0119】

次に、ステップS430において、動作処理部150は、省略対象として指定されているコマンドを生成する。例えば、動作処理部150は、上記のようにコマンド記憶部152から取得している上記コマンドを、対象装置に所定の動作を実行させるためのコマンド

50

として決定する。

【 0 1 2 0 】

最後に、ステップ S 4 4 0 において、動作処理部 1 5 0 は、特定のコマンドを省略対象として指定する。

【 0 1 2 1 】

< 5 . 第 5 の実施形態 >

本発明の第 5 の実施形態に係る情報処理装置は、本発明の第 1 の実施形態に係る情報処理装置が有する機能に対して、ユーザにより入力操作の一部が行われた場合に、さらなる入力操作の候補をユーザに提示することを可能とする機能が、さらに追加されたものである。また、ユーザにより入力操作の一部が行われた場合に、コマンドにより動作が実行される前における当該動作の対象の状態をユーザに提示することを可能とする機能が、さらに追加されたものである。

10

【 0 1 2 2 】

[5 - 1 . 情報処理装置の構成]

以下では、図 2 6 ~ 3 0 を用いて、本発明の第 5 の実施形態に係る情報処理装置の構成について説明する。

【 0 1 2 3 】

図 2 6 は、本発明の第 5 の実施形態に係る情報処理装置 1 0 0 の機能構成を示すブロック図である。図 2 6 を参照すると、情報処理装置 1 0 0 は、音声入力情報取得部 1 1 0、ジェスチャ入力情報取得部 1 2 0、音声認識部 1 3 0、音声記憶部 1 3 2、ジェスチャ認識部 1 4 0、ジェスチャ記憶部 1 4 2、動作処理部 1 5 0、コマンド記憶部 1 5 2 および時系列管理部 1 8 0 を有する。

20

【 0 1 2 4 】

このうち、音声認識部 1 3 0、ジェスチャ認識部 1 4 0 およびコマンド記憶部 1 5 2 は、第 1 の実施形態として [1 - 1 . 情報処理装置の構成] において説明したとおりである。よって、新たに追加されている時系列管理部 1 8 0、並びに第 1 の実施形態と一部の機能が相違する音声入力情報取得部 1 1 0、ジェスチャ入力情報取得部 1 2 0、音声記憶部 1 3 2、ジェスチャ記憶部 1 4 2 および動作処理部 1 5 0 の相違箇所を中心に説明する。

【 0 1 2 5 】

音声入力情報取得部 1 1 0 は、音声による入力操作から音声入力情報を取得すると、音声入力情報を取得したことを示す音声取得済情報を時系列管理部 1 8 0 へ出力する。

30

【 0 1 2 6 】

ジェスチャ入力情報取得部 1 2 0 は、体の一部若しくは全部の動きまたは状態による入力操作からジェスチャ入力情報を取得すると、ジェスチャ入力情報を取得したことを示すジェスチャ取得済情報を時系列管理部 1 8 0 へ出力する。

【 0 1 2 7 】

音声記憶部 1 3 2 は、入力パターンを、例えばデジタル化された音声情報や音声に関する特徴量等の音声入力情報との比較が可能な形式で記憶する。それに加えて、音声記憶部 1 3 2 は、さらに入力パターンを、当該入力パターンに対応する入力操作をユーザが理解できる文字情報等の形式でも記憶する。音声記憶部 1 3 2 は、動作処理部 1 5 0 からの要求に応じて当該入力パターンを動作処理部 1 5 0 に出力する。

40

【 0 1 2 8 】

ジェスチャ記憶部 1 4 2 は、入力パターンを、例えば手の動きに関する動画像や手の動きに関する特徴量等の、ジェスチャ入力情報との比較が可能な形式で記憶する。それに加えて、ジェスチャ記憶部 1 4 2 は、さらに入力パターンを、文字情報や入力操作を示す動画像または静止画像等の、当該入力パターンに対応する入力操作をユーザが理解できる形式でも記憶する。ジェスチャ記憶部 1 3 2 は、動作処理部 1 5 0 からの要求に応じて当該入力パターンを動作処理部 1 5 0 に出力する。

【 0 1 2 9 】

時系列管理部 1 8 0 は、音声入力情報およびジェスチャ入力情報の取得状況を時系列で

50

記憶する。また、時系列管理部 180 は、動作処理部 150 からの要求に応じて音声入力情報およびジェスチャ入力情報の取得状況を動作処理部 150 に出力する。時系列管理部 180 は、例えば上記音声取得済情報およびジェスチャ取得済情報により、音声入力情報およびジェスチャ入力情報の取得状況を時系列で把握し得る。

【0130】

動作処理部 150 は、コマンドの生成のために必要な意味情報のうち 1 種類以上の意味情報が認識されていない場合に、認識されていない意味情報の候補を特定し、当該候補の意味情報を意味する入力操作を、対象装置またはその他の装置の表示画面に表示するように制御する。

【0131】

動作処理部 150 は、例えば、所定の時間内に、音声認識部 130 またはジェスチャ認識部 140 のいずれか一方のみから意味情報の入力を受けた場合に、残りの意味情報を認識するための入力情報が取得されているかを時系列管理部 180 に確認する。そして、当該入力情報が取得されていない場合に、動作処理部 150 は、コマンド記憶部 152 から、既に認識されている意味情報と組合されて記憶されている意味情報を、認識されていない意味情報の候補として取得する。次に、動作処理部 150 は、例えば、音声記憶部 132 またはジェスチャ記憶部 142 から、上記候補の意味情報と関連付けられた入力パターンを取得する。そして、動作処理部 150 は、取得した入力パターンに基づいて、当該入力パターンに対応する入力操作をユーザが理解できる形式で、対象装置またはその他の装置の表示画面に表示するように制御する。表示される当該入力操作が、コマンドを生成するためのユーザによる入力操作の候補となる。

【0132】

図 27 は、入力操作の候補を表示する表示画面の一例を示す。図 27 には、ユーザが「手を上に挙げる」という手の動きによる入力操作を行った場合の表示画面の例が示されている。図 3 を参照すると、「手を上に挙げる」という入力操作により、「パラメータを挙げる」という意味情報が、ジェスチャ認識部 140 により認識される。よって、動作処理部 150 は、「パラメータを挙げる」という意味情報をジェスチャ認識部 140 から入力される。さらに図 5 を参照すると、コマンド記憶部 152 のコマンド辞書には、「動作の対象はチャンネル」、「動作の対象は音量」、「動作の対象は画面の輝度」という意味情報が、「手を上に挙げる」という意味情報と組合されて記憶されている。よって、動作処理部 150 は、「動作の対象はチャンネル」、「動作の対象は音量」、「動作の対象は画面の輝度」という意味情報の候補をコマンド記憶部 152 から取得する。また、図 2 を参照すると、音声記憶部 132 の音声認識辞書には、「ちゃんねる」、「ばんぐみ」、「おんりょう」という入力パターンが、「動作の対象はチャンネル」、「動作の対象は音量」、「動作の対象は画面の輝度」という意味情報と関連付けられて記憶されている。よって、動作処理部 150 は、「ちゃんねる」、「ばんぐみ」、「おんりょう」という入力パターンを音声記憶部 132 から取得する。そして、動作処理部 150 は、図 27 に示されるとおり、「チャンネル」、「音量」、「明るさ」という音声による入力操作の候補を表示画面に表示するように制御する。

【0133】

図 28 は、入力操作の候補を表示する表示画面他の例を示す。図 28 には、ユーザが「おんりょう」という音声による入力操作を行った場合の表示画面の例が示されている。動作処理部 150 は、上記と同様の処理を行った上で、図 28 に示されるとおり、「手を上に挙げる」、「手を下にさげる」という手の動きによる入力操作の候補を表示画面に表示するように制御する。

【0134】

なお、動作処理部 150 は、コマンドの生成のために必要な意味情報のうち 1 種類以上の意味情報が認識されていない場合に、認識されていない意味情報の候補を特定し、当該候補の意味情報と既に認識された意味情報により生成されるコマンドを特定し、当該コマンドにより所定の動作が実行される前の、対象装置に関する当該動作の対象の状態を、対

10

20

30

40

50

象装置またはその他の装置の表示画面に表示するように制御してもよい。

【 0 1 3 5 】

動作処理部 1 5 0 は、例えば、上記のように入力操作の候補を表示する場合と同じ処理により、認識されていない意味情報の候補として取得する。次に、動作処理部 1 5 0 は、例えば、コマンド記憶部 1 5 2 から、既に認識されている意味情報と上記候補の意味情報との組合せに対応するコマンドを取得する。そして、動作処理部 1 5 0 は、当該コマンドにより所定の動作が実行される前の、対象装置に関する当該動作の対象の状態を、表示画面に表示するように制御する。

【 0 1 3 6 】

図 2 9 は、対象装置に関する動作の対象の状態を表示する表示画面の一例を示す。図 2 9 には、ユーザが「手を上に挙げる」という手の動きによる入力操作が行った場合の表示画面の例が示されている。図 2 7 の場合と同様に、動作処理部 1 5 0 は、

「パラメータを上げる」という意味情報をジェスチャ認識部 1 4 0 から入力される。また、図 2 7 の場合と同様に、動作処理部 1 5 0 は、「動作の対象はチャンネル」、「動作の対象は音量」、「動作の対象は画面の輝度」という意味情報の候補をコマンド記憶部 1 5 2 から取得する。図 5 を参照すると、コマンド記憶部 1 5 2 のコマンド辞書には、「チャンネルを昇順に変える」、「音量を上げる」、「画面の輝度を上げる」というコマンドが「パラメータを上げる」という既に認識された意味情報と、「動作の対象はチャンネル」、「動作の対象は音量」、「動作の対象は画面の輝度」という候補の意味情報との組合せに対応して記憶されている。よって、動作処理部 1 5 0 は、「チャンネルを昇順に変える」、「音量を上げる」、「画面の輝度を上げる」というコマンドをコマンド記憶部 1 5 2 から取得する。そして、動作処理部 1 5 0 は、図 2 9 に示されるとおり、「チャンネルを昇順に変える」、「音量を上げる」、「画面の輝度を上げる」というコマンドにより動作が実行される前の、「チャンネル」、「音量」、「画面の輝度」の状態を、表示画面に表示するよう制御する。

【 0 1 3 7 】

図 3 0 は、対象装置に関する動作の対象の状態を表示する表示画面の他の例を示す。図 3 0 には、ユーザが「おんりょう」という音声による入力操作を行った場合の表示画面の例が示されている。動作処理部 1 5 0 は、上記と同様の処理を行った上で、「音量を上げる」、「音量を下げる」というコマンドにより動作が実行される前の、「音量」の状態を、表示画面に表示するよう制御する。

【 0 1 3 8 】

[5 - 2 . 処理の流れ]

以下では、図 3 1 を用いて、本発明の第 5 の実施形態に係るコマンド生成処理について説明する。図 3 1 は、第 5 の実施形態に係るコマンド生成処理を示すフローチャートである。この中で、ステップ S 3 1 0、ステップ S 3 2 0、ステップ S 3 3 0、ステップ S 3 4 0、ステップ S 3 5 0 およびステップ S 3 6 0 は、[1 - 2 . 処理の流れ] にて説明した第 1 の実施形態に係るコマンド生成処理と同じである。よって、ここでは新たに追加されているステップ S 4 1 0、ステップ S 4 5 0、ステップ S 4 6 0、ステップ S 4 7 0、ステップ S 4 8 0 およびステップ S 4 9 0 を中心に説明する。

【 0 1 3 9 】

ステップ S 4 1 0 において、動作処理部 1 5 0 は、コマンドを生成するための 2 種類の意味情報のうち一方の意味情報が認識されているかを判定する。ここで、一方の意味情報が認識されていれば、処理はステップ S 4 5 0 へ進む。一方、いずれの意味情報も認識されていない場合には、本処理を終了する。

【 0 1 4 0 】

ステップ S 4 5 0 において、動作処理部 1 5 0 は、意味情報を認識するための他方の入力情報が存在するかを、時系列管理部 1 8 0 に確認する。ここで、他方の入力情報が既に存在していれば、処理はステップ S 4 8 0 へ進む。一方、他方の入力情報が未だ存在していなければ、処理はステップ S 4 6 0 へ進む。

【 0 1 4 1 】

ステップ S 4 6 0 において、動作処理部 1 5 0 は、認識されていない意味情報の候補を特定し、当該候補の意味情報を意味する入力操作を、対象装置またはその他の装置の表示画面に表示するように制御する。

【 0 1 4 2 】

ステップ S 4 7 0 において、例えばユーザがさらなる入力操作を所定の時間内に行うと、音声入力情報取得部 1 1 0 またはジェスチャ入力情報取得部 1 2 0 は、当該入力操作により音声入力情報またはジェスチャ入力情報を取得する。

【 0 1 4 3 】

ステップ S 4 8 0 において、音声認識部 1 3 0 またはジェスチャ認識部 1 4 0 は、取得された音声入力情報またはジェスチャ入力情報から、上記他方の意味情報を認識する。

10

【 0 1 4 4 】

ステップ S 4 9 0 において、動作処理部 1 5 0 は、他方の意味情報が認識されているかを判定する。ここで、他方の意味情報が認識されていれば、処理ステップはステップ S 3 4 0 へ進む。一方、他方の意味情報が認識されていない場合には、本処理を終了する。

【 0 1 4 5 】

< 6 . 発明の各実施形態に係る情報処理装置のハードウェア構成 >

次に、図 3 2 を参照しながら、本発明の各実施形態に係る情報処理装置 1 0 0 のハードウェア構成について、詳細に説明する。図 3 2 は、本発明の各実施形態に係る情報処理装置 1 0 0 のハードウェア構成の一例を示すブロック図である。

20

【 0 1 4 6 】

情報処理装置 1 0 0 は、主に、CPU 9 0 1 と、ROM 9 0 3 と、RAM 9 0 5 と、を備える。また、情報処理装置 1 0 0 は、更に、ホストバス 9 0 7 と、ブリッジ 9 0 9 と、外部バス 9 1 1 と、インターフェース 9 1 3 と、入力装置 9 1 5 と、出力装置 9 1 7 と、ストレージ装置 9 1 9 と、ドライブ 9 2 1 と、接続ポート 9 2 3 と、通信装置 9 2 5 とを備える。

【 0 1 4 7 】

CPU 9 0 1 は、演算処理装置および制御装置として機能し、ROM 9 0 3、RAM 9 0 5、ストレージ装置 9 1 9、またはリムーバブル記録媒体 9 2 7 に記録された各種プログラムに従って、情報処理装置 1 0 0 内の動作全般またはその一部を制御する。ROM 9 0 3 は、CPU 9 0 1 が使用するプログラムや演算パラメータ等を記憶する。RAM 9 0 5 は、CPU 9 0 1 が使用するプログラムや、プログラムの実行において適宜変化するパラメータ等を一次記憶する。これらは CPU バス等の内部バスにより構成されるホストバス 9 0 7 により相互に接続されている。

30

【 0 1 4 8 】

ホストバス 9 0 7 は、ブリッジ 9 0 9 を介して、PCI (Peripheral Component Interconnect / Interface) バス等の外部バス 9 1 1 に接続されている。

【 0 1 4 9 】

入力装置 9 1 5 は、例えば、マイク、カメラ等ユーザの入力操作から入力情報を取得するための手段である。また、入力装置 9 1 5 は、例えば、マウス、キーボード、タッチパネル、ボタン、スイッチおよびレバー等ユーザが操作する操作手段である。また、入力装置 9 1 5 は、例えば、赤外線やその他の電波を利用したリモートコントロール手段（いわゆる、リモコン）であってもよいし、情報処理装置 1 0 0 の操作に対応した携帯電話や PDA 等の外部接続機器 9 2 9 であってもよい。さらに、入力装置 9 1 5 は、例えば、上記の操作手段を用いてユーザにより入力された情報に基づいて入力信号を生成し、CPU 9 0 1 に出力する入力制御回路等から構成されている。情報処理装置 1 0 0 のユーザは、この入力装置 9 1 5 を操作することにより、情報処理装置 1 0 0 に対して各種のデータを入力したり処理動作を指示したりすることができる。

40

【 0 1 5 0 】

50

出力装置 917 は、取得した情報をユーザに対して視覚的または聴覚的に通知することが可能な装置で構成される。このような装置として、CRTディスプレイ装置、液晶ディスプレイ装置、プラズマディスプレイ装置、ELディスプレイ装置およびランプ等の表示装置や、スピーカおよびヘッドホン等の音声出力装置や、プリンタ装置、携帯電話、ファクシミリ等がある。出力装置 917 は、例えば、情報処理装置 100 が行った各種処理により得られた結果を出力する。具体的には、表示装置は、情報処理装置 100 が行った各種処理により得られた結果を、テキストまたはイメージで表示する。他方、音声出力装置は、再生された音声データや音響データ等からなるオーディオ信号をアナログ信号に変換して出力する。

【0151】

ストレージ装置 919 は、情報処理装置 100 の記憶部の一例として構成されたデータ格納用の装置である。ストレージ装置 919 は、例えば、HDD (Hard Disk Drive) 等の磁気記憶部デバイス、半導体記憶デバイス、光記憶デバイス、または光磁気記憶デバイス等により構成される。このストレージ装置 919 は、CPU 901 が実行するプログラムや各種データ、および入力装置 915 または外部から取得した音響信号データや画像信号データ等を格納する。

【0152】

ドライブ 921 は、記録媒体用リダライタであり、情報処理装置 100 に内蔵、あるいは外付けされる。ドライブ 921 は、装着されている磁気ディスク、光ディスク、光磁気ディスク、または半導体メモリ等のリムーバブル記録媒体 927 に記録されている情報を読み出して、RAM 905 に出力する。また、ドライブ 921 は、装着されている磁気ディスク、光ディスク、光磁気ディスク、または半導体メモリ等のリムーバブル記録媒体 927 に記録を書き込むことも可能である。リムーバブル記録媒体 927 は、例えば、DVDメディア、HD-DVDメディア、Blu-rayメディア等である。また、リムーバブル記録媒体 927 は、コンパクトフラッシュ (登録商標) (Compact Flash: CF)、フラッシュメモリ、または、SDメモリカード (Secure Digital memory card) 等であってもよい。また、リムーバブル記録媒体 927 は、例えば、非接触型 IC チップを搭載した IC カード (Integrated Circuit card) または電子機器等であってもよい。

【0153】

接続ポート 923 は、機器を情報処理装置 100 に直接接続するためのポートである。接続ポート 923 の一例として、USB (Universal Serial Bus) ポート、IEEE 1394 ポート、SCSI (Small Computer System Interface) ポート等がある。接続ポート 923 の別の例として、RS-232C ポート、光オーディオ端子、HDMI (High-Definition Multimedia Interface) ポート等がある。この接続ポート 923 に外部接続機器 929 を接続することで、情報処理装置 100 は、外部接続機器 929 から直接音響信号データや画像信号データを取得したり、外部接続機器 929 に音響信号データや画像信号データを提供したりする。

【0154】

通信装置 925 は、例えば、通信網 931 に接続するための通信デバイス等で構成された通信インターフェースである。通信装置 925 は、例えば、有線または無線 LAN (Local Area Network)、Bluetooth (登録商標)、または WUSB (Wireless USB) 用の通信カード等である。また、通信装置 925 は、光通信のルータ、ADSL (Asymmetric Digital Subscriber Line) 用のルータ、または、各種通信のモデム等であってもよい。この通信装置 925 は、例えば、インターネットや他の通信機器との間で、例えば TCP/IP 等の所定のプロトコルに則して信号等を送受信することができる。また、通信装置 925 に接続される通信網 931 は、有線または無線によって接続されたネットワーク等により構成され、例えば、インターネット、家庭内 LAN、赤外線通信、ラジオ波通信または衛

10

20

30

40

50

星通信等であってもよい。

【0155】

以上、本発明の実施形態に係る情報処理装置100の機能を実現可能なハードウェア構成の一例を示した。上記の各構成要素は、汎用的な部材を用いて構成されていてもよいし、各構成要素の機能に特化したハードウェアにより構成されていてもよい。従って、本実施形態を実施する時々の技術レベルに応じて、適宜、利用するハードウェア構成を変更することが可能である。

【0156】

<7.まとめ>

ここまで、図1~図32を用いて、本発明の各実施形態について説明した。第1の実施形態によれば、様々な効果が得られ得る。まず、2種類以上の入力操作を組合せることにより、ユーザが覚えなければならない入力操作の数を減らすことが可能となる。例えば、音声による入力操作と手の動きによる入力操作が組合せられる場合に、ユーザは、5個の音声による入力操作および5個の手の動きによる入力操作の計10個の入力操作を覚えることにより、最大でその組合せの数である25個のコマンドを生成させ得る。一方で、手の動きによる入力操作のみが用いられる場合に、ユーザが25個のコマンドを生成させるために、ユーザは25個の手の動きによる入力操作を覚えなければならない。

10

【0157】

また、2種類以上の入力操作を組合せることにより、入力操作の種類毎の入力パターン数が減るため、入力操作から意図していない入力パターンが特定され、その結果意図していない意味情報が認識されるという誤入力の可能性を低減し得る。

20

【0158】

また、例えば1種類の入力操作が動作の内容を示す意味情報を意味し、別の1種類の入力操作が動作の対象を示す意味情報を意味する場合に、ユーザは、各入力操作が意味し得る意味情報を想定しやすいため、入力操作をより容易に覚え得る。

【0159】

また、例えば複数の入力パターンに同一の意味情報が関連付けられる場合に、ユーザが必ず覚えなければならない入力操作は減るため、ユーザが入力操作を覚える負担は軽減され得る。

【0160】

また、第2の実施形態によれば、第1の実施形態に関する上記効果に加えて、ユーザは、入力操作により、対象装置に単に所定の動作を実行させるだけでなく、対象装置に所望の実行量で動作を実行させ得る。これにより、簡単な入力操作により、より細かい操作指示を示すコマンドを生成でき、対象装置をより正確に動作させることができる。

30

【0161】

また、第3の実施形態によれば、第1の実施形態に関する上記効果に加えて、各ユーザは容易に入力操作を行い得る。例えば、ユーザID毎に予め定められた入力パターンを用いる場合、あるいはユーザID毎に予め定められた優先度を用いる場合に、ユーザの特性を考慮してコマンドが生成されるため、ユーザが使用しない入力操作が誤って特定され、意図しない意味情報が認識される可能性が低減され得る。またユーザが使用する入力操作が正しく特定され、意図した意味情報が認識される可能性が高まり得る。

40

【0162】

また、第4の実施形態によれば、第1の実施形態に関する上記効果に加えて、ユーザは入力操作の一部を省略し得る。これにより、ユーザにとっての入力操作の負担が軽減され得る。

【0163】

また、第5の実施形態によれば、第1の実施形態に関する上記効果に加えて、ユーザは、入力操作の一部を行うと、コマンドを生成するための残りの入力操作を把握し得る。また、ユーザは、入力操作の一部を行うと、コマンドにより動作が実行される前における当該動作の対象の状態を把握し得る。したがって、ユーザは次の入力操作のための参考情報

50

を得ることができるため、ユーザにとっての利便性が向上し得る。

【0164】

なお、上記第1～第5の実施形態において、各部の動作は互いに関連しており、互いの関連を考慮しながら、一連の動作及び一連の処理として置き換えることができる。これにより、情報処理装置の実施形態を、情報処理装置のコマンド生成方法の実施形態及び情報処理装置が有する機能をコンピュータに実現させるためのプログラムの実施形態とすることができる。

【0165】

以上、添付図面を参照しながら本発明の好適な実施形態について詳細に説明したが、本発明はかかる例に限定されない。本発明の属する技術の分野における通常の知識を有する者であれば、特許請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本発明の技術的範囲に属するものと了解される。

【0166】

例えば、各実施形態では、入力情報から意味情報を認識するために、予め入力操作をパターン化した入力パターンを用いる例が説明されたが、本発明はかかる例に限定されない。情報処理装置は、入力情報から直接的に意味情報を認識し、または別の種類の情報を介して入力情報から意味情報を認識してもよい。

【0167】

また、各実施形態では、入力パターン、意味情報、コマンド等の情報が情報処理装置内に記憶されている例が説明されたが、本発明はかかる例に限定されない。各情報は、情報処理装置に接続された別の装置に記憶され、情報処理装置は、適宜当該別の装置から各情報を取得してもよい。

【0168】

また、各実施形態では、2種類以上の入力操作として音声による入力操作と体の一部若しくは全部の動きまたは状態による入力操作が用いられているが、本発明はかかる例に限定されない。2種類の入力操作ではなく、3種類以上の入力操作が用いられてもよい。また、例えばリモートコントローラ、マウス、キーボード、タッチパネル等の、音声または体の一部若しくは全部の動きまたは状態以外を用いた入力操作が用いられてもよい。

【0169】

また、各実施形態は、理解しやすくするために分けて説明されているが、本発明はかかる例に限定されない。各実施形態は適宜組合せてもよい。例えば、第2の実施形態と第3の実施形態を組合せて、情報処理装置は、変化量変換部および個人判別部の両方を有してもよい。この場合に、例えば、変化量記憶部はユーザ毎の変化量変換辞書を記憶してもよく、変化量変換部は、特定されたユーザIDに応じて動作の実行量を示す実行量情報を認識してもよい。

【符号の説明】

【0170】

100	情報処理装置	
110	音声入力情報取得部	40
120	ジェスチャ入力情報取得部	
130	音声認識部	
132	音声記憶部	
140	ジェスチャ認識部	
142	ジェスチャ記憶部	
150	動作処理部	
152	コマンド記憶部	
160	変化量変換部	
170	個人判別部	

10

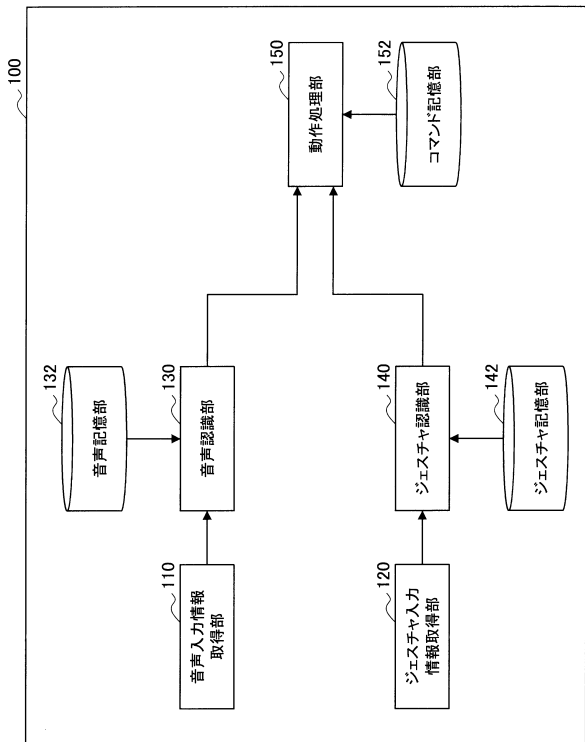
20

30

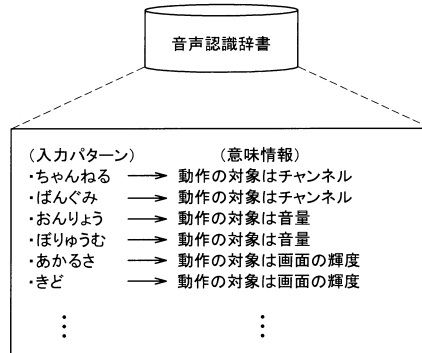
40

50

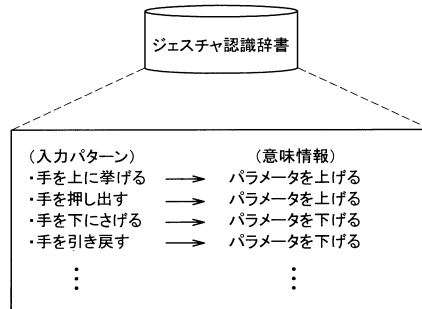
【図1】



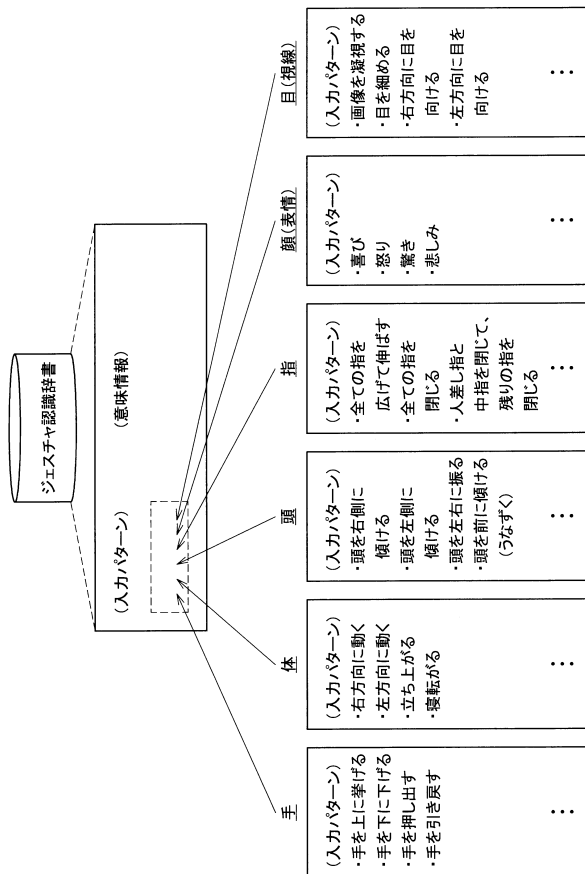
【図2】



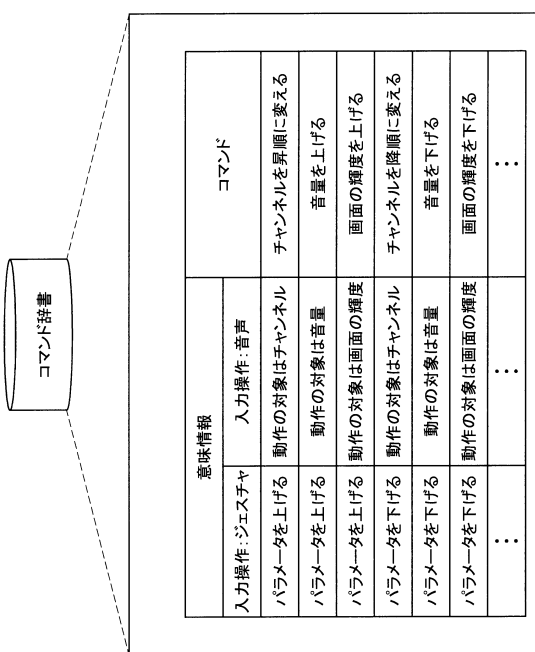
【図3】



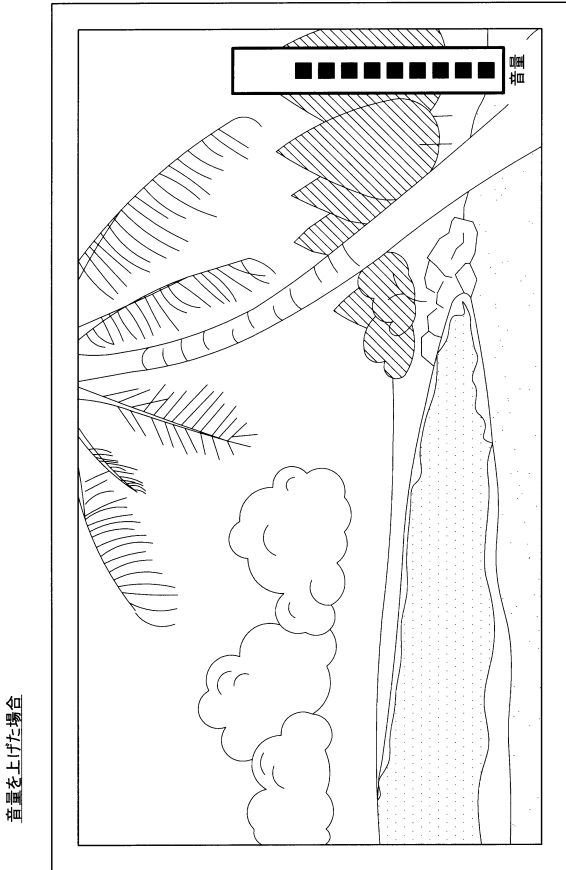
【図4】



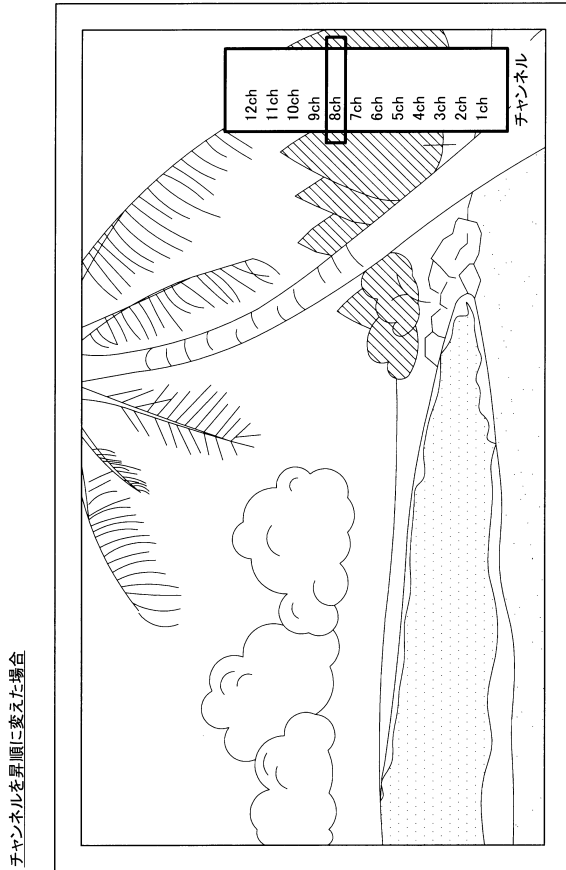
【図5】



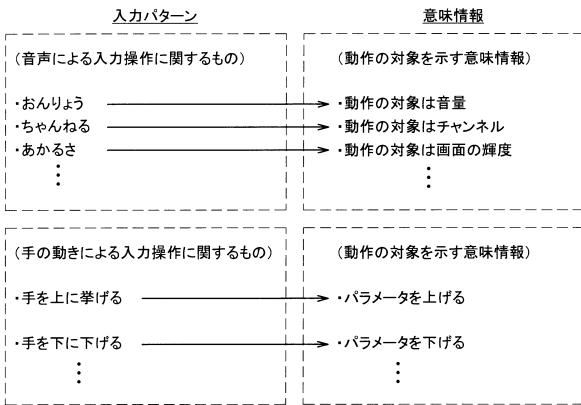
【図6】



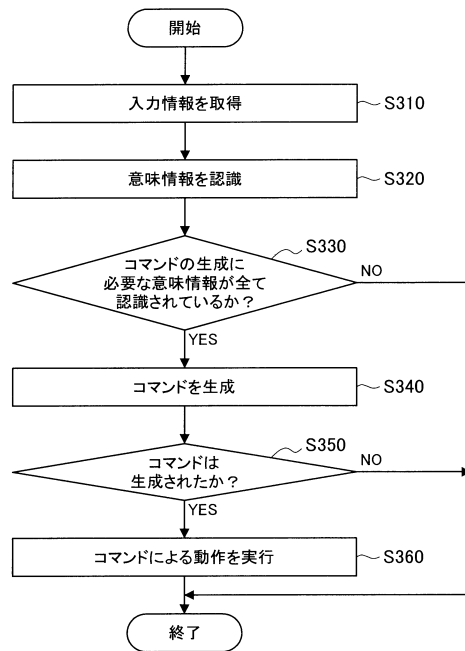
【図7】



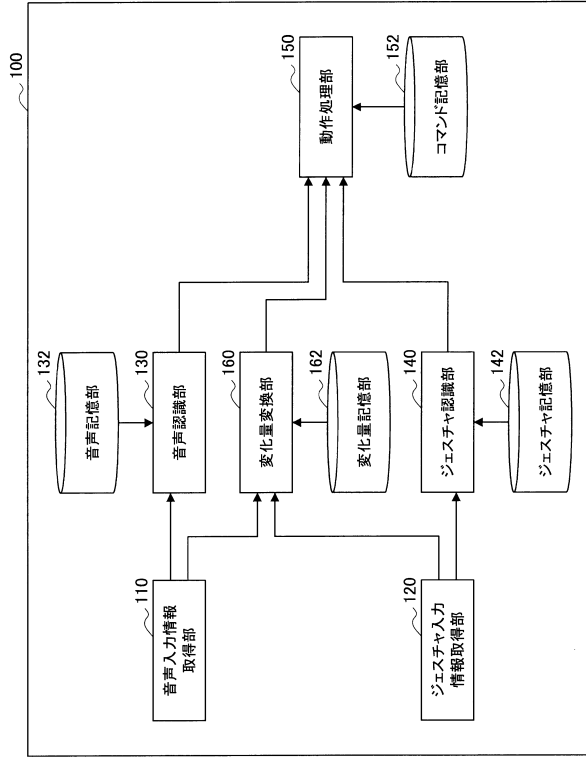
【図8】



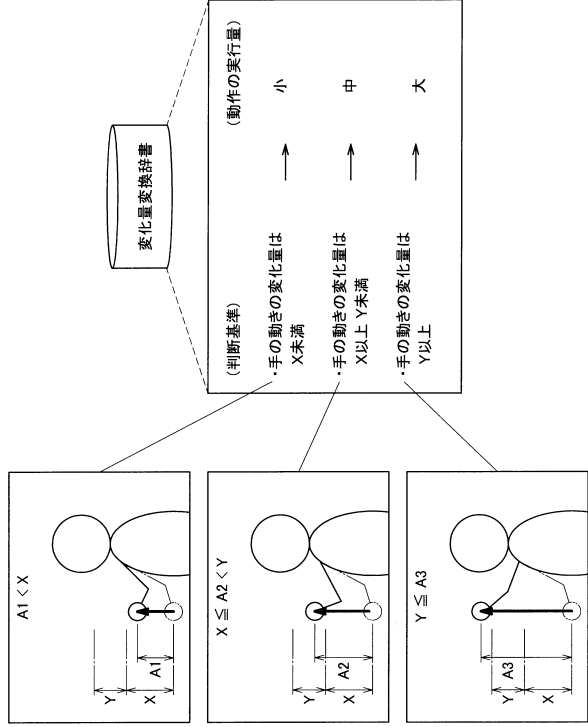
【図9】



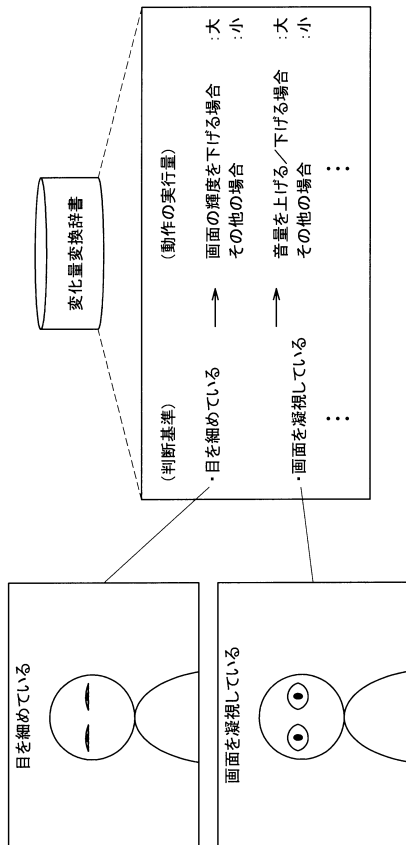
【図10】



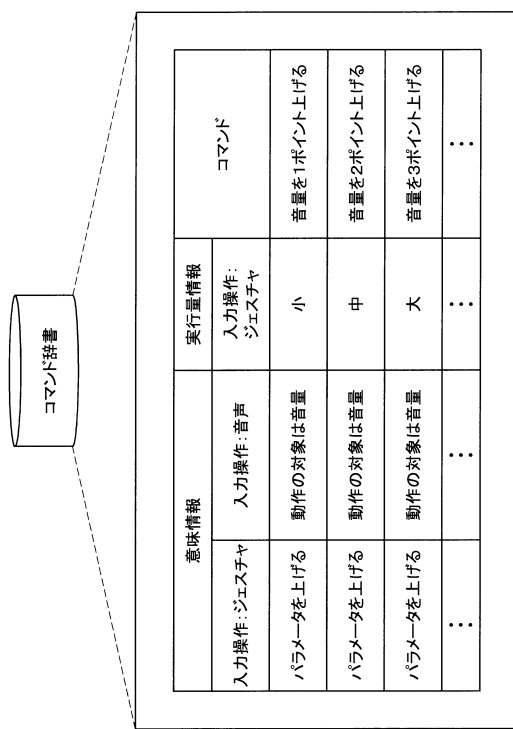
【図11】



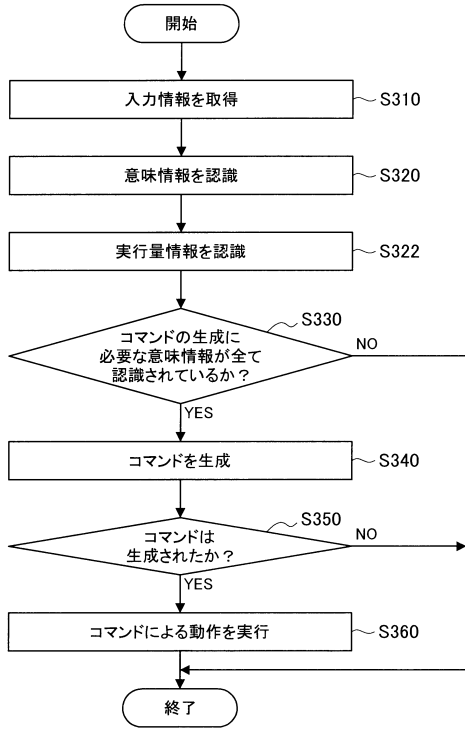
【図12】



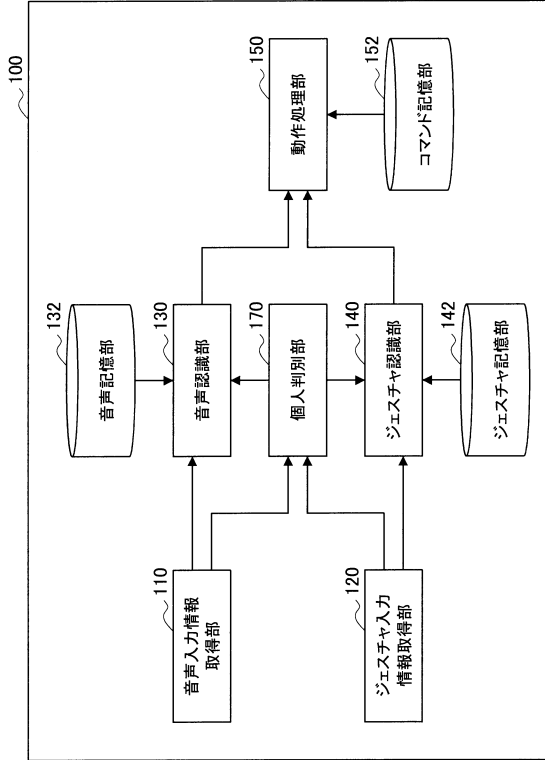
【図13】



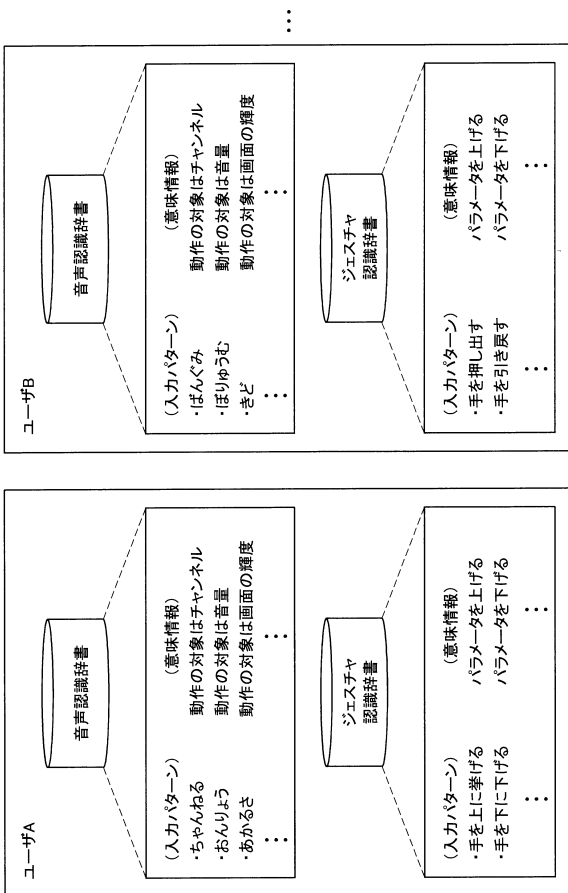
【図14】



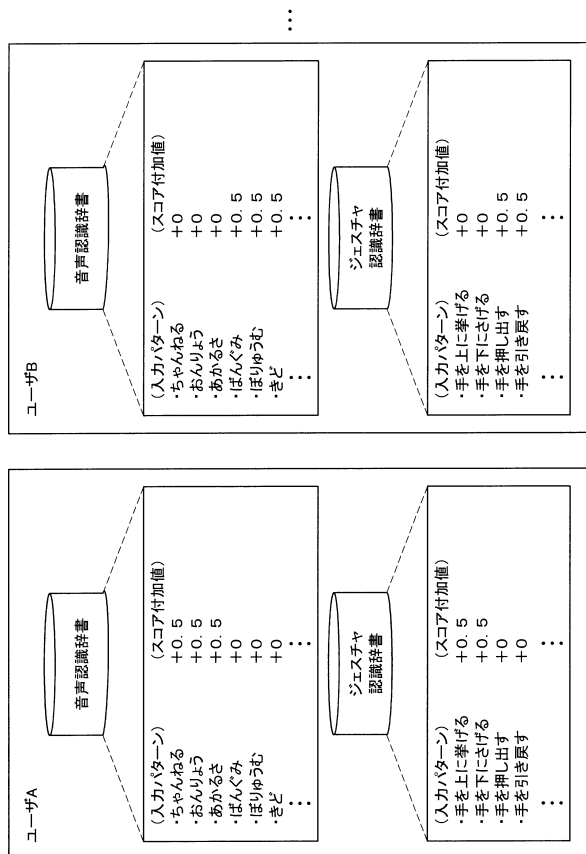
【図15】



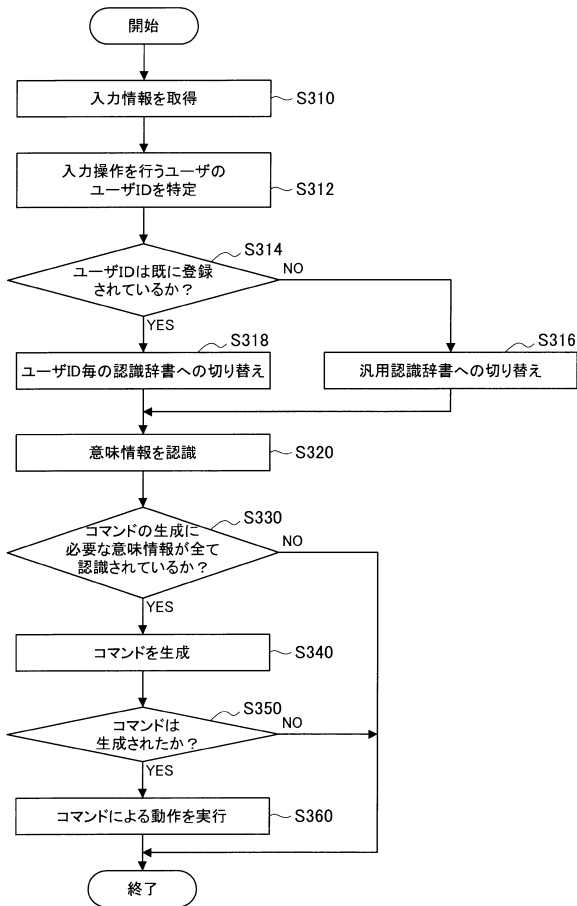
【図16】



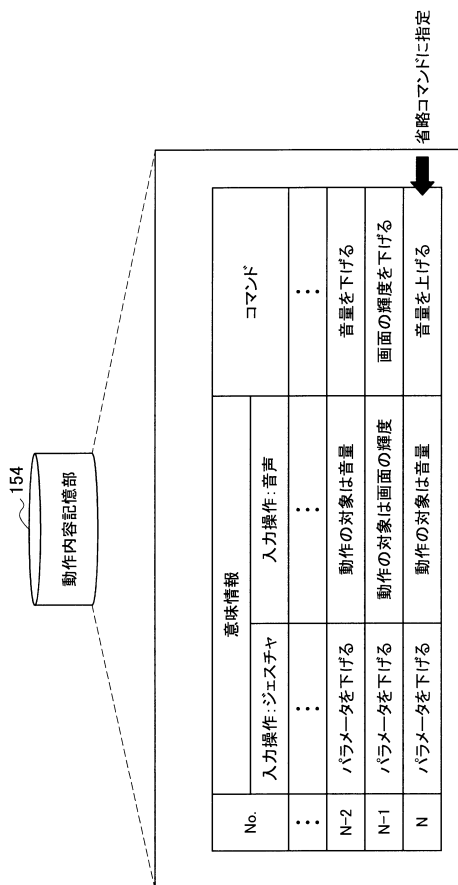
【図17】



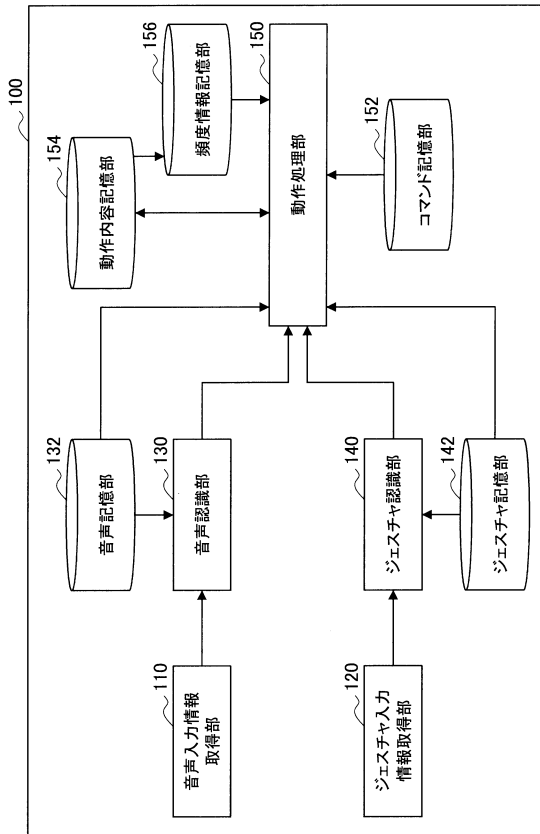
【図18】



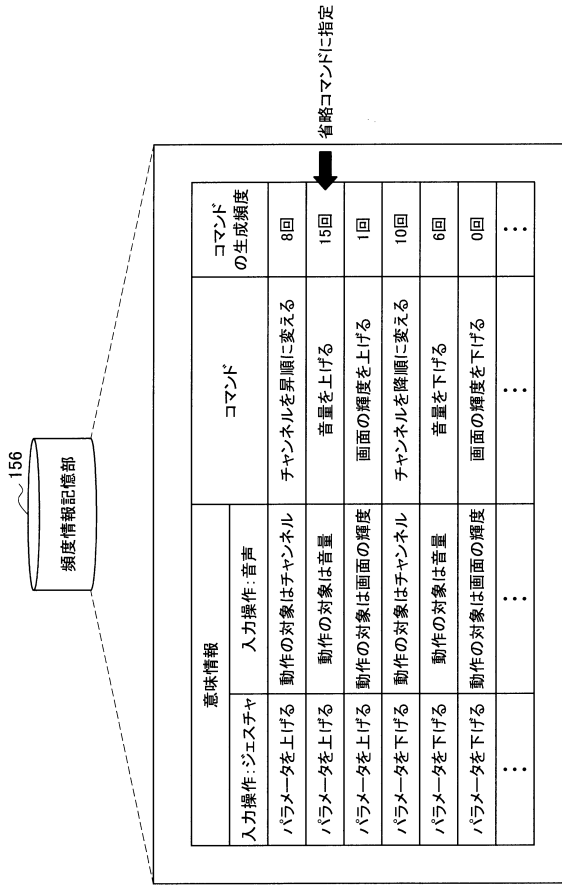
【図20】



【図19】



【図21】



【図 2 2】

意味情報		コマンド	省略対象 識別情報
入力操作:ジェスチャ	入力操作:音声		
パラメータを上げる	動作の対象はチャンネル	チャンネルを昇順に変える	
パラメータを上げる	動作の対象は音量	音量を上げる	○
パラメータを上げる	動作の対象は画面の輝度	画面の輝度を上げる	
パラメータを下げる	動作の対象はチャンネル	チャンネルを降順に変える	
パラメータを下げる	動作の対象は音量	音量を下げる	
パラメータを下げる	動作の対象は画面の輝度	画面の輝度を下げる	
⋮	⋮	⋮	⋮

コマンド辞書

【図 2 3】

どの操作コマンドについて片方の入力を省略しますか？

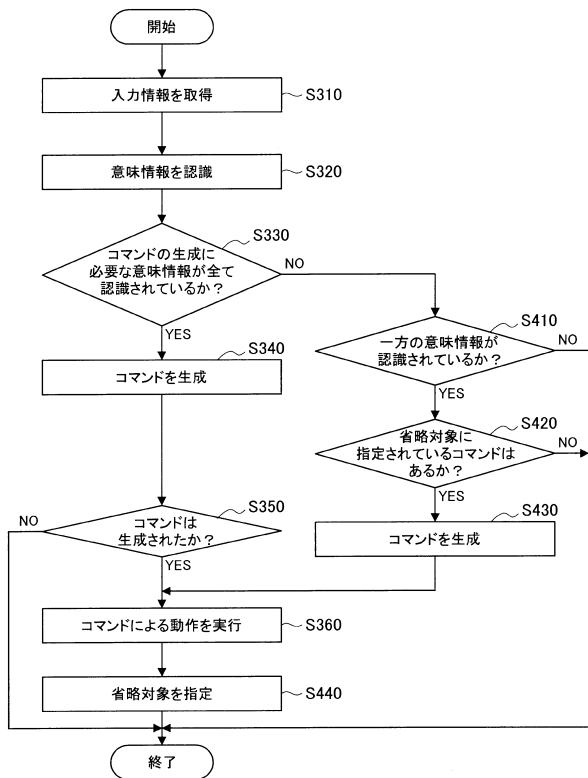
入力操作		コマンド	コマンド の生成頻度
ジェスチャ	音声		
手を上に挙げる	音量	音量を上げる	15回
手を下にさげる	チャンネル	チャンネルを降順に変える	10回
手を上に挙げる	チャンネル	チャンネルを昇順に変える	8回
手を下にさげる	音量	音量を下げる	6回
手を上に挙げる	明るさ	画面の輝度を上げる	1回
手を下にさげる	明るさ	画面の輝度を下げる	0回

【図 2 4】

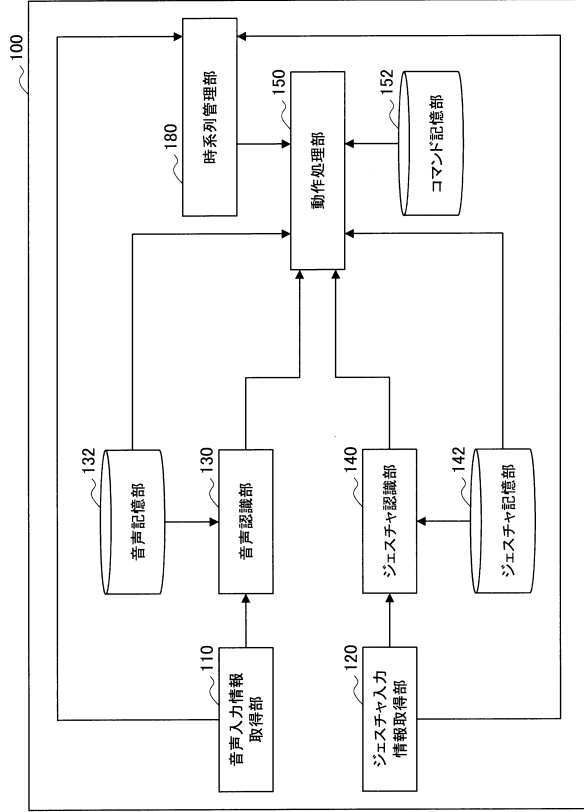


音量を上げた場合

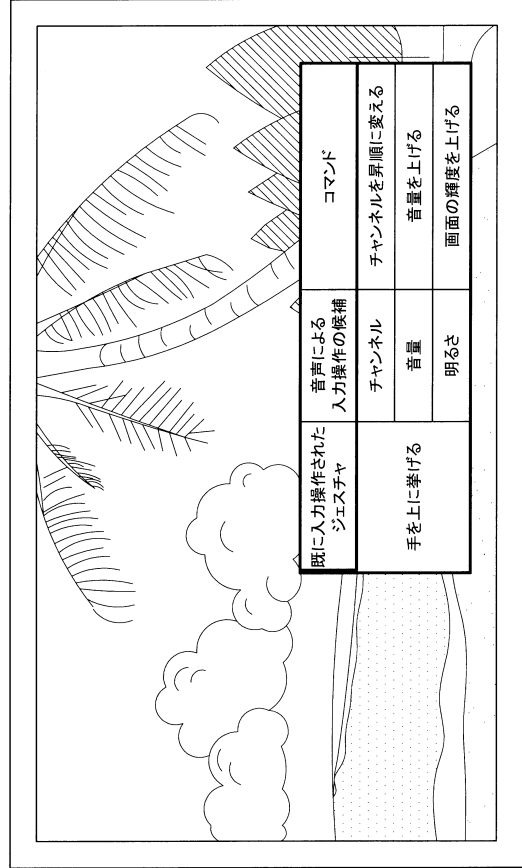
【図 2 5】



【図 26】

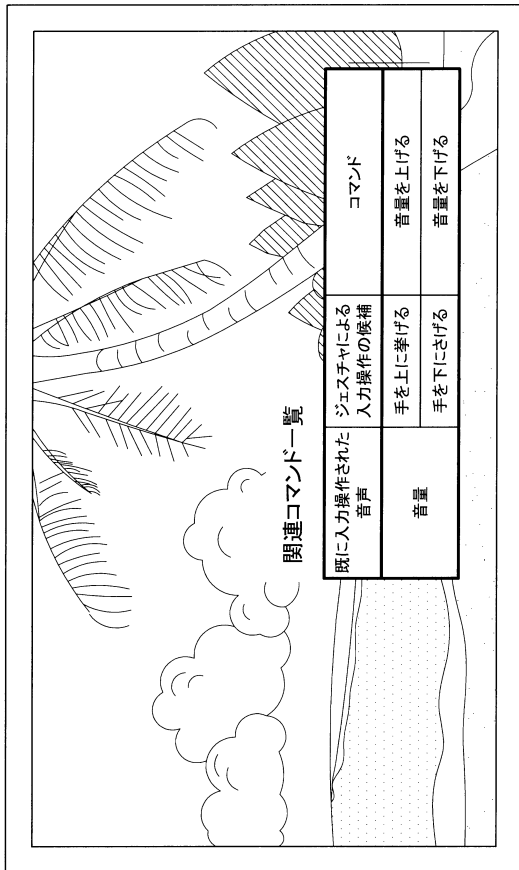


【図 27】



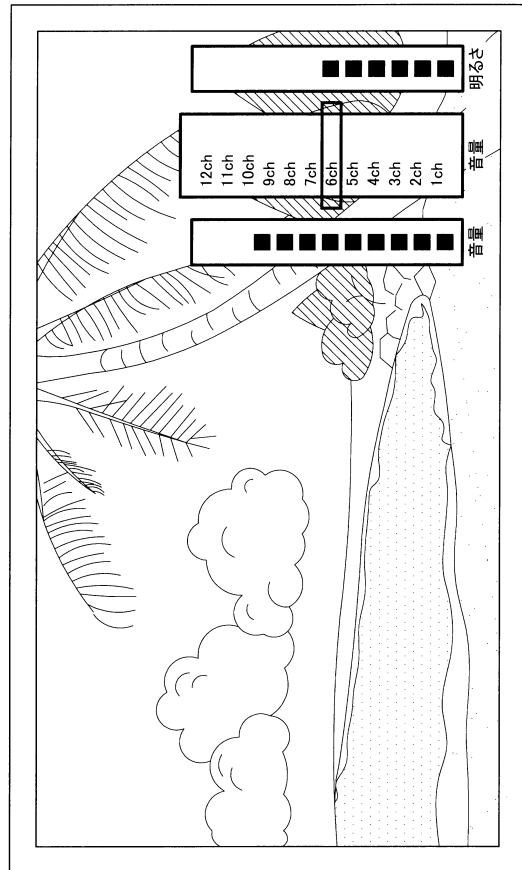
「手を上に挙げる」ジェスチャ(手の動き)を行った場合

【図 28】



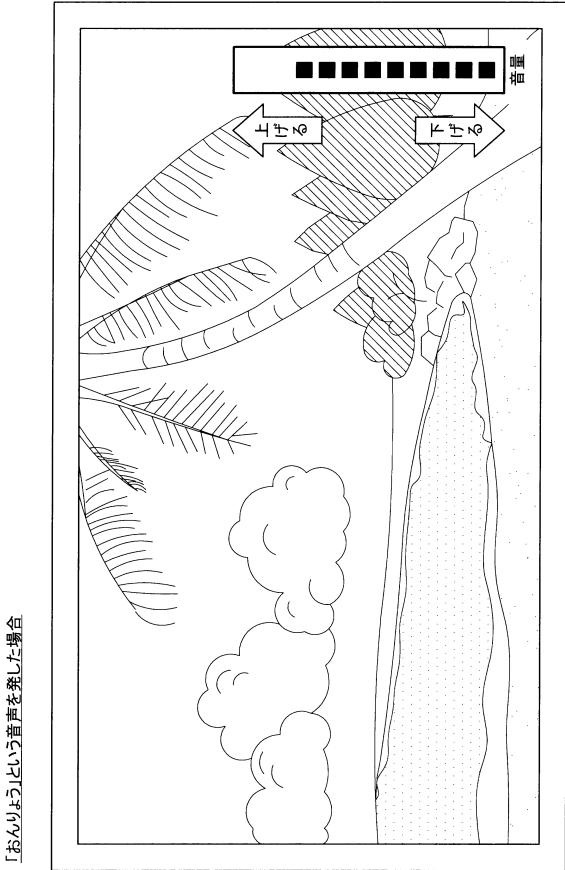
「おんりょう」という音声を発した場合

【図 29】



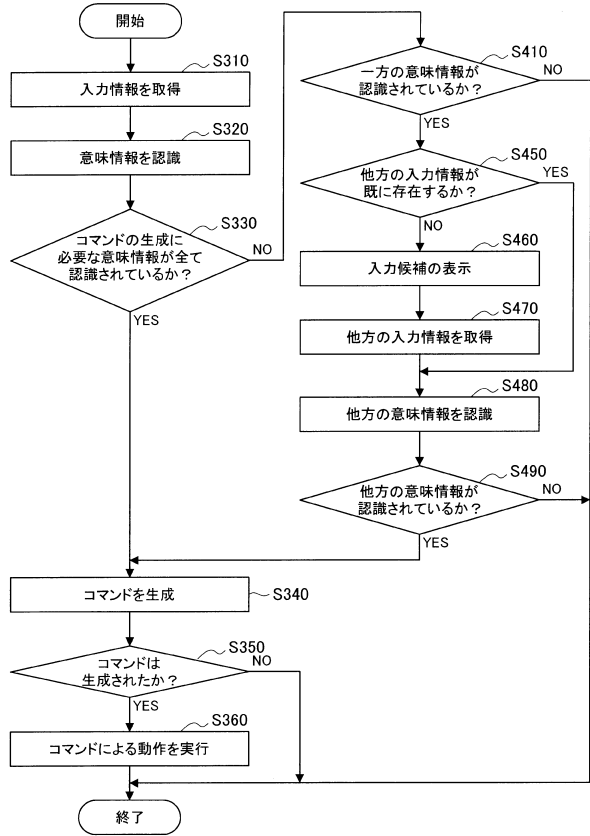
「手を上に挙げる」ジェスチャ(手の動き)を行った場合

【図30】

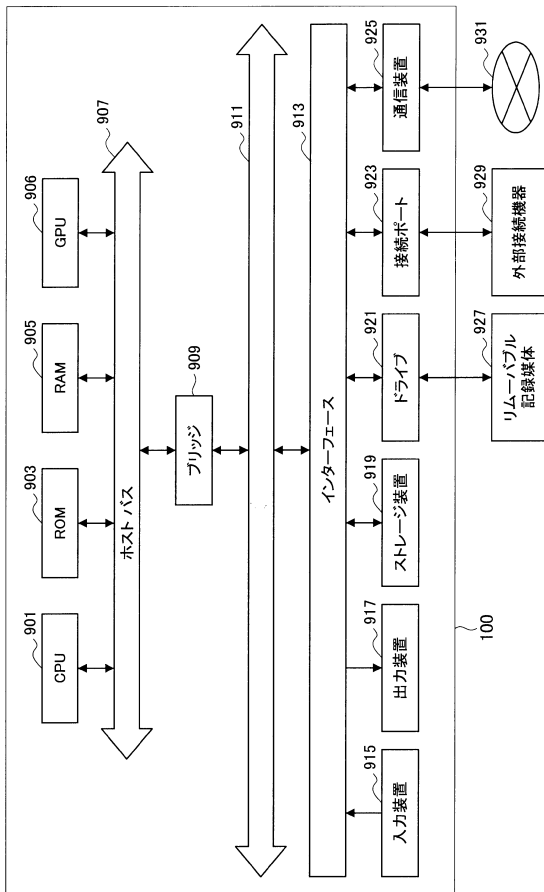


「おんりよう」という音声を発した場合

【図31】



【図32】



フロントページの続き

審査官 菅原 浩二

- (56)参考文献 特開2005-178473(JP,A)
特開2010-157192(JP,A)
特開2002-251235(JP,A)
特開2002-091492(JP,A)
特開2000-339305(JP,A)
再公表特許第2008/149482(JP,A1)
特開平10-031551(JP,A)
米国特許出願公開第2008/0252595(US,A1)
特開2007-283968(JP,A)
再公表特許第03/025859(JP,A1)
特開平11-288296(JP,A)
特開2008-180833(JP,A)
特開2008-293252(JP,A)
特開2002-062962(JP,A)
米国特許出願公開第2008/0270896(US,A1)

(58)調査した分野(Int.Cl., DB名)

G06F	3/01
G06F	3/16
H04N	5/00