

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2015-76040
(P2015-76040A)

(43) 公開日 平成27年4月20日 (2015.4.20)

(51) Int.Cl.		F I		テーマコード (参考)
G06F 3/16	(2006.01)	G06F 3/16	320B	
G10L 15/00	(2013.01)	G06F 3/16	320H	
		G06F 3/16	330E	
		G06F 3/16	330K	
		G10L 15/00	200P	

審査請求 有 請求項の数 4 O L (全 10 頁)

(21) 出願番号 特願2013-213690 (P2013-213690)
(22) 出願日 平成25年10月11日 (2013.10.11)

(71) 出願人 311012169
NECパーソナルコンピュータ株式会社
東京都千代田区外神田四丁目14番1号
秋葉原UDX
(74) 代理人 100084250
弁理士 丸山 隆夫
(72) 発明者 藤川 玲二
東京都品川区大崎一丁目11番1号 NEC
パーソナルコンピュータ株式会社内
(72) 発明者 原田 雅彦
東京都品川区大崎一丁目11番1号 NEC
パーソナルコンピュータ株式会社内

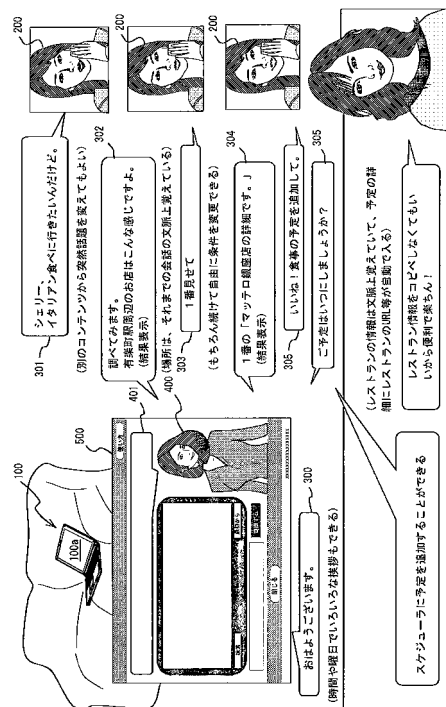
(54) 【発明の名称】 情報処理方法、情報処理装置、及びプログラム

(57) 【要約】

【課題】対話検索に引き続きスケジューラに追記可能な情報処理方法、情報処理装置、及びプログラムを提供する。

【解決手段】入力された音声情報から認識されたテキスト情報に基づいて特定される所定のコマンドを実行する情報処理方法であって、所定のコマンドおよび実行結果を記憶し、スケジュールの追加コマンドに対応するテキスト情報が認識されると、記憶された所定のコマンドおよび実行結果から時間情報と場所情報とを抽出し、抽出された時間情報と場所情報とに基づいてスケジュールを追加する。

【選択図】 図5



【特許請求の範囲】**【請求項 1】**

入力された音声情報から認識されたテキスト情報に基づいて特定される所定のコマンドを実行する情報処理方法であって、

前記所定のコマンドおよび実行結果を記憶し、

スケジュールの追加コマンドに対応するテキスト情報が認識されると、前記記憶された所定のコマンドおよび実行結果から時間情報と場所情報とを抽出し、

前記抽出された時間情報と場所情報とに基づいてスケジュールを追加することを特徴とする情報処理方法。

【請求項 2】

前記記憶された所定のコマンドおよび実行結果から時間情報または場所情報が抽出できないとき、抽出できない情報への回答を促す質問を発話し、

その後入力された音声情報から認識されたテキスト情報に基づいて特定される前記抽出できなかった時間情報または場所情報を加えてスケジュールを追加することを特徴とする請求項 1 記載の情報処理方法。

【請求項 3】

入力された音声情報から認識されたテキスト情報に基づいて特定される所定のコマンドを実行する情報処理装置であって、

前記所定のコマンドおよび実行結果を記憶する記憶手段と、

スケジュールの追加コマンドに対応するテキスト情報が認識されると、前記記憶された所定のコマンドおよび実行結果から時間情報と場所情報とを抽出する抽出手段と、

前記抽出された時間情報と場所情報とに基づいてスケジュールを追加する修正手段と、を備えたことを特徴とする情報処理装置。

【請求項 4】

入力された音声情報から認識されたテキスト情報に基づいて特定される所定のコマンドを実行する情報処理装置のコンピュータが、

記憶手段に、前記所定のコマンドおよび実行結果を記憶する手順と、

抽出手段に、スケジュールの追加コマンドに対応するテキスト情報が認識されると、前記記憶された所定のコマンドおよび実行結果から時間情報と場所情報とを抽出する手順と、

修正手段に、前記抽出された時間情報と場所情報とに基づいてスケジュールを追加する手順と、

を実行させるためのプログラム。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、情報処理方法、情報処理装置、及びプログラムに関する。

【背景技術】**【0002】**

近年、テレビ受像器やパーソナルコンピュータ等の電子機器に対するユーザ・コマンドの入力を支援する対話型操作支援システムが開発されている（例えば、特許文献 1 参照）。

【0003】

特許文献 1 に記載の発明は、「対話型操作支援システム及び対話型操作支援方法、並びに記憶媒体」に関する発明であり、具体的には、「音声合成やアニメーションによるリアクションを行なう擬人化されたアシスタントと呼ばれるキャラクターのアニメーションをユーザ・インターフェースとすることにより、ユーザに対して親しみを持たせると同時に複雑な命令への対応やサービスへの入り口を提供することができる。また、自然言語に近い感じの命令体系を備えているので、ユーザは、通常の会話と同じ感覚で機器の操作を容易に行なうことができる」ものである。

10

20

30

40

50

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2002-41276号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、特許文献1には、対話検索に引き続きスケジューラに追記することはできなかった。

そこで、本発明の目的は、対話検索に引き続きスケジューラに追記可能な情報処理方法、情報処理装置、及びプログラムを提供することにある。

10

【課題を解決するための手段】

【0006】

上記課題を解決するため、請求項1に記載の発明は、入力された音声情報から認識されたテキスト情報に基づいて特定される所定のコマンドを実行する情報処理方法であって、前記所定のコマンドおよび実行結果を記憶し、スケジュールの追加コマンドに対応するテキスト情報が認識されると、前記記憶された所定のコマンドおよび実行結果から時間情報と場所情報とを抽出し、前記抽出された時間情報と場所情報とに基づいてスケジュールを追加することを特徴とする。

20

【発明の効果】

【0007】

本発明によれば、対話検索に引き続きスケジューラに追記可能な情報処理方法、情報処理装置、及びプログラムの提供を実現できる。

【図面の簡単な説明】

【0008】

【図1】一実施形態に係る情報処理装置としてのパーソナルコンピュータ100のブロック図である。

【図2】図1に示したパーソナルコンピュータの主要部のブロック図の一例である。

【図3】(a)はスケジューラへの追記の動作を説明するためのフローチャートの一例であり、(b)はスケジューラへの追記の動作を説明するためのフローチャートの他の一例である。

30

【図4】ユーザが音声でパーソナルコンピュータに問いかけをしている状態を示す図である。

【図5】ユーザとパーソナルコンピュータとの対話によるスケジュール修正の説明図である。

【発明を実施するための形態】

【0009】

次に実施の形態について述べる。

<構成>

図1は、一実施形態に係る情報処理装置としてのパーソナルコンピュータ100のブロック図である。

40

同図に示すパーソナルコンピュータ(以下、PC)100は、マイクロフォン101、増幅回路102、104、スピーカ103、表示装置105、キーボード106、マウス107、光学読取装置108、制御手段109、HDD(Hard Disk Drive)110、ネットワーク接続部111、I/O(Input/Output)112、及びバスライン113を有する。

【0010】

マイクロフォン101は、ユーザの音声を電気信号に変換する機能を有する。マイクロフォン101としては、例えばコンデンサマイクロフォンが挙げられるが、ダイナミックマイクロフォンでもよい。

増幅回路102は、マイクロフォン101からの電気信号を増幅する回路である。

50

スピーカ 103 は、電気信号を音声に変換する機能を有する。スピーカ 103 は、主にアパターの声をユーザへ伝達する機能を有する。アパターは、本実施形態では女性であるが、限定されるものではない。

増幅回路 104 は、音声信号を、スピーカ 103 を駆動させるレベルまで増幅する回路である。

表示装置 105 は、アパターやアパターの発話内容を文字で表示した吹き出しを含む画像や文字等を表示する機能を有する。表示装置 105 としては、例えば、液晶表示素子が挙げられる。

キーボード 106 は、文字、数字、符号を入力する入力装置である。

マウス 107 は、入力装置の一種であり、机上を移動させることで表示装置 105 のカーソルを移動させる等の機能を有する。

光学読取装置 108 は、C D (Compact Disk)、D V D (Digital Versatile Disc) や C D - R (Compact Disc-Recordable) 等の光学媒体を読み取る機能を有する。

【0011】

制御手段 109 は、P C 100 を統括制御機能、及び音声処理機能を有する素子であり、例えば C P U (Central Processing Unit) が挙げられる。音声処理機能とは、主に入力した音声をテキストデータとして出力し、解析し、合成する機能である。

制御手段 109 は、それぞれソフトウェアで構成される入力制御手段 109 a、音声認識手段 109 b、音声解析手段 109 c、検索手段 109 d、音声合成手段 109 e、及び修正手段 109 f を有する。

【0012】

入力制御手段 109 a は、マイクロフォン 101 に入力された音声に変換された信号を解析して得られたコマンドに基づいて処理させる機能の他、キーボード 106 からのキー入力、及びマウス 107 からのクリックやドラッグ等による信号を文字表示、数字表示、符号表示、カーソル移動、コマンド等に変換する機能を有する。

音声認識手段 109 b は、マイクロフォン 101 からの信号をテキストデータとして出力する機能を有し、クライアント型音声認識部 203 である。

音声解析手段 109 c は、テキストデータを解析する機能を有する。音声解析手段 109 c は、ユーザから音声による問いかけがあると、その問いかけに関するテキストデータを解析する。

【0013】

検索手段 109 d は、ネットワーク 207 を介してインターネット検索する手段である。検索手段 109 d は、ユーザから検索の指示があると、予め設定されたブラウザでネットワークに接続し、予め設定されたインターネット検索サービスに接続し、キーワード検索する機能を有する。

音声合成手段 109 e は、クライアント型音声合成部 210 であり、人間の音声を人工的に作り出す機能を有する。音声はアパターの年齢性別に対応した音質が設定されている。音声合成手段 109 e の出力は、バスライン 113、及び増幅回路 104 を経て出力手段としてのスピーカ 103 から発音される。

修正手段 109 f は、対話検索の文脈上覚えておき、その文脈から詳細情報を把握しておき、スケジューラに追記する機能を有する。スケジューラの作成は、例えばユーザ 200 からの「いいね」、もしくは「追記して」の音声をトリガとして作成する。「文脈上覚えておき」とは、例えば「スケジューラに追記」をコマンドとして実行するとき利用できるよう、その前に検索した情報等（レストラン情報やイベント情報）を、所定期間一時記憶しておくことである。従って、「追記して」をトリガとして、記憶している情報を追記可能となる。

【0014】

H D D 110 は、記憶装置の一種であり、R O M (Read Only Memory) エリア、及び R A M (Random Access Memory) エリアを有する。R O M エリアは制御プログラムを格納するエリアであり、R A M エリアはメモリとして用いられるエリアである。

10

20

30

40

50

【 0 0 1 5 】

ネットワーク接続部 1 1 1 は、ネットワーク 2 0 7 を介して外部のサーバに接続する機能を有する公知の装置である。無線もしくは有線のいずれの手段を用いてもよい。

I / O 1 1 2 は、外部の電子機器、例えば U S B (Universal Serial Bus line) フラッシュメモリやプリンタを接続する機能を有する入出力装置である。

尚、P C 1 0 0 は、入力手段としてタッチパネルを有していてもよい。

【 0 0 1 6 】

図 2 は、図 1 に示したパーソナルコンピュータの主要部のブロック図の一例である。

図 2 において、本発明の実施形態における P C 1 0 0 は、マイクロフォン 1 0 1 から入力されたユーザの音声（電気信号）に変換されて、当該音声データが音声信号解釈部 2 0 2 によって解釈され、その結果がクライアント型音声認識部 2 0 3 において認識される。クライアント型音声認識部 2 0 3 は、認識した音声データをクライアントアプリケーション部 2 0 4 に渡す。

10

【 0 0 1 7 】

クライアントアプリケーション部 2 0 4 は、ユーザからの問い合わせに対する回答が、オフライン状態にあるローカルコンテンツ部 2 0 8 に格納されているか否かを確認し、ローカルコンテンツ部 2 0 8 に格納されている場合は、当該ユーザからの問い合わせに対する回答を、後述するテキスト読上部 2 0 9、クライアント型音声合成部 2 1 0 を経由して、スピーカ 1 0 3 から音声出力する。

【 0 0 1 8 】

ユーザからの問い合わせに対する回答が、ローカルコンテンツ部 2 0 8 に格納されていない場合は、P C 1 0 0 単独で回答を持ち合わせていないことになるので、インターネット等のネットワーク 2 0 7 に接続されるネットワーク接続部 1 1 1 を介して、インターネット上の検索エンジン等を用いてユーザからの問い合わせに対する回答を検索し、得られた検索結果を、テキスト読上部 2 0 9、クライアント型音声合成部 2 1 0 を経由して、スピーカ 1 0 3 から音声出力する。

20

【 0 0 1 9 】

クライアントアプリケーション部 2 0 4 は、ローカルコンテンツ部 2 0 8、又はネットワーク 2 0 7 から得られた回答をテキスト（文字）データに変換し、テキスト読上部 2 0 9 に渡す。テキスト読上部 2 0 9 は、テキストデータを読み上げ、クライアント型音声合成部 2 1 0 に渡す。クライアント型音声合成部 2 1 0 は、音声データを人間が認識可能な音声データに合成しスピーカ 1 0 3 に渡す。スピーカ 1 0 3 は、音声データ（電気信号）を音声に変換する。また、スピーカ 1 0 3 から音声を発するのに合わせて、表示装置 1 0 5 に当該音声に関連する詳細な情報を表示する。

30

【 0 0 2 0 】

< 動作 >

図 3 (a) は、スケジュールへの追記の動作を説明するためのフローチャートの一例であり、図 3 (b) は、スケジュールへの追記の動作を説明するためのフローチャートの他の一例である。

図 3 (a) において動作の主体は制御手段である。

40

音声認識が開始されると（ステップ S 1）、スケジュール追加コマンドが有るか否かを判断し（ステップ S 2）、スケジュール追加コマンドが無い場合（ステップ S 2 / N）、コマンドを実行し（ステップ S 3）、スケジュール追加コマンドがある場合（ステップ S 2 / Y）、ステップ S 5 に進む。

コマンド実行後、コマンド・実行結果を一時記憶し、ステップ S 1 に戻る（ステップ S 4）。

ステップ S 5 では一時記憶から時間・場所情報を抽出し（ステップ S 5）、スケジュール追加し終了する（ステップ S 6）。

尚、コマンドには例えばスケジュール追加、スケジュール変更、スケジュール取消等が挙げられる。また、スケジュール追加後、再度スケジュールの修正をしたい場合にはステ

50

ップ S 1 に戻ればよい。

図 3 (b) において動作の主体は制御手段である。

ステップ S 1 1 ~ S 1 4 は図 3 (a) のステップ S 1 ~ S 4 と同様のため、説明を省略し、ステップ S 1 5 から説明する。

ステップ S 1 5 では一時記憶から時間・場所情報を抽出したか否かを判断し、時間・場所情報を抽出しない場合 (ステップ S 1 5 / N)、不足情報を質問し、回答を音声認識し (ステップ S 1 6)、スケジュールを追加して終了する (ステップ S 1 7)。

時間場所情報を抽出した場合 (ステップ S 1 5 / Y)、スケジュールに追加して終了する (ステップ S 1 7)。

スケジュール追加後、再度スケジュールの修正をしたい場合にはステップ S 1 1 に戻ればよい。

10

【 0 0 2 1 】

図 4 は、ユーザが音声で P C に問いかけをしている状態を示す図である。図 5 は、ユーザと P C との対話によるスケジュール修正の説明図である。

例えば、図 4 に示すユーザ 2 0 0 がドレスラーのチェストに座ってメイクをしている場合について述べる。このときユーザ 2 0 0 は両手がふさがっており、かつ P C 1 0 0 から離れている。ユーザ 2 0 0 はメイクを続けながらソファーに載置された P C 1 0 0 に対し、特定のキーワードとしてのウェイクアップキーワードである「シェリー」と言うと、P C 1 0 0 は、判別手段としての制御手段が判別し、コマンドとしての問いかけに対する応答動作を開始し、例えば「お呼びでしょうか？」と返事をする。

20

P C 1 0 0 のモニタ 1 0 0 a には、図 5 に示すようなアバター 4 0 0 及びアバター 4 0 0 の吹き出し 4 0 1 を含むウィンドウ 5 0 0 が最大限のサイズで表示される。

【 0 0 2 2 】

ここで、アバター 4 0 0 は、吹き出し 3 0 0 のような「おはようございます。」等の挨拶を発音するように設定されている。挨拶は時間や曜日で異なるように設定されている。

【 0 0 2 3 】

ユーザ 2 0 0 はドレスラーの前でメイクを続けながら、ソファー上の P C 1 0 0 に対し、「シェリー、イタリアン食べに行きたいんだけど。」3 0 1 と問いかける。すなわちユーザ 2 0 0 は、両手がふさがった状態であっても音声による検索の要求を行うことができる (このとき別のコンテンツから突然話題をかえてもよい)。

30

この要求に対して、P C 1 0 0 からアバター 4 0 0 に対応した音声で「調べてみます。有楽町駅周辺のお店はこんな感じですよ。」3 0 2 と応答すると共に、モニタ 1 0 0 a に検索結果を表示する (有楽町駅はユーザ 2 0 0 の話の中に出現した場所であり、P C 1 0 0 が記憶しているものとする。)。

検索結果がモニタ 1 0 0 a に表示されていることを確認するため、ユーザ 2 0 0 はドレスラーからソファーまで歩いて移動し、モニタ 1 0 0 a を見ているものとする。

ユーザ 2 0 0 はモニタ 1 0 0 a に表示された検索結果に対して「1 番見せて」3 0 3 と P C 1 0 0 に問いかける (続けて自由に条件を変更できる)。

P C 1 0 0 は、ユーザ 2 0 0 の問いかけに対し、「1 番の『マッテロ銀座店の詳細です。』」と発音しながら、モニタ 1 0 0 a に結果を表示する。

40

この結果の表示に対して、ユーザ 2 0 0 は「いいね！食事の予定を追加して。」3 0 6 と言うと、P C 1 0 0 は、ユーザ 2 0 0 からの検索の内容がユーザ 2 0 0 の外出を伴うと判断し、スケジュールに追記するものと判断し、「ご予約はいつにしましょうか？」3 0 5 とユーザ 2 0 0 に質問する。この後、ユーザ 2 0 0 は P C 1 0 0 に予定を言うと、P C 1 0 0 はスケジュールに予定を追加する。尚、レストランの情報は文脈上覚えていて、予定の詳細にレストランの U R L (Uniform Resource Locator 等の情報が自動で追記される)。

以上において、本実施形態によれば、対話検索に引き続きスケジュールへの追記が可能となる。

【 0 0 2 4 】

50

< プログラム >

以上で説明した本発明に係る情報装置は、コンピュータで処理を実行させるプログラムによって実現されている。コンピュータとしては、例えばパーソナルコンピュータやワークステーションなどの汎用的なものが挙げられるが、本発明はこれに限定されるものではない。よって、一例として、プログラムにより本発明の機能を実現する場合の説明を以下で行う。

【 0 0 2 5 】

例えば、

情報処理装置のコンピュータに、

入力された音声情報に予め定められたキーワードが含まれるか否かを判別する手順と、
音声情報から所定のテキスト情報を認識する手順と、
認識する手順により認識された所定のテキスト情報に基づいて特定される所定のコマンドを実行する手順と、

判別する手順により予め定められたキーワードが含まれると判別されたとき、コマンドを実行する手順を起動させる処理と、
を実行させるためのプログラムであって、

コンピュータが、

表示手段に、ナビゲータ画像を含むウィンドウと、検索結果と、を表示する手順と、
検索手段に、ユーザから音声による問いかけがあると、前記ナビゲータ画像に対応した音声で応答すると共に問いかけに関する検索を開始する手順と、

修正手段に、検索結果が得られると、修正指示がある場合にはユーザのスケジュールに修正を行う手順と、
を実行させるためのプログラムが挙げられる。

また、

入力された音声情報から認識されたテキスト情報に基づいて特定される所定のコマンドを実行する情報処理装置のコンピュータが、

記憶手段に、所定のコマンドおよび実行結果を記憶する手順と、

抽出手段に、スケジュールの追加コマンドに対応するテキスト情報が認識されると、前記記憶された所定のコマンドおよび実行結果から時間情報と場所情報とを抽出する手順と

、
修正手段に、抽出された時間情報と場所情報とに基づいてスケジュールを追加する手順と、
を実行させるためのプログラムでもよい。

【 0 0 2 6 】

これにより、プログラムが実行可能なコンピュータ環境さえあれば、どこにおいても本発明にかかる情報処理装置を実現することができる。

このようなプログラムは、コンピュータに読み取り可能な記憶媒体に記憶されていてもよい。

【 0 0 2 7 】

< 記憶媒体 >

ここで、記憶媒体としては、例えば C D - R O M、フレキシブルディスク (F D)、C D - R 等のコンピュータで読み取り可能な記憶媒体、フラッシュメモリ、R A M、R O M、F e R A M 等の半導体メモリや H D D が挙げられる。

【 0 0 2 8 】

C D - R O M は、Compact Disc Read Only Memory の略である。フレキシブルディスクは、Flexible Disk : F D を意味する。F e R A M は、Ferroelectric RAM の略で、強誘電体メモリを意味する。

【 0 0 2 9 】

以上において、本発明によれば、入力された音声情報に予め定められたキーワードが含まれるか否かを判別する判別手段と、音声情報から所定のテキスト情報を認識する音声認

10

20

30

40

50

識手段と、音声認識手段により認識された所定のテキスト情報に基づいて特定される所定のコマンドを実行するコマンド実行手段と、判別手段により予め定められたキーワードが含まれると判別されたとき、コマンド実行手段を起動させる起動手段と、を備え、ナビゲータ画像を含むウィンドウと、検索結果と、を表示する表示手段と、ユーザから音声による問いかけがあると、ナビゲータ画像に対応した音声で応答すると共に問いかけに関する検索を開始する検索手段と、検索結果が得られると、修正指示がある場合にはユーザのスケジュールに修正を行う修正手段と、を備えたことにより、対話検索に引き続きスケジューラに追記可能な情報処理方法、情報処理装置、及びプログラムの提供を実現できる。

【 0 0 3 0 】

尚、上述した実施の形態は、本発明の好適な実施の形態の一例を示すものであり、本発明はそれに限定されることなく、その要旨を逸脱しない範囲内において、種々変形実施が可能である。例えば、本実施形態ではユーザから音声による検索の内容が外出を伴う場合で説明したが、本発明はこれに限定されるものではなく、ユーザへの来客を伴う場合であってもスケジュールに追記するように構成してもよい。さらに、ユーザから検索結果に対して、修正指示がある場合にもスケジュールに追記するように構成してもよい。

【 符号の説明 】

【 0 0 3 1 】

- 1 0 0 パーソナルコンピュータ (P C 、 情報処理装置)
- 1 0 0 a モニタ
- 1 0 1 マイクロフォン
- 1 0 2 、 1 0 4 増幅回路
- 1 0 3 スピーカ
- 1 0 5 表示装置
- 1 0 6 キーボード
- 1 0 7 マウス
- 1 0 8 光学読取装置
- 1 0 9 制御手段
- 1 0 9 a 入力制御手段
- 1 0 9 b 音声認識手段
- 1 0 9 c 音声解析手段
- 1 0 9 d 検索手段
- 1 0 9 e 音声合成手段
- 1 0 9 f 修正手段
- 1 1 0 H D D
- 1 1 1 ネットワーク接続部
- 1 1 2 I / O
- 1 1 3 バスライン
- 2 0 0 ユーザ
- 2 0 2 音声信号解釈部
- 2 0 3 クライアント型音声認識部
- 2 0 4 クライアントアプリケーション部
- 2 0 9 テキスト読上部
- 2 1 0 クライアント型音声合成部
- 3 0 0 、 3 0 1 、 3 0 2 、 3 0 3 、 3 0 4 、 3 0 5 、 4 0 1 吹き出し
- 4 0 0 アバター
- 5 0 0 ウィンドウ

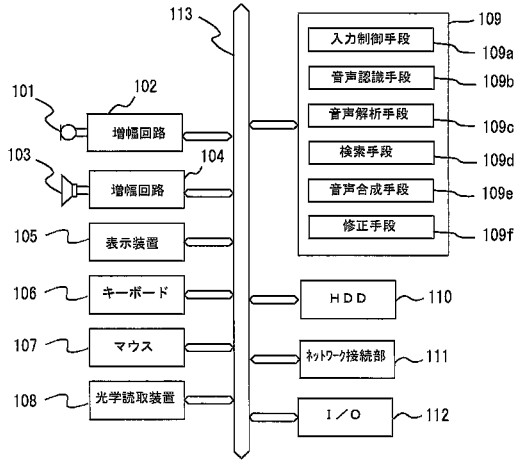
10

20

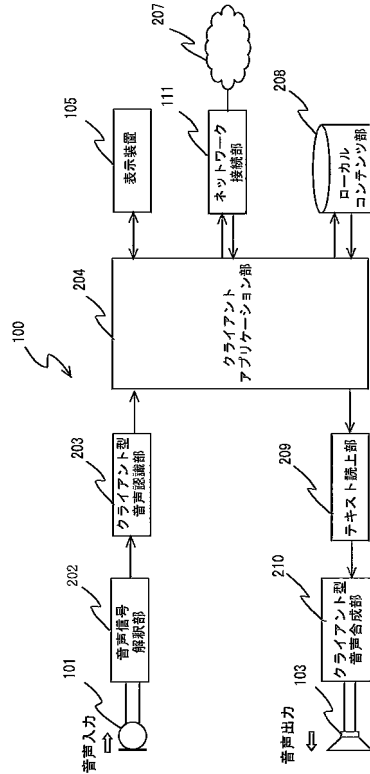
30

40

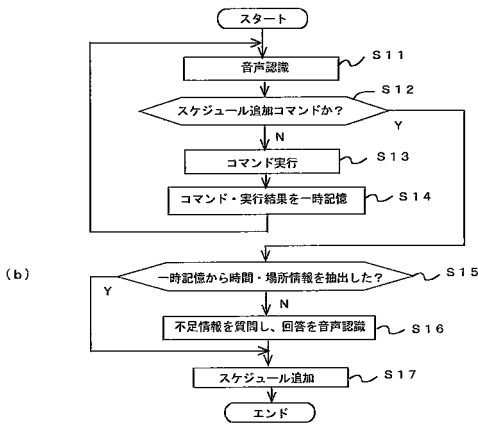
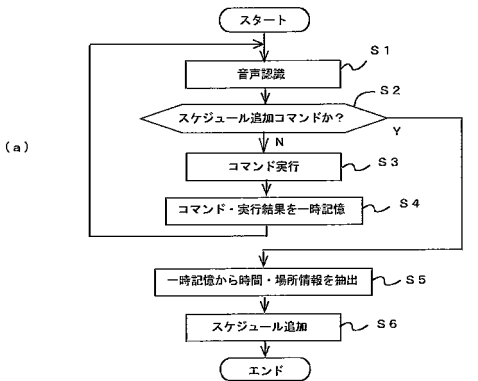
【図 1】



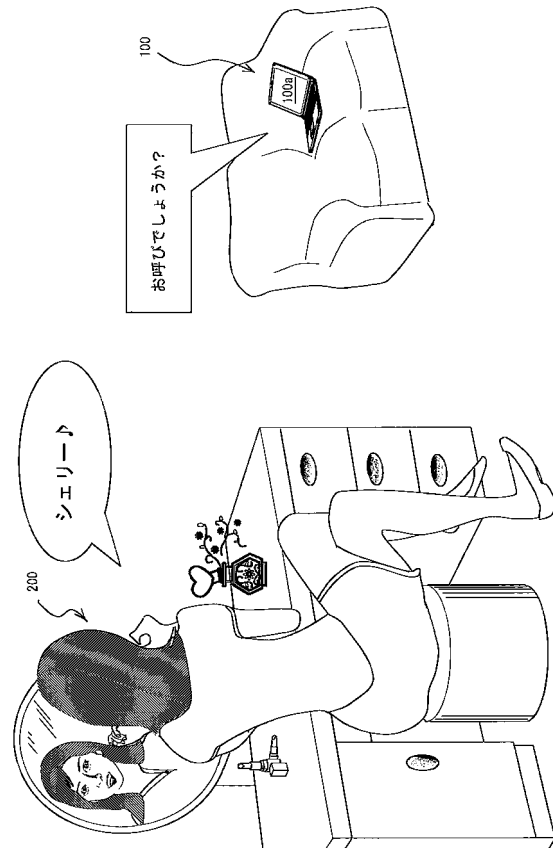
【図 2】



【図 3】



【図 4】



【 図 5 】

