

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6915169号

(P6915169)

(45) 発行日 令和3年8月4日(2021.8.4)

(24) 登録日 令和3年7月16日(2021.7.16)

(51) Int.Cl. F I  
**G 0 6 F 3 / 0 1 (2006.01)** G 0 6 F 3 / 0 1 5 1 0

請求項の数 19 (全 35 頁)

|                    |                               |           |                      |
|--------------------|-------------------------------|-----------|----------------------|
| (21) 出願番号          | 特願2020-546363 (P2020-546363)  | (73) 特許権者 | 502208397            |
| (86) (22) 出願日      | 令和1年6月14日(2019.6.14)          |           | グーグル エルエルシー          |
| (65) 公表番号          | 特表2021-514091 (P2021-514091A) |           | G o o g l e L L C    |
| (43) 公表日           | 令和3年6月3日(2021.6.3)            |           | アメリカ合衆国 カリフォルニア州 94  |
| (86) 国際出願番号        | PCT/US2019/037171             |           | 043 マウンテン ビュー アンフィシ  |
| (87) 国際公開番号        | W02020/005575                 |           | アター パークウェイ 1600      |
| (87) 国際公開日         | 令和2年1月2日(2020.1.2)            |           | 1600 Amphitheatre P  |
| 審査請求日              | 令和2年10月29日(2020.10.29)        |           | arkway 94043 Mounta  |
| (31) 優先権主張番号       | 16/017,394                    |           | in View, CA U. S. A. |
| (32) 優先日           | 平成30年6月25日(2018.6.25)         | (74) 代理人  | 100108453            |
| (33) 優先権主張国・地域又は機関 | 米国 (US)                       |           | 弁理士 村山 靖彦            |
| 早期審査対象出願           |                               | (74) 代理人  | 100110364            |
|                    |                               |           | 弁理士 実広 信哉            |
|                    |                               | (74) 代理人  | 100133400            |
|                    |                               |           | 弁理士 阿部 達彦            |

最終頁に続く

(54) 【発明の名称】 スマートデバイスの識別および制御

(57) 【特許請求の範囲】

【請求項 1】

1つまたは複数のデータ処理装置によって実施される方法であって、  
ユーザのモバイルデバイスのカメラによって取り込まれた画像のための画像データを受信するステップと、

前記画像データに基づいて、前記画像が、1つまたは複数の追加のスマートデバイスを制御するように構成されるスマートデバイスまたは前記スマートデバイスのための物理的コントロールのうちの少なくとも1つを示すと決定するステップと、

前記画像が、1つまたは複数の追加のスマートデバイスを制御するように構成されるスマートデバイスまたは前記スマートデバイスのための物理的コントロールのうちの少なくとも1つを示すとの決定にตอบสนองして、前記1つまたは複数の追加のスマートデバイスの各々を制御するための1つまたは複数のユーザインターフェースコントロールを識別するステップと、

前記モバイルデバイスのディスプレイにおいて、前記1つまたは複数の追加のスマートデバイスを制御するための前記1つまたは複数のユーザインターフェースコントロールを生成および提示するステップと、

前記モバイルデバイスの前記ディスプレイにおいて、前記1つまたは複数のユーザインターフェースコントロールのうちの所与のユーザインターフェースコントロールとのユーザ対話を検出するステップと、

前記検出されたユーザ対話に基づいて、前記所与のユーザインターフェースコントロー

10

20

ルによって制御される所与の追加のスマートデバイスを制御するステップと、を含む方法。

【請求項2】

前記スマートデバイスが、仮想アシスタント制御デバイスを備え、

前記所与の追加のスマートデバイスを制御するステップが、前記所与の追加のスマートデバイスに動作可能に接続された前記仮想アシスタント制御デバイスに、データを送信するステップであって、前記データが、前記検出されたユーザ対話に対応する動作を、前記動作を実施するために前記所与の追加のスマートデバイスに送ることを、前記仮想アシスタント制御デバイスに行わせる、ステップを含む、請求項1に記載の方法。

【請求項3】

前記画像が前記モバイルデバイスにおいて提示されるときの前記モバイルデバイスが、前記仮想アシスタント制御デバイスを制御することが可能であると決定するステップをさらに含む、請求項2に記載の方法。

【請求項4】

前記画像が前記モバイルデバイスにおいて提示されるときの前記モバイルデバイスが、前記仮想アシスタント制御デバイスを制御することが可能であると決定するステップが、

前記画像が前記モバイルデバイスにおいて提示されるときの前記モバイルデバイスが、前記ユーザのアカウントを用いて登録された特定の仮想アシスタント制御デバイスのしきい値距離内であると決定することによって、前記仮想アシスタント制御デバイスが、前記ユーザの前記アカウントを用いて登録された前記特定の仮想アシスタント制御デバイスであると決定するステップを含む、請求項3に記載の方法。

【請求項5】

前記画像が前記モバイルデバイスにおいて提示されるときの前記モバイルデバイスが、前記仮想アシスタント制御デバイスを制御することが可能であるとの決定に応答して、前記スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを提示することを決定するステップをさらに含む、請求項3に記載の方法。

【請求項6】

前記画像データに基づいて、前記画像が、1つまたは複数の追加のスマートデバイスを制御するように構成される前記スマートデバイスまたは前記スマートデバイスのための前記物理的コントロールのうちの少なくとも1つを示すと決定するステップが、

前記画像が前記ユーザの仮想アシスタント制御デバイスの画像を示すと決定するステップを含む、請求項1に記載の方法。

【請求項7】

前記仮想アシスタント制御デバイスが、

前記ユーザから音声コマンドを受信する仮想アシスタントスマートスピーカースピーカーデバイスを備え、

前記仮想アシスタントスマートスピーカースピーカーデバイスのスピーカーを使用して、前記ユーザに情報を提供する、請求項6に記載の方法。

【請求項8】

前記スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを識別するステップが、

前記仮想アシスタントスマートスピーカースピーカーデバイスが現在実施中であるタスクまたはサービスを決定するステップと、

前記タスクまたはサービスを制御するための1つまたは複数のユーザインターフェースコントロールを選択するステップとを含む、請求項7に記載の方法。

【請求項9】

前記モバイルデバイスにおいて、前記1つまたは複数の追加のスマートデバイスを制御するための前記1つまたは複数のユーザインターフェースコントロールを生成および提示するステップが、

前記モバイルデバイスのカメラのビューファインダーの視野上に、拡張現実において前記1つまたは複数のユーザインターフェースコントロールを重ね合わせるステップを含む、請求項1に記載の方法。

【請求項10】

前記画像が、1つまたは複数の追加のスマートデバイスを制御するように構成されるスマートデバイスまたは前記スマートデバイスのための物理的コントロールのうちの少なくとも1つを示すとの決定に応答して、前記1つまたは複数の追加のスマートデバイスの各々を制御するための1つまたは複数のユーザインターフェースコントロールを識別するステップが、

前記各追加のスマートデバイスのための登録ファイルを取得するステップであって、前記登録ファイルが、追加のスマートデバイスのタイプと、前記追加のスマートデバイスを制御するための利用可能なユーザインターフェースコントロールとを指定するデータを含む、ステップを含む、請求項1に記載の方法。

10

【請求項11】

前記画像データに基づいて、前記受信された画像データが、1つまたは複数の追加のスマートデバイスを制御するように構成されるスマートデバイスまたは前記スマートデバイスのための物理的コントロールのうちの少なくとも1つを示すと決定するステップが、

前記スマートデバイスが前記ユーザのデバイスであることを確認するように、前記ユーザにプロンプトを出すステップと、

前記ユーザから確認を受信するステップとを含む、請求項1に記載の方法。

20

【請求項12】

システムであって、

1つまたは複数のデータ処理装置と、

前記1つまたは複数のデータ処理装置とデータ通信しているメモリ記憶装置とを備え、前記メモリ記憶装置が、前記1つまたは複数のデータ処理装置によって実行可能な命令であって、その実行時に、前記1つまたは複数のデータ処理装置に動作を実行させる、命令を記憶し、前記動作が、

ユーザのモバイルデバイスのカメラによって取り込まれた画像のための画像データを受信すること、

30

前記画像データに基づいて、前記画像が、1つまたは複数の追加のスマートデバイスを制御するように構成されるスマートデバイスまたは前記スマートデバイスのための物理的コントロールのうちの少なくとも1つを示すと決定すること、

前記画像が、1つまたは複数の追加のスマートデバイスを制御するように構成されるスマートデバイスまたは前記スマートデバイスのための物理的コントロールのうちの少なくとも1つを示すとの決定に応答して、前記1つまたは複数の追加のスマートデバイスの各々を制御するための1つまたは複数のユーザインターフェースコントロールを識別すること、

前記モバイルデバイスのディスプレイにおいて、前記1つまたは複数の追加のスマートデバイスを制御するための前記1つまたは複数のユーザインターフェースコントロールを生成および提示すること、

40

前記モバイルデバイスの前記ディスプレイにおいて、前記1つまたは複数のユーザインターフェースコントロールの所与のユーザインターフェースコントロールとのユーザ対話を検出すること、ならびに

前記検出されたユーザ対話に基づいて、前記所与のユーザインターフェースコントロールによって制御される所与の追加のスマートデバイスを制御することを含む、システム。

【請求項13】

前記動作が、

前記画像が前記モバイルデバイスにおいて提示されるときの前記モバイルデバイスが、

50

仮想アシスタント制御デバイスを制御することが可能であると決定すること  
をさらに含む、請求項12に記載のシステム。

【請求項14】

前記画像が前記モバイルデバイスにおいて提示されるときの前記モバイルデバイスが、  
前記仮想アシスタント制御デバイスを制御することが可能であると決定することが、

前記画像が前記モバイルデバイスにおいて提示されるときの前記モバイルデバイスが、  
前記ユーザのアカウントを用いて登録された特定の仮想アシスタント制御デバイスのしき  
い値距離内であると決定することによって、前記仮想アシスタント制御デバイスが、前記  
ユーザの前記アカウントを用いて登録された前記特定の仮想アシスタント制御デバイスで  
あると決定することを含む、請求項13に記載のシステム。

10

【請求項15】

前記動作が、

前記画像が前記モバイルデバイスにおいて提示されるときの前記モバイルデバイスが、  
前記仮想アシスタント制御デバイスを制御することが可能であるとの決定に応答して、前  
記スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロ  
ールを提示することを決定することをさらに含む、請求項13に記載のシステム。

【請求項16】

前記画像データに基づいて、前記画像が、1つまたは複数の追加のスマートデバイスを  
制御するように構成される前記スマートデバイスまたは前記スマートデバイスのための前  
記物理的コントロールのうちの少なくとも1つを示すと決定することが、

20

前記画像が前記ユーザの仮想アシスタント制御デバイスの画像を示すと決定すること  
を含む、請求項12に記載のシステム。

【請求項17】

前記モバイルデバイスにおいて、前記スマートデバイスを制御するための前記1つまた  
は複数のユーザインターフェースコントロールを生成および提示することが、

前記モバイルデバイスのカメラのビューファインダーの視野上に、拡張現実において前  
記1つまたは複数のユーザインターフェースコントロールを重ね合わせることを含む、請  
求項12に記載のシステム。

【請求項18】

コンピュータプログラムが符号化された非一時的コンピュータ記憶媒体であって、前記  
コンピュータプログラムが、データ処理装置によって実行されると、前記データ処理装置  
に動作を実施させる命令を備え、前記動作が、

30

ユーザのモバイルデバイスのカメラによって取り込まれた画像のための画像データを受  
信すること、

前記画像データに基づいて、前記画像が、1つまたは複数の追加のスマートデバイスを  
制御するように構成されるスマートデバイスまたは前記スマートデバイスのための物理的  
コントロールのうちの少なくとも1つを示すと決定すること、

前記画像が、1つまたは複数の追加のスマートデバイスを制御するように構成されるス  
martデバイスまたは前記スマートデバイスのための物理的コントロールのうちの少なく  
とも1つを示すとの決定に応答して、前記1つまたは複数の追加のスマートデバイスの各々  
を制御するための1つまたは複数のユーザインターフェースコントロールを識別すること

40

、  
前記モバイルデバイスのディスプレイにおいて、前記1つまたは複数の追加のスマート  
デバイスを制御するための前記1つまたは複数のユーザインターフェースコントロールを  
生成および提示すること、

前記モバイルデバイスの前記ディスプレイにおいて、前記1つまたは複数のユーザイン  
ターフェースコントロールの所与のユーザインターフェースコントロールとのユーザ対話  
を検出すること、ならびに

前記検出されたユーザ対話に基づいて、前記所与のユーザインターフェースコントロ  
ールによって制御される所与の追加のスマートデバイスを制御すること

50

を含む、非一時的コンピュータ記憶媒体。

【請求項19】

前記スマートデバイスが、仮想アシスタント制御デバイスを備え、  
前記所与の追加のスマートデバイスを制御することが、前記所与の追加のスマートデバイスに動作可能に接続された前記仮想アシスタント制御デバイスに、データを送信することとあって、前記データが、前記検出されたユーザ対話に対応する動作を、前記動作を実施するために前記所与の追加のスマートデバイスに送ることを、前記仮想アシスタント制御デバイスに行わせる、送信することを含む、請求項12に記載のシステム。

【発明の詳細な説明】

【背景技術】

10

【0001】

スマートフォンなど、多数のモバイルデバイスには、写真およびデジタル画像を取り込むためのカメラが装備されている。モバイルデバイスはまた、多種多様な情報へのアクセスも提供する。この情報は、モバイルデバイス上で稼働しているウェブブラウザまたはネイティブアプリケーション内で見られることがある。ユーザはまた、画像を使用して、画像内に示されたオブジェクトまたはロケーションに関連する追加の情報を取得することもできる。

【0002】

仮想アシスタントは、ユーザのためにタスクまたはサービスを実施するアプリケーションである。仮想アシスタントは、音声アシスタントデバイス(たとえば、スマートスピーカー、または他の音声により制御されるデバイス)内に統合され得る。いくつかの仮想アシスタントはまた、ホームオートメーションデバイスなど、他のスマートデバイスを制御するスマートデバイスの役割もする。たとえば、ユーザは、ライトを制御するように構成される音声アシスタントデバイスに、音声コマンドを与えることによって、ライトの動作を制御することができる。

20

【発明の概要】

【課題を解決するための手段】

【0003】

本明細書は、スマートデバイスを識別し、識別されたスマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを提示することに関する技術について説明する。

30

【0004】

概して、本明細書で説明する主題の1つの発明的態様は、方法において具現化され得、方法は、ユーザのモバイルデバイスのカメラによって取り込まれた画像のための画像データを受信するステップと、画像データに基づいて、画像がスマートデバイスまたはスマートデバイスのための物理的コントロールのうちの少なくとも1つを示すと決定するステップと、画像がスマートデバイスまたはスマートデバイスのための物理的コントロールのうちの少なくとも1つを示すとの決定に応答して、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを識別するステップと、モバイルデバイスのディスプレイにおいて、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを生成および提示するステップと、モバイルデバイスのディスプレイにおいて、1つまたは複数のユーザインターフェースコントロールのうちの少なくとも1つとのユーザ対話を検出するステップと、検出されたユーザ対話に基づいて、スマートデバイスを制御するステップとを含む。この態様の他の実装形態は、対応する装置、方法、システム、およびコンピュータ記憶デバイス上に符号化された、方法のアクションを実施するように構成されたコンピュータプログラムを含む。

40

【0005】

これらの実装形態および他の実装形態は、場合によっては、以下の特徴のうちの1つまたは複数の各々を含み得る。いくつかの態様では、スマートデバイスを制御するステップが、スマートデバイスに動作可能に接続された仮想アシスタント制御デバイスに、データを

50

送信するステップであって、データが、検出されたユーザ対話に対応する動作を、動作を実施するためにスマートデバイスに送ることを、仮想アシスタント制御デバイスに行わせる、ステップを含む。

#### 【0006】

いくつかの態様では、画像データに基づいて、画像がスマートデバイスまたはスマートデバイスのための物理的コントロールのうちの少なくとも1つを示すと決定するステップが、画像が、他のスマートデバイスを制御するように構成される仮想アシスタント制御デバイスを示すと決定するステップを含む。いくつかの態様では、方法は、画像がモバイルデバイスにおいて提示されるときモバイルデバイスが、仮想アシスタント制御デバイスを制御することが可能であると決定するステップをさらに含み得る。いくつかの態様では、画像がモバイルデバイスにおいて提示されるときモバイルデバイスが、仮想アシスタント制御デバイスを制御することが可能であると決定するステップが、画像がモバイルデバイスにおいて提示されるときモバイルデバイスが、ユーザのアカウントを用いて登録された特定の仮想アシスタント制御デバイスのしきい値距離内であると決定することによって、仮想アシスタント制御デバイスが、ユーザのアカウントを用いて登録された特定の仮想アシスタント制御デバイスであると決定するステップを含む。いくつかの態様では、方法は、画像がモバイルデバイスにおいて提示されるときモバイルデバイスが、仮想アシスタント制御デバイスを制御することが可能であるとの決定に応答して、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを提示することを決定するステップをさらに含み得る。

#### 【0007】

いくつかの態様では、画像データに基づいて、画像がスマートデバイスまたはスマートデバイスのための物理的コントロールのうちの少なくとも1つを示すと決定するステップが、画像がユーザの仮想アシスタント制御デバイスの画像を示すと決定するステップを含む。いくつかの態様では、仮想アシスタント制御デバイスが、ユーザから音声コマンドを受信する仮想アシスタントスマートスピーカーデバイスを含み、仮想アシスタントスマートスピーカーデバイスのスピーカーを使用して、ユーザに情報を提供する。いくつかの態様では、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを識別するステップが、仮想アシスタントスマートスピーカーデバイスが現在実施中であるタスクまたはサービスを決定するステップと、タスクまたはサービスを制御するための1つまたは複数のユーザインターフェースコントロールを選択するステップとを含む。

#### 【0008】

いくつかの態様では、モバイルデバイスにおいて、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを生成および提示するステップが、モバイルデバイスのカメラのビューファインダーの視野上に、拡張現実において1つまたは複数のユーザインターフェースコントロールを重ね合わせるステップを含む。

#### 【0009】

いくつかの態様では、画像がスマートデバイスまたはスマートデバイスのための物理的コントロールのうちの少なくとも1つを示すとの決定に応答して、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを識別するステップが、スマートデバイスのための登録ファイルを取得するステップであって、登録ファイルが、スマートデバイスのタイプと、スマートデバイスを制御するための利用可能なユーザインターフェースコントロールとを指定するデータを含む、ステップを含む。

#### 【0010】

本明細書で説明する主題は、以下の利点のうちの1つまたは複数を実現するように、特定の実施形態において実装され得る。それを通してスマートデバイスが画像内で識別され、仮想ユーザインターフェースコントロールが識別されたスマートデバイスのために提示される、プラットフォームを容易にすることによって、プラットフォームは、オブジェクト、それらの能力、および/またはオブジェクトが現在動作中であるコンテキストに合わ

せてカスタマイズまたは適合される、仮想ユーザインターフェースコントロールの提示を可能にする。

#### 【0011】

従来のユーザインターフェースでは、ユーザは、正しいデータ/機能を発見するため、または音声アシスタントデバイスに複数の音声コマンドを与えるために、何回もビューをスクロールして切り替える必要があり得る。本システムの仮想アシスタントアプリケーションは、スマートデバイスごとに別個のアプリケーションを発見する必要なしに、ユーザが(たとえば、拡張現実技法を使用する)モバイルデバイス上でスマートデバイスを見ている間、スマートデバイスのための仮想ユーザインターフェースコントロールを提示することができる。たとえば、ユーザは、ユーザの自宅、オフィス、または他のロケーションにおいて、仮想アシスタント制御デバイス(たとえば、スマートスピーカーなどの音声アシスタントデバイス)が音楽を再生中である間、モバイルデバイスのカメラを仮想アシスタント制御デバイスに向けることができる。仮想アシスタントアプリケーションは(たとえば、単独で、またはリモートシステムの助けをかりて)、カメラから受信された画像データに基づいて、仮想アシスタント制御デバイスを検出し、ユーザが現在再生されているオーディオを制御することを可能にするユーザインターフェースコントロールを提示することができる。画像データ内で検出されたスマートデバイスと、スマートデバイスの現在のモードまたはアクションとに基づいて、ユーザインターフェースコントロールを選択することによって、仮想アシスタントアプリケーションは、スマートデバイスのより効率的および効果的な制御を提供する、カスタマイズされたユーザインターフェースを提供することができる。たとえば、音声制御を使用して、音楽の音量または再生の調整を試行する際には、音量または再生の適切な変化を得るために、複数の音声コマンド、およびスマートスピーカーによる調整が必要となることがある。モバイルデバイスによって提示されたユーザインターフェースコントロールは、ユーザがより迅速にかつより直観的に同じ変更を行うことを可能にすることができる。

#### 【0012】

本明細書で説明するシステムおよび技法は、モバイルデバイスのカメラからの画像データ(たとえば、単一のフレーム画像、連続したビデオ、画像のストリームなど)から、スマートデバイスを認識することができる。スマートデバイスが識別されると、モバイルデバイスは、結果にインデックス付けすることができる。これによって、ユーザが複数の要求を行う必要がなくなり、それによって、次に、コンピューティングシステムがスマートデバイスを識別するための要求を処理し、どのユーザインターフェースコントロールが特定のスマートデバイスアプリケーションのために提示されるべきであるか(たとえば、音楽再生のためのオーディオコントロール、ライトのためのライトスイッチコントロールなど)を決定する回数が低減される。これが多数のユーザにわたって集約されることで、コンピュータへの処理需要を低減することによって、ユーザインターフェースコントロールを選択する仮想アシスタント管理システムのコンピュータの機能を著しく改善することができる。仮想アシスタント管理システムは、ネットワークを介してアクセスされ得るので、要求の数を低減することによって、消費される帯域幅の量も低減され、他のネットワークトラフィックのための帯域幅を可能にすること、および/またはネットワークの速度を増すことができる。

#### 【0013】

スマートデバイスに基づいて、適切なユーザインターフェースコントロールを選択することによって、ユーザは、ユーザがモバイルデバイスのカメラを同じスマートデバイスに向けるたびに、コントロールを選択しなければならない負担がなくなる。これによって、ユーザによるより迅速な要求、およびユーザのためのより高い柔軟性が可能になる。たとえば、仮想アシスタントアプリケーションを使用することによって、ユーザは、カメラがスマートデバイスに向いているようにモバイルデバイスを保持し、モバイルデバイスまたはスマートデバイスとさらに対話する必要なしに、スマートデバイスのいくつかの特徴を制御するためのユーザインターフェースコントロールが提示され得る。したがって、本明

細書で説明するシステムおよび技法は、スマートデバイスを制御するためのガイド付きヒューマンマシン対話プロセスを提供する。

【0014】

上記の主題の様々な特徴および利点について、以下で図に関して説明する。追加の特徴および利点は、本明細書で説明する主題、および特許請求の範囲から明らかである。

【図面の簡単な説明】

【0015】

【図1A】仮想アシスタントアプリケーションがスマートデバイスを識別し、スマートデバイスを制御するためのユーザインターフェースコントロールを提示する、例示的な環境のブロック図である。

10

【図1B】仮想アシスタントアプリケーションが、識別されたスマートデバイスのためのユーザインターフェースコントロールを提示する、例示的なプロセスの例示的なシステムフロー図である。

【図2】受信された画像データに基づいて、識別されたスマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを提示する、モバイルデバイスの例示的なスクリーンショットのシーケンスの図である。

【図3】受信された画像データに基づいて、識別されたスマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを提示する、モバイルデバイスの例示的なスクリーンショットの別のシーケンスの図である。

【図4】仮想アシスタントアプリケーションを使用して、スマートデバイスを識別し、識別されたスマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを提示するための、例示的なプロセスのフロー図である。

20

【図5】本開示で説明する方法、システム、およびプロセスを実装するために使用され得る、例示的なコンピュータシステムのブロック図である。

【発明を実施するための形態】

【0016】

様々な図面における同様の参照番号および名称は、同様の要素を示す。

【0017】

仮想アシスタントアプリケーションを使用して、画像データ(たとえば、単一のフレーム画像、連続したビデオ、画像のストリームなど)に基づいて、スマートデバイス(または、スマートデバイスのための物理的コントロール)を識別し、識別されたスマートデバイスごとに、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを提示するためのシステム、方法、およびコンピュータプログラム製品について説明する。たとえば、スマートデバイスまたはスマートデバイスのための物理的コントロールは、モバイルデバイスのカメラのビューファインダーを表す画像データ内で識別(たとえば、オブジェクト認識技法を使用して認識)され得る。それに応答して、スマートデバイスを制御するためのユーザインターフェースコントロールは、ユーザがスマートデバイスを制御することができるように、たとえば、拡張現実技法を使用して、ビューファインダー内で提示され得る。スマートデバイスは、1つまたは複数のネットワークを介して他のデバイスに接続される電子デバイスである。スマートデバイスは、自律的に、かつ/または他のスマートデバイスとともに対話的に動作することができる。例示的なスマートデバイスには、いくつかの例を挙げれば、仮想アシスタント制御デバイス(たとえば、スマートスピーカーなどの音声制御デバイス)、ホームオートメーションデバイス(たとえば、スマートライト、スマートアプライアンス、サーモスタットなど)、スマートテレビ、およびスマートラジオが含まれる。仮想アシスタント制御デバイスは、ユーザのためにタスクおよび/またはサービスを実施し、他のスマートデバイスを制御する、統合された仮想アシスタントを含む電子デバイスである。例示的な仮想アシスタント制御デバイスは、音声コマンドに応答して、タスクを実施し、他のスマートデバイスを制御する、スマートスピーカーである。

30

40

【0018】

50



たとえば、ユーザは、モバイルデバイスのカメラをスマートスピーカーに向けて、スマートスピーカー、またはそのためにスマートスピーカーが制御するように構成されている他のスマートデバイス(たとえば、ユーザのアカウントを用いて登録された他のスマートデバイス)を制御するための、ユーザインターフェースコントロールを見ることができる。別の例では、ユーザは、モバイルデバイスのカメラをスマートライトまたはライトスイッチに向けて、スマートライトを調整するためのユーザインターフェースコントロールを見ることができる。

#### 【0019】

いくつかの態様では、モバイルデバイスは、仮想アシスタント制御デバイス(たとえば、スマートスピーカー)を通して、スマートデバイスに接続し、その場合、各々が同じネットワークに接続されている。いくつかの態様では、モバイルデバイスは、仮想アシスタント制御デバイスに、または仮想アシスタント制御デバイスを通して接続することなしに、スマートデバイスに接続することができ、たとえば、その場合、その特定のスマートデバイスは、モバイルデバイスと同じネットワーク(たとえば、ワイヤレスネットワーク)を介して、データを通信するように構成される。

#### 【0020】

いくつかの態様によれば、ユーザは、ユーザのアカウントを用いて、仮想アシスタント制御デバイスを登録することができる。ユーザはまた、仮想アシスタント制御デバイスが各登録されたスマートデバイスを制御することができるように、アカウントを用いて、他のスマートデバイスを登録することもできる。別の例では、ユーザは、たとえば、仮想アシスタント制御デバイスにおいて、1つまたは複数のスマートデバイスを制御するように、仮想アシスタント制御デバイスを構成することができる。次いで、ユーザは、それ自体のスピーカーにおいて音楽を再生するように、仮想アシスタント制御デバイスに要求するか、あるいは、ライトが、ユーザのアカウントを用いて登録されているか、またはさもなければ仮想アシスタント制御デバイスによって制御される、スマートデバイスである(または、スマートデバイスに接続される)場合、ある室内でライトをオフにするように、仮想アシスタント制御デバイスに要求することができる。仮想アシスタント制御デバイスは、ワイヤレスネットワークを介して、スマートデバイスに制御データを送ることによって、またはスマートデバイスに制御データを中継する仮想アシスタント管理システムに、制御データを送ることによって、スマートデバイスを制御することができる。

#### 【0021】

本明細書で説明する仮想アシスタントアプリケーションは、画像がモバイルデバイスにおいて提示されるときの(または、画像が取り込まれたときの)モバイルデバイスのロケーションが、ユーザのアカウントを用いて登録された仮想アシスタント制御デバイスのしきい値距離内であると決定することによって、画像データ内で検出されたスマートデバイス(仮想アシスタント制御デバイス、またはスマートライトスイッチなど)が、モバイルデバイスのユーザに関連付けられるかを決定することができる。たとえば、システムは、モバイルデバイスの地理的ロケーション、およびユーザのアカウントを用いて登録されている仮想アシスタント制御デバイスのロケーションを使用して、画像データによって表された特定のスマートデバイスがユーザのスマートデバイスであるかを決定することができる。次いで、仮想アシスタントアプリケーションは、生成されたインターフェースコントロールを通して、仮想アシスタント制御デバイスと特定のスマートデバイスとの間の接続を使用して、スマートデバイスを制御することができる。スマートデバイスのためのアプリケーションは、特定のスマートデバイスを構成し、WiFiネットワークなどのネットワークに接続させるために使用され得る。仮想アシスタント制御デバイスのためのアプリケーションは、同じWiFiネットワーク上でスマートデバイスを接続および制御し、特定のスマートデバイスをさらに構成する(たとえば、ニックネームを追加する、部屋に割り当てるなどの)ために使用され得る。

#### 【0022】

いくつかの態様によれば、仮想アシスタントアプリケーションは、仮想アシスタント制

10

20

30

40

50

御デバイスなしに、スマートデバイスに接続することができる。たとえば、仮想アシスタントアプリケーションは、モバイルデバイスのユーザが、スマートデバイスが登録されているホームネットワークへのアクセス(たとえば、WiFiパスワードを通してホームネットワークにアクセスするための許可など)を有するかを決定することができる。本明細書で説明する仮想アシスタントアプリケーションシステムは、画像がモバイルデバイスにおいて提示されるとき(または、画像が取り込まれたときの)モバイルデバイスのロケーションが、スマートデバイスのロケーション(同じネットワークなど)内であり、ユーザのアカウントを用いて登録されていると決定することによって、画像データ内で認識されたスマートデバイス(スマートライトスイッチ、またはスマートサーモスタットなど)が、ユーザに登録されているかを決定することができる。たとえば、システムは、モバイルデバイスの地理的ロケーションを使用し、仮想アシスタント管理システムを通して登録ファイルにアクセスして、モバイルデバイスが特定のスマートデバイスにアクセスすることを可能にされるかを決定することができる。登録ファイルは、限定はしないが、識別および制御情報を含む、特定のスマートデバイスについてのデータを含む。仮想アシスタントアプリケーションは、登録ファイル内のデータを使用して、特定のスマートデバイスと通信し、特定のスマートデバイスを制御することができる。

#### 【0023】

本明細書で説明するユーザインターフェースコントロールは、モバイルデバイスのユーザインターフェースによって、たとえば、モバイルデバイス上で実行している仮想アシスタントアプリケーションによって、生成および提示され得る。本システムは、拡張現実(AR)モジュールなどを使用して、モバイルデバイスのユーザインターフェースによって提示されたライブ画像(たとえば、デジタルビューファインダー)上に、または以前に取り込まれた画像上に、ユーザインターフェースコントロールをオーバーレイすることができる。たとえば、ユーザは、モバイルデバイスのカメラのビューファインダー内でスマートライトスイッチのライブビューを有し、ライブ画像上に重ね合わせられたユーザインターフェースコントロールを仮想的に見ることができる。

#### 【0024】

動作時に、仮想アシスタントアプリケーションは、画像データと、そこにおいて仮想アシスタントアプリケーションが実行中であるモバイルデバイスのロケーションを指定するロケーションデータとを受信する。仮想アシスタントアプリケーションの(または、リモートサーバの)オブジェクト認識器構成要素は、画像データを解析し、識別されたオブジェクトデータを生成することができる。識別されたオブジェクトデータは、オブジェクトが占有する画像の領域を指定する。いくつかの態様では、オブジェクト認識器構成要素は、オブジェクトがスマートデバイスであるかを決定し、識別されたオブジェクトデータにその情報を埋め込むことができる。仮想アシスタントアプリケーション(または、リモートサーバ)のスマートデバイスコントロールセクタは、オブジェクト認識器構成要素から、識別されたオブジェクトデータを受信し、画像データが、スマートデバイスなど、指定されたオブジェクトを示すかを決定する。スマートデバイスが、識別されたオブジェクトデータから識別される場合、スマートデバイスコントロールセクタは、モバイルデバイスからロケーションデータを受信する。次いで、スマートデバイスコントロールセクタは、画像が提示される(または、画像のための画像データが取り込まれた)ときのモバイルデバイスのロケーションが、ユーザのアカウントを用いて登録された仮想アシスタント制御デバイスのロケーションのしきい値距離(たとえば、50メートル、100メートル、または別の適切な距離)内であると決定することができる。いくつかの態様では、システムは、スマートデバイスがユーザのアカウントに関連付けられるか、および、仮想アシスタント制御デバイスがスマートデバイスを制御することが可能であるかを決定することができる。

#### 【0025】

いくつかの態様では、スマートデバイスが、識別されたオブジェクトデータから識別される場合、スマートデバイスコントロールセクタは、モバイルデバイスからロケーショ

10

20

30

40

50

ンデータを受信し、次いで、たとえば、音声アシスタント制御デバイスがネットワーク上で発見されない場合、登録ファイルを処理して、スマートデバイスがユーザのアカウントに登録されているかを決定する。登録ファイルは、(既知である場合は)ユーザデバイス上に記憶され得るか、またはスマートデバイス登録エンジンなど、仮想アシスタント管理システム内でバックエンドサーバを通してアクセスされ得る。

#### 【0026】

モバイルデバイスのロケーションが、ユーザのアカウントを用いて登録された仮想アシスタント制御デバイスのロケーションのしきい値距離内であると、スマートデバイスコントロールセレクトアが決定した後、スマートデバイスコントロールセレクトアは、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを識別(たとえば、選択)し、インターフェースコントロール識別データを生成する。インターフェースコントロール識別データは、識別されたスマートデバイスを制御することができる、1つまたは複数のコントロールを指定するデータを含む。たとえば、スマートデバイスがライトスイッチである場合、コントロールは、オンアイコン、オフアイコン、および/または調光器を含み得る。インターフェースコントロール識別データは、インターフェース生成器によって受信される。インターフェースコントロール識別データによって、対象のスマートデバイスに属すると指定された各ユーザインターフェースコントロールについて、インターフェース生成器は、ユーザインターフェースにおいてユーザインターフェースコントロールを提示する提示データを生成する。

#### 【0027】

受信された画像データは、リアルタイムで処理され、カメラが移動する場合でも、カメラのレンズの視野内で対象のオブジェクトを識別し続けることになる、ライブの連続したビデオからのものであり得る。いくつかの態様では、受信された画像データは、単一のフレームショット、または録画されたビデオであり得る。

#### 【0028】

これらの特徴および追加の特徴について、以下でより詳細に説明する。

#### 【0029】

図1Aは、仮想アシスタントアプリケーション116がスマートデバイスを識別し、スマートデバイスを制御するためのユーザインターフェースコントロールを提示する、例示的な環境100Aのブロック図である。図1Bは、仮想アシスタントアプリケーション116が、識別されたスマートデバイスのためのユーザインターフェースコントロールを提示する、例示的なプロセスの例示的なシステムフロー図である。

#### 【0030】

仮想アシスタントアプリケーション116は、モバイルデバイス110上にインストールされ、および/またはモバイルデバイス110によって実行され得る。モバイルデバイス110は、データ通信ネットワーク150を介してデータを送受信することが可能である、電子デバイスである。例示的なモバイルデバイス110には、スマートフォン、タブレットコンピューティングデバイス、ウェアラブルコンピューティングデバイス(たとえば、スマートウォッチ)、およびネットワーク150を介してデータを送受信することができる他のデバイスが含まれる。ネットワーク150には、ローカルエリアネットワーク(LAN)、ワイドエリアネットワーク(WAN)、インターネット、モバイルネットワーク、またはそれらの組合せが含まれ得る。

#### 【0031】

仮想アシスタントアプリケーション116は、特定のプラットフォームまたは特定のデバイスのために開発されたネイティブアプリケーションとして実装され得る。仮想アシスタントアプリケーション116は、モバイルデバイス110のユーザのためにタスクまたはサービスを実施することができる。たとえば、仮想アシスタントアプリケーション116は、ユーザの音声コマンドに応答する(たとえば、要求された情報を提供する)、ユーザのスマートデバイスを制御する、コンテンツ(たとえば、音楽またはビデオ)を再生することなどができる。ユーザによるスマートデバイスの制御をより容易かつより効率的にするために、仮

想アシスタントアプリケーション116は、モバイルデバイス110のディスプレイ上にグラフィカルユーザインターフェースコントロールを提示することができる。

【0032】

仮想アシスタントアプリケーション116は、モバイルデバイス110のカメラ111の視野内のシーンを表す画像データ内で認識されたスマートデバイスのためのユーザインターフェースコントロールを提示することができる。たとえば、仮想アシスタントアプリケーション116は、スマートデバイスを識別し、識別されたスマートデバイスのためのユーザインターフェースコントロールを提示するかを決定することができる。仮想アシスタントアプリケーション116は、識別されたスマートデバイスが、ユーザのアカウントを用いて登録されているか、および/またはモバイルデバイス110が、ユーザのアカウントを用いて登録された仮想アシスタント制御デバイスのしきい値距離内であるかに基づいて、識別されたスマートデバイスのためのユーザインターフェースコントロールを提示するかを決定することができる。仮想アシスタントアプリケーション116はまた、識別されたスマートデバイスのために、どのユーザインターフェースコントロールを使用するか(たとえば、音楽のためのオーディオコントロール、スマートデバイスライトスイッチのためのトグルコントロールなど)を選択し、選択されたコントロールをもつインターフェースを、たとえば、画像データ(たとえば、画像またはカメラ111のビューファインダー内に示されたオブジェクト)上のグラフィックオーバーレイとして提示することもできる。

【0033】

仮想アシスタントアプリケーション116は、たとえば、スマートデバイスを識別するためのユーザ要求を受信することなしに、連続して、モバイルデバイス110のカメラ111から受信された画像データ内で、スマートデバイスを検出することを試行することができる。たとえば、仮想アシスタントアプリケーション116は、(画像データに基づいて)モバイルデバイス110のカメラ111のビューファインダー内でオブジェクトを検出および/または認識し、ユーザがカメラ111をスマートデバイスに向けているという事実を、識別されたスマートデバイスを制御するための要求として解釈することができる。

【0034】

いくつかの実装形態では、仮想アシスタントアプリケーション116は、1つまたは複数の入力を介して、スマートデバイスコントロールを位置特定および選択するためのコマンドを受信することができる。たとえば、仮想アシスタントアプリケーション116は、モバイルデバイス110のマイクロフォン113から音声要求を受信することができる。仮想アシスタントアプリケーション116は、マイクロフォン113から入力されたオーディオを受信し、(オーディオが、話された言葉を含む場合)オーディオをテキストに変換し、仮想アシスタントアプリケーション116にテキストを提供することができる、音声認識器120を含み得る。

【0035】

いくつかの実装形態では、仮想アシスタントアプリケーション116は、ユーザが、本明細書で説明するプロセスを開始するために、モバイルデバイス110上に表示された仮想アシスタントアプリケーション116のインターフェース上のボタンと対話する(たとえば、タッチする)ことによって、スマートデバイスを識別するためのコマンドを受信することができる。いくつかの実装形態では、仮想アシスタントアプリケーション116はまた、たとえば、モバイルデバイス110の物理的またはタッチキーパッドを使用する際にタイプされた、テキスト要求を受信することもできる。

【0036】

仮想アシスタントアプリケーション116は、スマートデバイスが画像内に(たとえば、画像を表すピクセルデータ内に)存在するかを決定し、識別されたスマートデバイス、その能力、および/またはスマートデバイスが現在動作中であるコンテキストに基づいて、ユーザインターフェースコントロールを選択する、スマートデバイスコントロールセクタ130を含む。いくつかの態様では、仮想アシスタントアプリケーション116は、仮想アシスタントアプリケーション116がアクティブである間、オブジェクト認識器125(本明細書で

さらに説明する)から、識別されたオブジェクトデータ126を受信することができる。たとえば、仮想アシスタントアプリケーション116は、仮想アシスタントアプリケーション116が起動されるとき、カメラ111からビューファインダーのためのピクセルデータを取得することを開始することができる。スマートデバイスコントロールセクタ130は、スマートデバイスコントロールセクタ130がユーザインターフェースコントロールを選択するために十分な情報を有するまで、データを監視することができる。たとえば、スマートデバイスコントロールセクタ130は、スマートデバイスコントロールセクタ130が、ユーザインターフェースコントロールを提示するか、およびどのコントロールを選択するかについての判断を行うために、十分なデータを取得するまで、画像のストリーム、および/またはロケーションデータ124などの他のデータについて、識別されたオブジェクトデータ126を監視することができる。

10

#### 【0037】

いくつかの実装形態では、スマートデバイスコントロールセクタ130は、カメラ111のビューファインダーのための画像データ123に基づいて、オブジェクト認識器125によって生成された、識別されたオブジェクトデータ126に基づいて、ユーザインターフェースコントロールを選択することができる。画像データ123は、カメラ111のビューファインダーの現在のシーンを表す、ピクセルデータを含み得る。仮想アシスタントアプリケーション116は、仮想アシスタントアプリケーション116が起動された後、カメラ111から画像データ123を取得することができる。たとえば、仮想アシスタントアプリケーション116は、ピクセルデータセットのストリームを取得することができる。各ピクセルデータセットは、特定の時点のためのビューファインダーのピクセルを表すことができる。各ピクセルデータセット内のピクセルデータは、ビューファインダーの各ピクセルの視覚的特性(たとえば、色、強度、輝度など)を指定するデータを含み得る。

20

#### 【0038】

スマートデバイスコントロールセクタ130は、スマートデバイス(または、スマートデバイスのための物理的コントロール)がビューファインダー内で(たとえば、ピクセルデータセットのうちの1つまたは複数において)検出されるか、および、そうである場合、スマートデバイスの識別、またはスマートデバイスのクラスに基づいて、ユーザインターフェースコントロールを選択することができる。たとえば、仮想アシスタントアプリケーション116は、ピクセルデータ(または、画像)内でスマートデバイス(または、それらの物理的コントロール)を検出および認識(たとえば、識別)することを試行する、オブジェクト認識器125を含み得る。オブジェクト認識器125は、エッジ検出および/または他のオブジェクト認識技法を使用して、スマートスピーカー、アプライアンス、テレビ、物理的コントロール(たとえば、ライトスイッチ、ドアノブ、サーモスタット、オープン/コントロールなど)、および/または他のタイプのオブジェクトなど、様々なオブジェクトを検出することができる。いくつかのスマートデバイスでは、スマートデバイス、およびスマートデバイスのための物理的コントロールは、同じであり得る(たとえば、スマートサーモスタット、スマートコーヒーメーカーなど)。以下で説明するように、オブジェクト認識器125は、スマートデバイスコントロールセクタ130に、もしあれば、画像データ123内で識別されたスマートデバイスを識別する、識別されたオブジェクトデータ126を提供することができる。

30

40

#### 【0039】

いくつかの実装形態では、オブジェクト認識器125は、ピクセルデータセットがオブジェクトの1つまたは複数の特定のクラス(たとえば、カテゴリー)内のオブジェクトを含むかを決定する、粗い分類器(coarse classifier)を含む。たとえば、粗い分類器は、実際のオブジェクトの認識の有無にかかわらず、ピクセルデータセットが特定のクラス(たとえば、照明コントロールなど、スマートデバイスのクラス)のオブジェクトを含むことを検出し得る。いくつかの態様では、オブジェクト認識器125は、画像データ123から、スマートデバイスが画像データ内にあるかを決定し、識別されたスマートデバイスに固有のデータを生成することができる。

50

## 【0040】

粗い分類器は、画像がオブジェクトのクラスを示す1つまたは複数の特徴を含むか否かに基づいて、オブジェクトのクラスの存在を検出することができる。粗い分類器は、そのオブジェクトのクラス内のオブジェクトの存在を検出するために、計算量の低い解析を実施するための軽量モデルを含み得る。たとえば、粗い分類器は、オブジェクトの各クラスについて、画像がオブジェクトのクラス内に入るオブジェクトを示すかを決定するために、画像内に示された制限された視覚特徴のセットを検出することができる。具体的な例では、粗い分類器は、画像が、以下のクラス、すなわち、スマートスピーカー、アプライアンス、テレビ、物理的コントロール(たとえば、ライトスイッチ、ドアノブ、サーモスタット、オープン/コンロコントロール)のうちの1つまたは複数に分類されるオブジェクトを示すかを検出することができる。

10

## 【0041】

いくつかの実装形態では、粗い分類器は、画像の視覚特徴に基づいて、画像を分類するために、トレーニングされた機械学習モデル(たとえば、畳み込みニューラルネットワーク)を使用する。たとえば、機械学習モデルは、それらのそれぞれのクラスでラベル付けされる、ラベル付けされた画像を使用してトレーニングされ得る。機械学習モデルは、オブジェクトのクラスの特定のセットのうちの0個以上に画像を分類するようにトレーニングされ得る。機械学習モデルは、入力として、画像の視覚特徴に関連するデータを受信し、オブジェクトのクラスの特定のセット内のオブジェクトのクラスのうちの0個以上への分類を出力することができる。

20

## 【0042】

粗い分類器は、オブジェクトのクラスが画像内で検出されたかを指定するデータを出力することができる。粗い分類器はまた、オブジェクトのクラスの存在が画像内で検出されたという信頼性を示す信頼値、および/または、実際のオブジェクト、たとえば、特定のタイプのスマートスピーカーが画像内に示されるという信頼性を示す信頼値を出力することもできる。

## 【0043】

図1Bに示されるように、たとえば、ユーザは、モバイルデバイス110のカメラ111をオブジェクト115に向けている。オブジェクト115は、おそらく、ユーザが仮想アシスタントアプリケーション116にそのためのコントロールを生成および提示してほしいと望むスマートデバイスである。仮想アシスタントアプリケーション116を使用して、ユーザは、オブジェクト115がカメラ111の視野内にあるように、および、オブジェクト115が仮想アシスタントアプリケーション116のユーザインターフェース136内でユーザに提示されるように、モバイルデバイス110を保持している。たとえば、仮想アシスタントアプリケーション116のユーザインターフェース136は、カメラ111のためのビューファインダーを含み得る。具体的な例では、仮想アシスタントアプリケーション116が、画像内で検出されたスマートデバイスのためのユーザインターフェースコントロールを仮想アシスタントアプリケーション116が提示することができる画像モード内にあり、かつ/または画像内で検出されたオブジェクトに関連するコンテンツを提供するとき、仮想アシスタントアプリケーション116は、ビューファインダーを提示することができる。

30

40

## 【0044】

カメラ111の視野内にある(かつ、ユーザインターフェース136のビューファインダー内で提示される)シーンを表す画像データ123が、オブジェクト認識器125によって受信される。ロケーションデータもまた、たとえば、モバイルデバイス110の全地球測位システム(GPS)受信機から、スマートデバイスコントロールセクタ130によって受信され得る。オブジェクト認識器125は、スマートデバイスコントロールセクタ130に、オブジェクトが画像データ123内で検出されたかを指定する、識別されたオブジェクトデータ126と、オブジェクトが認識された場合(特定のスマートデバイスなど)、認識されたオブジェクトを識別するデータとを提供することができる。粗い分類器が使用されるとき、オブジェクト認識器125は、スマートデバイスコントロールセクタ130に、オブジェクトのクラスのうち

50

の少なくとも1つ内のオブジェクトの存在が検出されたか否か、およびそうである場合、検出されたクラスを指定するデータを提供することができる。上記で説明したように、仮想アシスタントアプリケーション116は、ピクセルデータセットのストリームを受信することができる。この例では、オブジェクト認識器125は、各ピクセルデータセットを評価し、ストリーム内のピクセルデータセットの各々(または、少なくとも一部分)について、スマートデバイスコントロールセクタ130に、このデータ(オブジェクトが検出されたか、およびいずれかの認識されたオブジェクトの識別情報を指定するデータ)を提供することができる。

#### 【0045】

いくつかの実装形態では、オブジェクト認識器125は、モバイルデバイス110から受信された画像データ内のオブジェクトを認識するために、トレーニングされた機械学習モデル(たとえば、畳み込みニューラルネットワーク)を使用する。たとえば、機械学習モデルは、それらのそれぞれのスマートデバイス、および/またはスマートデバイスのための物理的コントロールでラベル付けされる、ラベル付けされた画像を使用してトレーニングされ得る。機械学習モデルは、画像データによって表された画像内に示されたスマートデバイス、および/またはスマートデバイスのための物理的コントロールを認識し、それを識別するデータを出力するようにトレーニングされ得る。機械学習モデルは、入力として、画像の視覚特徴に関連するデータを受信し、画像内に示されたスマートデバイス、またはそのための物理的コントロールが画像内に示されるスマートデバイスを識別するデータを出力することができる。

#### 【0046】

スマートデバイスが認識される場合、スマートデバイスコントロールセクタ130は、認識された実際のスマートデバイス、ピクセルデータセット内で検出されたスマートデバイスのクラス、スマートデバイスの能力、および/またはスマートデバイスが動作中であるコンテキストに基づいて、認識されたスマートデバイスのためのユーザインターフェースコントロールを選択することができる。たとえば、スマートデバイスが(ビューファインダーのためのピクセルデータセットに基づいて)ビューファインダー内で検出される場合、ビューファインダー内に検出可能なスマートデバイスがない場合よりも、ビューファインダー内にあるものに基づいて、ユーザがユーザインターフェースコントロールを要求中である可能性がより高い。したがって、スマートデバイスコントロールセクタ130は、スマートデバイスまたはスマートデバイスのクラスがピクセルデータ内で検出される場合、ユーザインターフェースコントロールを選択することができる。

#### 【0047】

各スマートデバイスまたはスマートデバイスのクラスは、1つまたは複数の対応するユーザインターフェースコントロールを含み得る。各スマートデバイスまたはスマートデバイスのクラスのためのユーザインターフェースコントロールは、ユーザインターフェースコントロールインデックス140内に記憶され得る。たとえば、スマートライトは、ユーザがライトをオンおよびオフにすることを可能にする、対応するユーザインターフェースコントロールを有し得る。調光可能なライトは、調光不可能なライトとは異なるユーザインターフェースコントロールを有し得る。たとえば、調光可能なライトのためのユーザインターフェースは、ユーザがライトの強度を調整するために調整することができる、回転調光器またはスライダーバーを含み得る。ユーザがカメラ111をトグルライトスイッチに向け、オブジェクト認識器125が、カメラ111がトグルライトスイッチに向いていることを検出する場合、スマートデバイスコントロールセクタ130は、ライトが、オンおよびオフにすることはできるが、調光することはできないライトであると決定することができる。したがって、スマートデバイスコントロールセクタ130は、ユーザがライトをオンおよびオフにするが、ライトの強度を調整しないことを可能にする、ユーザインターフェースコントロールを選択することができる。同様に、オブジェクト認識器125がライト調光器を認識する場合、スマートデバイスコントロールセクタ130は、ユーザがライトの強度を調整することを可能にする、ユーザインターフェースコントロールを選択することがで

きる。

#### 【0048】

仮想アシスタント制御デバイスなど、いくつかのスマートデバイスは、複数の異なるタスクを実施するか、または複数のアプリケーションを実行することができる。スマートデバイスコントロールセクタ130は、スマートデバイスが現在動作中であるコンテキストに基づいて、これらのデバイスのためのユーザインターフェースコントロールを選択することができる。たとえば、オブジェクト認識器125がカメラ111のためのビューファインダー内でスマートスピーカーを認識するとき、スマートスピーカーが音楽を再生中である場合、スマートデバイスコントロールセクタ130は、音楽を制御するためのユーザインターフェースコントロール、たとえば、音量の調整、異なるソングの再生、早送りなどのためのコントロールを選択することができる。スマートスピーカーが現在、アプライアンス(たとえば、オーブン)を制御中である場合、スマートデバイスコントロールセクタ130は、ユーザが(たとえば、オーブンの温度を変更するために)アプリケーションを制御することを可能にする、ユーザインターフェースコントロールを選択することができる。

10

#### 【0049】

いくつかの実装形態では、スマートスピーカーは、スマートスピーカーを動作させるための対応するユーザインターフェースを有し得る。このユーザインターフェースは、グラフィカルディスプレイを装備してくるスマートスピーカーによって提示される、ユーザインターフェースに対応し得る。たとえば、いくつかのスマートスピーカーは、オプションのディスプレイを含み得る。この例では、ユーザが、ディスプレイを有していないスマートスピーカーを有する場合、ディスプレイによって提示されることになるユーザインターフェースに対応するユーザインターフェースは、ユーザがカメラ111をスマートスピーカーに向けるとき、仮想アシスタントアプリケーション116によって提示され得る。

20

#### 【0050】

仮想アシスタント制御デバイス(たとえば、スマートスピーカー)のためのユーザインターフェースは、ユーザのアカウントを用いて登録された各スマートデバイスのためのユーザインターフェースコントロールを含み得る。たとえば、仮想アシスタントアプリケーションは、カメラ111のビューファインダーを表す画像データ123内の仮想アシスタント制御デバイスの認識、およびモバイルデバイスが(以下で説明するように)仮想アシスタント制御デバイスのしきい値距離内であるとの決定にตอบสนองして、仮想アシスタント制御デバイスを制御するためのユーザインターフェースコントロールを選択および提示することができる。これらのユーザインターフェースコントロールは、ユーザが、ユーザのアカウントを用いて登録されたスマートデバイスから選択することを可能にすることができる。たとえば、ユーザが、スマートオーブンおよびスマートライトを制御するように、仮想アシスタント制御デバイスを構成した場合、ユーザインターフェースコントロールは、ユーザがこれらのデバイスのいずれかを選択することを可能にすることができる。選択にตอบสนองして、仮想アシスタントアプリケーション116は、選択されたデバイスを制御するためのユーザインターフェースコントロールを提示することができる。別の例では、仮想アシスタントアプリケーション116は、ユーザがマスタパネルから各スマートデバイスを制御することを可能にする、すべて(または、少なくとも複数のスマートデバイス)のためのマスタパネルを提示することができる。

30

40

#### 【0051】

いくつかの実装形態では、仮想アシスタントアプリケーション116は、ユーザが、登録されたスマートデバイスから選択すること、および、仮想アシスタント制御デバイスが画像データ123内で認識されるときに提示されることになるスマートデバイスのためのユーザインターフェースコントロールを構成することを、可能にすることができる。たとえば、仮想アシスタントアプリケーション116は、登録されたスマートデバイスのリストを提示し、ユーザがリストから選択することを可能にすることができる。それにตอบสนองして、仮想アシスタントアプリケーション116は、仮想アシスタント制御デバイスが画像データ123内で認識されるたびに、選択されたスマートデバイスのためのユーザインターフェースコ

50



ントロールを提示することができる。

#### 【0052】

仮想アシスタントアプリケーション116はまた、画像データ123内の他のオブジェクトの認識に応答して、ユーザがコンテンツまたはユーザインターフェースコントロールの提示を引き起こすことを可能にすることもできる。たとえば、仮想アシスタントアプリケーション116は、ユーザが、特定のウォールまたは他のオブジェクトが画像データ123内で検出されるときは常に、仮想アシスタントアプリケーション116のユーザインターフェース内でARにおいてコンテンツ(たとえば、1つまたは複数のウェブページまたは他のリソースから取得された気象情報、スポーツ情報、金融情報など)を提示する、ダッシュボードを構成することを可能にすることができる。別の例では、ユーザは、特定のウォールまたは他のオブジェクトが画像データ123内で検出されるときは常に、仮想アシスタントアプリケーション116のユーザインターフェース内でARにおいて、ライブのウェブページまたはアプリケーションを提示するように、仮想アシスタントアプリケーション116を構成することができる。ユーザは、カメラ111をウォールまたは特定のオブジェクトに向けること、コンテンツを特定のオブジェクトに割り当てるために、アイコンまたは他のユーザインターフェースコントロールを選択すること、および次いで(たとえば、1つまたは複数のURLを提供すること、ブックマークを選択することなどによって)コンテンツを選択することによって、これらの構成を行うことができる。

10

#### 【0053】

いくつかの実装形態では、ユーザは、スマートデバイスの画像と組み合わせて、音声コマンドを使用して、仮想アシスタントアプリケーション116によって制御されることになるスマートデバイスを登録することができる。たとえば、ユーザは、カメラ111をスマートデバイスに向けて、音声コマンド(たとえば、これは自分のステレオです)を話すことができる。それに応答して、仮想アシスタントアプリケーション116は、スマートデバイスの画像をスマートデバイスの名称(たとえば、ステレオ)に関連付けることができる。ユーザがカメラ111をスマートデバイスに向ける場合、仮想アシスタントアプリケーション116は、スマートデバイスのためのユーザインターフェースコントロールを識別し、ユーザがコントロールを使用して、スマートデバイスを制御することを可能にすることができる。

20

#### 【0054】

いくつかの実装形態では、仮想アシスタントアプリケーション116は、音声コマンドに  
応答して、特定のロケーションのためのスマートデバイスのインスタンスを作成する。た  
とえば、仮想アシスタントアプリケーション116はまた、音声コマンドの検出に応答して  
、スマートデバイスのロケーションを識別し、ロケーションをスマートデバイスの画像お  
よびスマートデバイスの名称に関連付けることもできる。スマートデバイスが(たとえば  
、スマートデバイスのロケーションからしきい値距離よりも大きい)異なるロケーション  
にあるとき、ユーザがカメラ111をスマートデバイス(または、その複製)に向ける場合、  
仮想アシスタントアプリケーション116は、スマートデバイスのためのユーザインターフ  
ェースコントロールを提示しないように決定することができる。モバイルデバイス110が  
スマートデバイスの近く(たとえば、スマートデバイスのしきい値距離内)にあるとき、ユ  
ーザがカメラ111をスマートデバイスに向ける場合、仮想アシスタントアプリケーション1  
16は、スマートデバイスのためのユーザインターフェースコントロールを提示することが  
できる。

30

40

#### 【0055】

スマートデバイスコントロールセクタ130はまた、たとえば、ストリームにおける、  
ピクセルデータセットのシーケンスに基づいて、ユーザインターフェースコントロールを  
選択することもできる。たとえば、ピクセルデータセット内で認識されたオブジェクトが  
、短い時間期間(たとえば、2~5秒)の間に変化する場合、ユーザがモバイルデバイス110  
をあちこちに動かしており、特定のスマートデバイスのためのコントロールを取得しよう  
としていない可能性が高いことがある。しかしながら、同じスマートデバイスがピクセル  
データセットのシーケンス内で認識される場合、ユーザが特定のスマートデバイスのため

50

のコントロールを要求している可能性がより高い。

#### 【0056】

スマートデバイスコントロールセクタ130はまた、たとえば、仮想アシスタント制御デバイスに対する、モバイルデバイス110のロケーションに基づいて、ユーザインターフェースコントロールを提示し、および/またはユーザインターフェースコントロールを選択するかを決定することもできる。たとえば、画像がモバイルデバイス110によって提示されているとき、モバイルデバイス110がユーザの仮想アシスタント制御デバイスの近く(たとえば、しきい値距離内)にある場合、ユーザがビューファインダー内のスマートデバイスのためのコントロールを要求しており、特定のユーザインターフェースコントロールが選択され得る可能性が高い。モバイルデバイスが、ユーザの仮想アシスタント制御デバイスから離れている(たとえば、しきい値距離よりも遠く離れている)場合、ビューファインダーが、ユーザに関連付けられたスマートデバイスを見ている可能性は低い。これはまた、仮想アシスタントアプリケーション116が、他のユーザの仮想アシスタント制御デバイスを制御することが可能ではないことがあるので、その別のユーザの仮想アシスタント制御デバイスが偶発的に画像内に取り込まれる状況において、仮想アシスタントアプリケーション116によるユーザインターフェースコントロールの提示を防止する。この状況では、ユーザインターフェースコントロールの提示は、ユーザを失望させることがある。

10

#### 【0057】

別の例では、ユーザが、たとえば、仮想アシスタント制御デバイスからリモートのロケーションからスマートデバイスを制御することができるよう、ユーザは、以前に取り込まれた画像にアクセスし得る。この例では、スマートデバイスコントロールセクタ130は、画像が取り込まれたときのモバイルデバイス110のロケーションにアクセスすることができる。このロケーションは、たとえば、メタデータとして、画像とともに記憶され得る。画像が取り込まれたときのモバイルデバイス110のロケーションが、仮想アシスタント制御デバイスのロケーションのしきい値距離内である場合、スマートデバイスコントロールセクタ130は、モバイルデバイス110が画像内で認識されたスマートデバイスを制御することができると決定し、スマートデバイスを制御するためのユーザインターフェースコントロールを提示することができる。

20

#### 【0058】

モバイルデバイス110が仮想アシスタント制御デバイスのしきい値距離内であるかを決定するために、仮想アシスタントアプリケーション116は、モバイルデバイスのGPS受信機から、モバイルデバイス110のためのロケーションデータを取得し、仮想アシスタント制御デバイスから、仮想アシスタント管理システム160から、または仮想アシスタント制御デバイスの初期構成からのいずれかで、仮想アシスタント制御デバイスのためのロケーションデータを取得することができる。一例では、仮想アシスタント制御デバイスは、仮想アシスタント管理システム160に、その現在のロケーションを指定するデータを周期的に送ることができる。仮想アシスタントアプリケーション116は、モバイルデバイス110のロケーションを仮想アシスタント制御デバイスのロケーションと比較して、モバイルデバイス110が仮想アシスタント制御デバイスのしきい値距離内であるかを決定することができる。

30

40

#### 【0059】

スマートデバイスコントロールセクタ130はまた、認識されたスマートデバイスに一致するスマートデバイスが、ユーザのアカウント(たとえば、ユーザの仮想アシスタントアカウント)を用いて登録されたかに基づいて、スマートデバイスのためのユーザインターフェースコントロールを提示するかを決定することもできる。このように、仮想アシスタントアプリケーション116は、仮想アシスタントアプリケーション116が制御するようにセットアップされていないスマートデバイスのためのユーザインターフェースコントロールを提示しない。たとえば、ユーザは、仮想アシスタントアプリケーション116を使用して、ユーザの自宅におけるスマートスピーカーなど、仮想アシスタント制御デバイスによる制御のためのスマートデバイスを登録することができる。仮想アシスタントアプリケー

50

ション116は、各登録されたスマートデバイスについてのデータをローカルに、たとえば、コントロールインデックス140または別のインデックス内に記憶することができる。データは、デバイスの名称、デバイスのクラス、デバイスの能力、デバイスが、たとえば、仮想アシスタント制御デバイスとデバイスとの間のワイヤレス接続を介して、どのように制御されるか、および/または他の適切なデータを含み得る。いくつかの実装形態では、デバイスのためのデータはまた、デバイスの画像、および/またはデバイスを制御するための物理的コントロールの画像も含み得る。仮想アシスタントアプリケーション116はまた、ユーザの仮想アシスタントアカウントを管理する仮想アシスタント管理システム160に、データを提供することもできる。

#### 【0060】

スマートデバイスコントロールセクタ130はまた、モバイルデバイス110のロケーション、たとえば、モバイルデバイス110の地理的ロケーションに基づいて、ユーザインターフェースコントロールを選択することもできる。たとえば、モバイルデバイス110は、モバイルデバイス110の地理的ロケーションを決定するGPS受信機を含み、モバイルデバイス110のロケーションを指定するロケーションデータ124を、仮想アシスタントアプリケーション116に提供することができる。スマートデバイスコントロールセクタ130は、地理的ロケーションを使用して、モバイルデバイスが、ユーザの仮想アシスタント制御デバイスが位置する、それらの自宅などの既知のロケーション、またはユーザが登録された仮想アシスタント制御デバイスを有する他のエリアにあるかを決定することができる。モバイルデバイス110が、仮想アシスタント制御デバイスに関連付けられたホームネットワークなど、これらのロケーションのうちの1つ(または、そこからしきい値距離内)にあるとき、スマートデバイスコントロールセクタ130は、それらが各々同じネットワークに関連付けられる場合、仮想アシスタント制御デバイスとともに、識別されたスマートデバイスを登録することができる。

#### 【0061】

スマートデバイスコントロールセクタ130は、識別されたスマートデバイスまたは同様のスマートデバイスのためのユーザインターフェースコントロールを使用するユーザの履歴に基づいて、識別されたスマートデバイスを制御するためのユーザインターフェースコントロールを選択することができる。スマートデバイスコントロールセクタ130はまた、ユーザによって指定されたユーザインターフェースコントロールのための選好に基づいて、識別されたスマートデバイスのためのユーザインターフェースコントロールを選択することもできる。たとえば、ユーザは、特定の状況について、どのユーザインターフェースコントロールをユーザが選好するかを選択することができる。具体的な例では、仮想アシスタントアプリケーション116は、同じオブジェクトまたは同じオブジェクトのクラスのための様々なユーザインターフェースコントロールを生成することができる。ユーザは、様々なユーザインターフェースコントロールの間で選択することができ、仮想アシスタントアプリケーション116は、ユーザのための選択を記憶することができる。

#### 【0062】

いくつかの態様によれば、仮想アシスタントアプリケーション116は、本明細書で説明するプロセスのうちの1つまたは複数を実施するために、ネットワーク150を介して、仮想アシスタント管理システム160に、画像データ123および/またはロケーションデータ124を送ることができる。たとえば、画像データ123およびロケーションデータ124を用いて、仮想アシスタント管理システム160は、受信された画像データがスマートデバイスを示すかを決定し、画像データ123が取り込まれたときのモバイルデバイス110のロケーションが、ユーザのアカウントを用いて登録された仮想アシスタント制御デバイス(たとえば、スマートスピーカー)のロケーションのしきい値距離内であると決定し、モバイルデバイス110が仮想アシスタント制御デバイスのしきい値距離内であった場合に、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを選択することができる。仮想アシスタントアプリケーション116は、仮想アシスタント管理システム160に、1つまたは複数のピクセルデータセットを送ることができる。ピクセルデータセットは

、スマートデバイスコントロールを選択するために使用されたビューファインダーのためのピクセルデータ、および/またはスマートデバイスコントロールが選択された後に取り込まれたピクセルデータセットを含み得る。たとえば、仮想アシスタントアプリケーション116は、スマートデバイスコントロールを選択するために使用されたピクセルデータセットを最初に送ることができる。ビューファインダーが、後に異なるスマートデバイスに向いている場合、仮想アシスタントアプリケーション116は、仮想アシスタント管理システム160に、異なるスマートデバイスのためのピクセルデータを送ることができる。

#### 【0063】

仮想アシスタント管理システム160は、1つまたは複数のフロントエンドサーバ161と、1つまたは複数のバックエンドサーバ162とを含む。フロントエンドサーバ161は、モバイルデバイス110からデータを受信し、バックエンドサーバ162にデータを提供することができる。フロントエンドサーバ161はまた、データの受信に応答して、モバイルデバイス110にコンテンツを送ることもできる。

#### 【0064】

バックエンドサーバ162は、モバイルデバイス110から受信されたデータに基づいて、ユーザインターフェースコントロールを選択し、モバイルデバイス110に提供する、スマートデバイスコントロール選択エンジン165を含む。本明細書で使用するエンジンという用語は、タスクのセットを実施するデータ処理装置を指す。スマートデバイスコントロール選択エンジン165は、複数のスマートデバイスおよび/またはスマートデバイスのクラスのためのユーザインターフェースコントロールを記憶する、スマートデバイスコントロール記憶ユニット166(たとえば、1つまたは複数のハードドライブ、フラッシュメモリなど)から、ユーザインターフェースコントロールを選択することができる。スマートデバイスコントロール選択エンジン165は、仮想アシスタントアプリケーション116のオブジェクト認識器125およびスマートデバイスコントロールセクタ130と同じ動作、または同様の動作を実施することができる。たとえば、スマートデバイスコントロールセクタエンジン165は、画像データ123がスマートデバイス(または、スマートデバイスのための物理的コントロール)を含むかを決定することができる、オブジェクト認識モジュールを含み、そうである場合、認識されたスマートデバイスのためのユーザインターフェースコントロールを選択することができる。

#### 【0065】

たとえば、モバイルデバイス110から受信された画像データ123が、仮想アシスタント制御デバイス(たとえば、スマートスピーカー)の画像を表し、モバイルデバイス110から受信されたオーディオが、音楽を表す場合、スマートデバイスコントロール選択エンジン165は、音声アシスタントデバイスによって表されたスマートデバイスコントロールを識別し、その特定の音声アシスタントデバイスに関連するオーディオのためのスマートデバイスコントロール(たとえば、図2に示されるような、調整可能な音量トグル、送る、戻す、タイトル、いいね/よくないねボタンなど)を選択することができる。モバイルデバイス110から受信された画像データ123が、ライトまたはライトコントローラ(たとえば、物理的なライトスイッチ)の画像を表す場合、スマートデバイスコントロール選択エンジン165は、ライトを制御するためのユーザインターフェースコントロール(たとえば、図3に示されるような、仮想の調整可能な調光器トグル、ON/OFFボタンなど)を選択することができる。

#### 【0066】

いくつかの実施形態では、スマートデバイスコントロール選択エンジン165がそのためのユーザインターフェースコントロールを生成している特定のスマートデバイスが、仮想アシスタント制御デバイスの登録されたユーザであるモバイルデバイス110のユーザのアカウントに登録されていると決定され、ユーザインターフェースコントロールが、仮想アシスタント制御デバイスが制御することが可能である、音声によりアクティブ化されるコントロールに関連付けられる。

#### 【0067】

バックエンドサーバ162はまた、ユーザのアカウントを管理するアカウント管理エンジン167も含む。アカウント管理エンジン167は、アカウントを有する各ユーザについて、ユーザの仮想アシスタント制御デバイス、および仮想アシスタント制御デバイスによって制御されるために登録されているスマートデバイスを指定する、登録データ168を記憶することができる。アカウント管理エンジン167は、各ユーザの仮想アシスタントアプリケーション116から、データを受信することができる。各スマートデバイスのための登録データは、デバイスの名称、デバイスのクラス、デバイスの能力、デバイスが、たとえば、仮想アシスタント制御デバイスとデバイスとの間のワイヤレス接続を介して、どのように制御されるか、デバイスの画像、および/またはデバイスを制御するための物理的コントロールの画像を含み得る。

10

**【0068】**

バックエンドサーバ162は、フロントエンドサーバ161に、スマートデバイスコントロール選択エンジン165によって選択されたユーザインターフェースコントロールを提供することができる。フロントエンドサーバ161は、次に、そこからスマートデバイスコントロールを選択するために使用されたデータが受信されたモバイルデバイス110に、スマートデバイスコントロールを提供することができる。

**【0069】**

いくつかの実施形態では、バックエンドサーバ162は、モバイルデバイス110から受信されたデータに応答して、登録ファイルを選択し、モバイルデバイス110に提供する、スマートデバイス登録エンジンを含み得る。スマートデバイス登録エンジンは、複数のスマートデバイスのための登録ファイルを選択することができる。たとえば、スマートデバイスコントロール選択エンジン165が、ピクセルデータ内で表されたスマートデバイスを識別することができる場合、スマートデバイス登録エンジンは、登録ファイルが、モバイルデバイス110のユーザのアカウントを用いて、そのスマートデバイスに関連付けられるかを決定することができる。たとえば、スマートデバイスのための相手先商標製造会社(OEM)は、ユーザがそのホームネットワークを通して、デバイスを登録することを要求することができる。いくつかの実施形態では、OEMは、ユーザの仮想アシスタント制御デバイスを通して、スマートデバイスの関連付けおよび登録を行うことができるか、またはスマートデバイスが、OEMによって管理されたネイティブアプリケーションもしくはウェブブラウザによってアクセスされ得る。いくつかの態様では、仮想アシスタント管理システム160は、OEMがユーザのホームネットワークを通してアクセスされることを可能にする、登録ファイルを記憶することができる。

20

30

**【0070】**

いくつかの実施形態では、仮想アシスタント管理システム160は、特定のスマートデバイスがモバイルデバイス110のユーザのアカウントに登録されているかを決定することによって、関連付けられた仮想アシスタント制御デバイスなしに、特定のモバイルデバイス110が仮想アシスタントアプリケーション116を通して、スマートデバイスを制御することを可能にすることができる。いくつかの実装形態では、ユーザは、その特定のスマートデバイスにアクセスするために、登録ファイル内で許可を与えられる必要があり得る。たとえば、ユーザは、仮想アシスタント制御デバイスを通して接続することなしに、OEMセットアッププロセスを通して、スマートデバイスライトコントローラをセットアップし得る。仮想アシスタントアプリケーション116は、スマートデバイス登録エンジンにアクセスし、識別されたスマートデバイスに関連付けられた登録ファイルを取り出し、モバイルデバイス110のロケーションデータ124、および登録ファイル内のユーザに関連付けられたアカウント情報に基づいて、スマートデバイスライトコントローラを制御することができる。いくつかの実施形態では、仮想アシスタントアプリケーションは、モバイルデバイス上で登録ファイルを記憶することができる。

40

**【0071】**

仮想アシスタントアプリケーション116は、グラフィカルユーザインターフェースを生成および提示する、ユーザインターフェース生成器135を含む。ユーザインターフェース

50

は、カメラ111のためのビューファインダー内に表された画像データ123内で認識されたスマートデバイスを制御するためのユーザインターフェースコントロールを含み得る。上記で説明したように、仮想アシスタントアプリケーション116のユーザインターフェースは、カメラ111のためのビューファインダーを提示することができる。図1Bに示されるように、スマートデバイスコントロールセクタ130は、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを選択した後、インターフェースコントロール識別データ131を生成し、ユーザインターフェース生成器135に送る。インターフェースコントロール識別データ131は、インターフェース生成器135が選択されたユーザインターフェースコントロールを提示するために使用することができる、選択されたユーザインターフェースコントロールのためのデータ(たとえば、異なるタイプのコントロール、コントロールとして使用される異なるアイコン、異なるレイアウトなど)を含む。ユーザインターフェース生成器135は、インターフェースコントロール識別データ131に少なくとも部分的に基づいて、モバイルデバイス110のディスプレイにおいて、ユーザインターフェース136を生成、更新、および提示する。たとえば、ユーザインターフェース生成器135は、(たとえば、カメラ111のビューファインダー内で)ライブ画像上に、選択されたユーザインターフェースコントロールを提示することができる。

10

#### 【0072】

ユーザインターフェース生成器135は、ユーザインターフェースを定義する提示データ138を生成し、モバイルデバイス110のディスプレイに提示データ138を出力することができる。提示データ138は、ユーザインターフェース136内の特定のロケーションにおいてユーザインターフェースコントロールを提示することを、ディスプレイに行わせるデータを含み得る。たとえば、提示データ138は、ユーザインターフェースコントロールがユーザインターフェース136内で提示されることになるロケーションを(たとえば、ピクセル座標を使用して)指定することができる。ユーザインターフェースコントロールのロケーションは、認識されたスマートデバイスの近くで、たとえば、ビューファインダー内で認識されたスマートデバイスの下に、またはそれに隣接して提示され得る。別の例では、ユーザインターフェースコントロールは、ビューファインダーの下に、またはそれに隣接して提示され得る。

20

#### 【0073】

次いで、ユーザは、スマートデバイスを制御するために、ユーザインターフェースコントロールと対話することができる。いくつかの実装形態では、仮想アシスタントアプリケーション116は、ユーザ対話を検出し、ユーザの仮想アシスタント制御デバイスに、ユーザ対話を指定するデータを提供する。たとえば、モバイルデバイス110は、ワイヤレスネットワークを介して、仮想アシスタント制御デバイスにデータを送信することができる。次いで、仮想アシスタント制御デバイスは、ワイヤレスネットワークを介して、デバイスを制御することができる。別の例では、仮想アシスタント制御デバイスは、仮想アシスタント管理システム160に制御データを送ることができ、仮想アシスタント管理システム160は、次に、たとえば、ネットワーク150を介して、スマートデバイスに制御データを送ることによって、制御データに基づいて、スマートデバイスを制御する。

30

#### 【0074】

いくつかの実装形態では、仮想アシスタントアプリケーション116は、たとえば、仮想アシスタント制御デバイスにデータを送ることなしに、ネットワーク150を介して、仮想アシスタント管理システム160に制御データを送る。次いで、仮想アシスタント管理システム160は、上記で説明したように、スマートデバイスに制御データを送ることができる。

40

#### 【0075】

いくつかの実装形態では、仮想アシスタントアプリケーション116は、複数のユーザが互いに通信すること、および識別されたオブジェクト上での通信目的のために、メッセージまたは他のコンテンツ媒体を残すことを可能にすることができる。たとえば、同じホームネットワーク(または、同じ仮想アシスタントアカウント)のユーザは、冷蔵庫などの識

50

別されたオブジェクトを登録し、別のユーザまたは複数のユーザが仮想アシスタントアプリケーション116を使用してアクセスし、ビューファインダー内で見ると、(たとえば、拡張現実技法を使用する)仮想メモを置くことができる。いくつかの実施形態では、ユーザは、仮想メモへのアクセスを有するための特定のユーザを指定することができるか、またはユーザは、ホームネットワーク上の誰もが仮想メモを見ることを可能にすることができる。

#### 【0076】

図2は、受信された画像データに基づいて、識別されたスマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを提示する、モバイルデバイスの例示的なスクリーンショット210および220のシーケンスを示す。第1のスクリーンショット210は、モバイルデバイスの例示的なユーザインターフェース212を示す。この例では、ユーザは、ユーザインターフェース212内で提示されるカメラのビューファインダーを通して、音声アシスタントデバイス215(たとえば、スマートスピーカー)を見ており、音声アシスタントデバイス215は、音楽を再生中である。ユーザインターフェース212は、仮想アシスタントアプリケーション、たとえば、図1の仮想アシスタントアプリケーション116によって生成および提示され得る。

#### 【0077】

第2のスクリーンショット220は、仮想アシスタントアプリケーション116のためのユーザインターフェース222を提示する。この例では、ユーザインターフェース222は、ユーザインターフェース222の現在のビュー上にオーバーレイされたユーザインターフェースコントロール224を提示する。具体的には、ユーザインターフェースコントロール224は、この例では、音声アシスタントデバイス215である、識別されたオブジェクトの周囲にウィンドウを提示する。仮想アシスタントアプリケーション116は、ビューファインダーを表す画像データから、音声アシスタントデバイス215を識別することができる。上記で説明したように、仮想アシスタントアプリケーション116はまた、モバイルデバイスのロケーションが、ユーザの音声アシスタントデバイスのロケーションのしきい値距離内であるかに基づいて、音声アシスタントデバイスのためのユーザインターフェースコントロールを提示するかを決定することもできる。

#### 【0078】

ユーザインターフェースコントロール224内に、音楽のオーディオを制御するためのいくつかの例示的なスマートデバイスコントロールがある。具体的には、音量トグルバー226が、音声アシスタントデバイス215の音量レベルを制御するために示されている。さらに、音楽制御ボタン228が、ユーザに提示され得るいくつかの異なるオプションとともに示されており、限定はしないが、一時停止/再生ボタン、いいねおよびよくないねボタン、前のソングボタン、次のソングボタン、ならびに現在再生されているオーディオ内の特定のポイントに移動するためのプレースホルダートグルバーを含む。音楽制御ボタン228のオプションは、音声アシスタントデバイス215によって使用されている特定の音楽アプリケーションに依存し得る。いくつかの態様では、利用可能なボタンは、すべての音楽再生アプリケーションに汎用的であり得る。いくつかの態様では、ボタンは、特定の音楽再生アプリケーションに合わせてカスタマイズされてよく、仮想アシスタント管理システム160によって更新され得る。

#### 【0079】

仮想アシスタントアプリケーション116は、音声アシスタントデバイス215のコンテキストに基づいて、音楽コントロールを選択することができる。たとえば、仮想アシスタントアプリケーション116は、音声アシスタントデバイス215と通信し、現在の動作モード、音声アシスタントデバイス215によって実施されている現在のタスク、または音声アシスタントデバイス215によって現在使用中のアプリケーション(たとえば、音楽アプリケーション)を要求することができる。音声アシスタントデバイス215は、仮想アシスタントアプリケーション116に、現在の動作モード、または実施されている現在のタスクを指定するデータを提供することができ、次に、仮想アシスタントアプリケーション116は、ユーザが

現在の動作モードまたは現在のタスクを制御することを可能にする、ユーザインターフェースコントロールを選択することができる。この例では、音声アシスタントデバイス215は、音楽を再生中であり、仮想アシスタントアプリケーション116は、音声アシスタントデバイス215が音楽を再生中であることを指定するデータの受信に応答して、音楽コントロールを選択した。いくつかの実装形態では、仮想アシスタントアプリケーション116は、検出されたオーディオに基づいて、音声アシスタントデバイス215が音楽を再生中であると決定することができる。

#### 【0080】

ユーザが、音楽制御ボタン228と対話(たとえば、音楽制御ボタン228を選択)する場合、仮想アシスタントアプリケーション116は、音声アシスタントデバイス215と通信して、音楽再生アプリケーションを制御することができる。いくつかの態様によれば、ユーザが、ビューファインダーを音声アシスタントデバイス215から離れるように移動させる場合、仮想アシスタントアプリケーション116は、ユーザインターフェース222からユーザインターフェースコントロール224を除去することになり、その理由は、もはや画像データ内でスマートデバイス、具体的には、音声アシスタントデバイス215を検出しないからである。いくつかの態様では、仮想アシスタントアプリケーション116は、ユーザインターフェース222内の音声アシスタントデバイス215の画像をフリーズさせることができる。たとえば、仮想アシスタントアプリケーション116は、それと対話すると、仮想アシスタントアプリケーション116が画像をフリーズさせることを引き起こす、ロック画面またはロック画像ボタンをユーザに提示することができる。この例では、ユーザは、カメラをスマートデバイスに直接向ける必要なしに、モバイルデバイス110を移動させることができるが、ユーザインターフェースコントロール224を使用して、スマートデバイスを制御し続けることができる。いくつかの実装形態では、仮想アシスタントアプリケーション116は、コントロールのうちの1つとのユーザ対話に応答して、画像をフリーズさせることができ、その理由は、これが、ユーザが音声アシスタントデバイス215を制御するために、それらのコントロールを使用することを望むことを示すからである。

#### 【0081】

いくつかの態様では、ユーザインターフェースコントロール224は、コンテンツ、または仮想アシスタントアプリケーション116がコンテンツを識別することを試行中であることを示すアニメーションを提示することができる。たとえば、ユーザインターフェース222は、仮想アシスタントアプリケーションがユーザに提示するためのスマートデバイスコントロールを識別中であることをシグナリングする、ループ状のアニメーションになるスロバー(throbber)アニメーション、または別のアニメーションを提示することができる。

#### 【0082】

図3は、受信された画像データに基づいて、識別されたスマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを提示する、モバイルデバイスの例示的なスクリーンショット310、320、および330の別のシーケンスを示す。第1のスクリーンショット310は、モバイルデバイスの例示的なインターフェース312を示す。この例では、ユーザは、ビューファインダーを通して、スマートデバイスのための物理的コントロール315を見ている。この例では、スマートデバイスはスマートライトであり、物理的コントロール315はライトスイッチである。

#### 【0083】

第2のスクリーンショット320および第3のスクリーンショット330は、仮想アシスタントアプリケーション116のために、それぞれユーザインターフェース322および332を提示する。これらの例では、ユーザインターフェース322および332は、ユーザインターフェース322の現在のビュー上にオーバーレイされたユーザインターフェースコントロール324を提示する。具体的には、ユーザインターフェースコントロール324は、スマートライトのための物理的コントロール315の下にウィンドウを提示する。仮想アシスタントアプリケーション116は、モバイルデバイスのカメラのためのビューファインダーを表す画像データから、物理的コントロール315を識別した。図示のように、ユーザインターフェースコン



トロール324内に、スマートデバイスの照明を制御するための3つの例示的なスマートデバイスコントロールがある。図示のように、スクリーンショット320は、以前にユーザによって「ON」として最初に選択されているような照明コントロールとともに、スマートデバイスを示す(または、スマートライトの現在の状態を表す)。いくつかの例では、ユーザは、ライトコントロールをオンにするために、「ON」トグルボタン325を選択することができる。スクリーンショット330に示されるように、モバイルデバイスのユーザが、ライトをオフにするために、仮想アシスタントアプリケーション116を使用することを選ぶ場合、ユーザは、「OFF」トグルボタン334を選択する。さらに、調光トグルバー326が、スマートライトの調光レベルを制御するために示されている。

#### 【0084】

これらの例では、スマートデバイスは、ユーザのアカウントを用いて登録されており、図2における音声アシスタントデバイス215など、仮想アシスタント制御デバイスと通信することができ、仮想アシスタント制御デバイスによって制御され得る。仮想アシスタントアプリケーション116は、ユーザのモバイルデバイスのカメラから、画像データを受信し、受信された画像データがスマートデバイスを示すと決定することになる。次に、仮想アシスタントアプリケーション116は、画像データが取り込まれたときのモバイルデバイスのロケーションが、図2における音声アシスタントデバイス215など、仮想アシスタント制御デバイスのロケーションのしきい値距離内であると決定することができる。さらに、仮想アシスタントアプリケーション116は、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを識別し、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロール(すなわち、インターフェースコントロール324)を提示することができる。

#### 【0085】

いくつかの態様によれば、利用可能なボタンは、すべての照明制御スマートデバイスに汎用的であり得る。いくつかの態様では、ボタンは、特定の照明制御スマートデバイスに合わせてカスタマイズされてよく、仮想アシスタント管理システム160によって更新され得る。

#### 【0086】

いくつかの態様によれば、他のスマートデバイスコントロールが利用可能である(たとえば、第2のライトコントローラなど、第2のスマートデバイスが画像データ内にある)場合、ユーザインターフェース322は、ユーザが他のスマートデバイスコントロールを選択することを可能にする、ユーザインターフェースコントロールを含み得る。

#### 【0087】

いくつかの態様によれば、画像データ内のスマートデバイスは、コンロまたは電子レンジなど、アプライアンスのタイマーまたは時計である。この例では、仮想アシスタントアプリケーション116は、タイマーを設定するためのユーザインターフェースコントロールを提示するか、または迅速にスマートデバイスの時計の時間を変更するため、もしくはアラームを設定するための容易なコントロールを提供することができる。

#### 【0088】

いくつかの態様によれば、画像データ内のスマートデバイスは、ディスプレイデバイス上に提示されるか、または背景上に投影される、デジタルダッシュボードアプリケーションである。たとえば、スマートデバイスは、リビングルームの壁において、仮想ダッシュボード(たとえば、気象、ニュース、スポーツ、金融など)を提示する、プロジェクトを使用することができる。モバイルデバイスは、ダッシュボードの画像データを獲得し、仮想アシスタントアプリケーション116にデータを提供することになり、仮想アシスタントアプリケーション116は、次いで、ダッシュボードを更新または変更するためのインターフェースコントロールを提供することができる。

#### 【0089】

図4は、仮想アシスタントアプリケーションを使用して、スマートデバイスを識別し、識別されたスマートデバイスを制御するための1つまたは複数のユーザインターフェース

10

20

30

40

50

コントロールを提示するための、例示的なプロセス400のフロー図である。プロセス400の動作は、たとえば、図1Aおよび図1Bのモバイルデバイス110など、1つまたは複数のデータ処理装置によって実施され得る。プロセス400の動作はまた、非一時的コンピュータ可読媒体上に記憶された命令としても実装され得る。命令の実行によって、1つまたは複数のデータ処理装置にプロセス400の動作を実施させる。

#### 【0090】

ユーザのモバイルデバイスのカメラからの画像データが受信される(402)。本明細書で説明するように、仮想アシスタントアプリケーション116のオブジェクト認識器125は、モバイルデバイスのカメラのビューファインダーのためのピクセルデータ(たとえば、図1Bに示されるような画像データ123)を受信することができる。画像データは、単一の画像、またはストリーミングライブビデオなどの連続した画像であり得る。

10

#### 【0091】

画像データが、モバイルデバイスのカメラから受信された後、受信された画像データがスマートデバイスまたはスマートデバイスのための物理的コントロールを示すかについての決定が行われる(404)。スマートデバイスは、仮想アシスタント制御デバイス(たとえば、音声アシスタントデバイス)、または別のタイプのスマートデバイスであり得る。たとえば、スマートデバイスは、たとえば、ユーザの仮想アシスタントアカウントを用いてスマートデバイスを登録すること、またはさもなければ、スマートデバイスを制御するように仮想アシスタント制御デバイスを構成することによって、仮想アシスタント制御デバイスによって制御されることが可能であるスマートデバイスであり得る。上記で説明したように、モバイルデバイス上の仮想アシスタントアプリケーションは、スマートデバイスまたはスマートデバイスのための物理的コントロールが画像データ内に示されているかの決定を行うことができる。いくつかの態様では、画像データは、リモートシステム、たとえば、仮想アシスタント管理システムによって解析され得、結果がモバイルデバイスに返送され得る。

20

#### 【0092】

いくつかの態様によれば、スマートデバイスが画像データ内に示されていると決定される場合、画像がモバイルデバイスにおいて提示されるときモバイルデバイスが、ユーザのアカウントを用いて登録された仮想アシスタント制御デバイスを制御することが可能であるかについての決定が行われる。たとえば、決定は、画像が提示されるときモバイルデバイスが、仮想アシスタント制御デバイスのしきい値距離内であるかに基づき得る。上記で説明したように、モバイルデバイスは、仮想アシスタントアプリケーションにロケーションデータを送ることができ、仮想アシスタントアプリケーションは、次いで、仮想アシスタント制御デバイス(たとえば、音声アシスタントデバイス)が極めて近接しているか、またはモバイルデバイスのロケーションのあらかじめ決定された距離内であるかを決定することができる。たとえば、仮想アシスタントアプリケーションは、モバイルデバイスのロケーションを、仮想アシスタント制御デバイスのロケーションと比較することができる。仮想アシスタント制御デバイスのロケーションは、(たとえば、仮想アシスタント制御デバイスが最初に登録されるときに決定された)指定されたロケーションであり得る。別の例では、仮想アシスタント制御デバイスは、(たとえば、仮想アシスタント制御デバイス上にインストールされているか、または仮想アシスタント制御デバイス内のGPS受信機を使用して)その現在のロケーションを識別するデータを、モバイルデバイスに、または(そこから仮想アシスタントアプリケーションがロケーションデータを取得することができる)仮想アシスタント管理システムに送信することができる。ロケーションデータの作成、およびスマートデバイスがユーザの仮想アシスタント制御デバイスのある距離内であるかの決定は、リモートシステム、たとえば、仮想アシスタント管理システムによって行うことができ、決定がモバイルデバイスに返送される。

30

40

#### 【0093】

いくつかの態様によれば、受信された画像が、ユーザのアカウントを用いて登録されたスマートデバイスを示すと決定することは、画像データがユーザの仮想アシスタント制御

50

デバイスの画像を示すと決定することを含む。いくつかの態様によれば、仮想アシスタント制御デバイスは、ユーザから音声コマンドを受信する仮想アシスタントスマートスピーカデバイスを備え、仮想アシスタントスマートスピーカデバイスのスピーカを使用して、ユーザに情報を提供する。

【0094】

いくつかの態様によれば、受信された画像が、ユーザのアカウントを用いて登録されたスマートデバイスを示すと決定することは、画像が、ユーザの仮想アシスタント制御デバイスによって制御されるスマートデバイスを示すと決定することを含む。たとえば、仮想アシスタントアプリケーションは、音声アシスタントデバイスと通信して、音声アシスタントデバイスが何のスマートデバイスとともにペアにされたかなど、音声アシスタントデバイスが何のスマートデバイスに対する制御を有するかを決定することができる。

10

【0095】

いくつかの態様によれば、受信された画像が、ユーザのアカウントを用いて登録されたスマートデバイスを示すと決定することは、ユーザの仮想アシスタントアカウントを用いて登録されたスマートデバイス、たとえば、ユーザがそれに対して制御するように仮想アシスタント制御デバイスを構成したスマートデバイスを指定するデータを取得することを含む。画像内で識別されたスマートデバイス(または、物理的コントロール)が、登録されたスマートデバイス(または、登録されたスマートデバイスのための物理的コントロール)に一致する場合、仮想アシスタントアプリケーションは、ユーザがモバイルデバイスのカメラを登録されたスマートデバイスに向けていると決定することができる。

20

【0096】

スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールが、画像がスマートデバイスまたはスマートデバイスのための物理的コントロールのうちの少なくとも1つを示すとの決定にตอบสนองして、識別される(406)。上記で説明したように、画像データが取り込まれたときのモバイルデバイスのロケーションが、ユーザのアカウントを用いて登録された仮想アシスタント制御デバイスのロケーションのある距離内であると、スマートデバイスコントロールセクタが決定した後、スマートデバイスコントロールセクタは、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを識別し、インターフェースコントロール識別データを生成する。インターフェースコントロール識別データは、識別されたスマートデバイスを制御することができ、いくつかのコントロールを指定するデータを含む。ユーザインターフェースコントロールは、リモートシステム、たとえば、仮想アシスタント管理システムによって選択され、モバイルデバイスに送られ得る。

30

【0097】

いくつかの態様によれば、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを識別することは、仮想アシスタント制御デバイスが現在実施中であるタスクまたはサービスを決定すること、および、その特定のタスクまたはサービスを制御するための1つまたは複数のユーザインターフェースコントロールを選択することを含む。たとえば、仮想アシスタント制御デバイスがスマートスピーカであり、スマートスピーカがアプライアンスを制御中である場合、アプライアンスを制御するためのコントロールが選択され得る。

40

【0098】

スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールが生成され、モバイルデバイスのユーザに提示される(408)。上記で説明したように、インターフェースコントロール識別データがインターフェース生成器によって受信され、インターフェースコントロール識別データによって、対象のスマートデバイスに属すると指定された各ユーザインターフェースコントロールについて、インターフェース生成器は、モバイルデバイス上のユーザインターフェース上に重ね合わせられたユーザインターフェースコントロールを提示する提示データを生成する。いくつかの態様では、連続したビデオ(たとえば、ライブ映像)の場合、ユーザインターフェースコントロールは、スマー

50

トデバイスが画像データ内で識別される間のみ、ユーザインターフェース上に重ね合わせられる。ユーザインターフェースコントロールは、リモートシステム、たとえば、仮想アシスタント管理システムによって選択され、表示のために、提示データとしてモバイルデバイスに送られ得る。

#### 【0099】

1つまたは複数のユーザインターフェースコントロールのうちの少なくとも1つとのユーザ対話が、モバイルデバイスのディスプレイにおいて検出される(410)。たとえば、仮想アシスタントアプリケーションは、対話を検出し、対話に基づいて、実施するための対応するアクションを決定することができる。

#### 【0100】

検出された、1つまたは複数のユーザインターフェースコントロールのうちの少なくとも1つとのユーザ対話に基づいて、スマートデバイスが制御される(412)。上記で説明したように、ユーザがインターフェースコントロールと対話する場合、ユーザは、ユーザ対話に基づいて、特定のスマートデバイスを制御することが可能である。たとえば、音楽アプリケーションが、スマートスピーカー上で稼働中である場合、ユーザは、モバイルデバイスのディスプレイ上でユーザに提示されたインターフェースコントロールと対話することによって、アプリケーションの制御(たとえば、音量レベル、次のトラックにスキップなど)を行うことができる。

#### 【0101】

モバイルデバイスは、たとえば、ワイヤレスネットワークを介して、スマートデバイスに制御データを送ることによって、スマートデバイスを直接的に制御することができる。スマートデバイスが、仮想アシスタント制御デバイスによって制御される場合、モバイルデバイスは、仮想アシスタント制御デバイスに制御データを送ることができる。次に、仮想アシスタント制御デバイスは、スマートデバイスを制御することができる。

#### 【0102】

いくつかの態様では、スマートデバイスを制御するための1つまたは複数のユーザインターフェースコントロールを提示することは、モバイルデバイスのカメラのビューファインダーの視野上に、拡張現実において1つまたは複数のユーザインターフェースコントロールを重ね合わせることを含む。たとえば、図2に示されるように、音楽制御ボタン228および音量トグルバー226は、ユーザが音声アシスタントデバイス215の音楽再生アプリケーションを制御するために選択することができる、モバイルデバイスのカメラのビューファインダー上に拡張現実において示されたユーザインターフェースコントロールである。

#### 【0103】

いくつかの態様では、連続したビデオ(たとえば、ライブ映像)の場合、ユーザインターフェースコントロールは、スマートデバイスが画像データ内で識別される間のみ、ユーザインターフェース上に重ね合わせられる。いくつかの態様では、モバイルデバイスがスマートデバイスを制御するために、スマートデバイスの画像データを連続して獲得する必要がなくなるように、スマートデバイスの静止フレームショットが取り込まれ、スマートデバイスを制御するために使用され得る。たとえば、ユーザは、音声アシスタントデバイスのスナップショットを撮り、次いで、音声アシスタントデバイスから離れるように移動し、仮想アシスタントアプリケーションを稼働させて、音声アシスタントデバイスを制御するためのユーザインターフェースコントロールを生成することを望むことがある。

#### 【0104】

図5は、上記で説明した動作を実施するために使用され得る、例示的なコンピュータシステム500のブロック図である。システム500は、プロセッサ510と、メモリ520と、記憶デバイス530と、入出力デバイス540とを含む。構成要素510、520、530、および540の各々は、たとえば、システムバス550を使用して相互接続され得る。プロセッサ510は、システム500内で実行するための命令を処理することが可能である。一実装形態では、プロセッサ510は、シングルスレッドプロセッサである。別の実装形態では、プロセッサ510は、マルチスレッドプロセッサである。プロセッサ510は、メモリ520内または記憶デバイス530上

10

20

30

40

50

で記憶された命令を処理することが可能である。

【0105】

メモリ520は、システム500内の情報を記憶する。一実装形態では、メモリ520は、コンピュータ可読媒体である。一実装形態では、メモリ520は、揮発性メモリユニットである。別の実装形態では、メモリ520は、不揮発性メモリユニットである。

【0106】

記憶デバイス530は、システム500のための大容量記憶を提供することが可能である。一実装形態では、記憶デバイス530は、コンピュータ可読媒体である。様々な異なる実装形態では、記憶デバイス530は、たとえば、ハードディスクデバイス、光ディスクデバイス、複数のコンピューティングデバイスによってネットワークを介して共有される記憶デバイス(たとえば、クラウド記憶デバイス)、または何らかの他の大容量記憶デバイスを含み得る。

10

【0107】

入出力デバイス540は、システム500のための入出力動作を提供する。一実装形態では、入出力デバイス540は、ネットワークインターフェースデバイス、たとえば、イーサネットカード、シリアル通信デバイス、たとえば、RS-232ポート、および/またはワイヤレスインターフェースデバイス、たとえば、802.11カードのうちの1つまたは複数を含み得る。別の実装形態では、入出力デバイスは、入力データを受信し、他の入出力デバイス、たとえば、キーボード、プリンタ、およびディスプレイデバイス560に出力データを送るように構成された、ドライバデバイスを含み得る。しかしながら、モバイルコンピューティングデバイス、モバイル通信デバイス、セットトップボックステレビクライアントデバイスなど、他の実装形態もまた使用され得る。

20

【0108】

例示的な処理システムについて、図5において説明したが、本明細書で説明する主題および機能的動作の実装形態は、本明細書で開示する構造およびそれらの構造的等価物を含む、他のタイプのデジタル電子回路において、またはコンピュータソフトウェア、ファームウェア、もしくはハードウェアにおいて、あるいはそれらのうちの1つまたは複数の組合せにおいて実装され得る。

【0109】

本明細書で説明する主題および機能的動作の実装形態は、本明細書で開示する構造およびそれらの構造的等価物を含む、デジタル電子回路において、有形に具現化されたコンピュータソフトウェアまたはファームウェアにおいて、コンピュータハードウェアにおいて、あるいはそれらのうちの1つまたは複数の組合せにおいて実装され得る。本明細書で説明する主題の実装形態は、1つまたは複数のコンピュータプログラム、すなわち、データ処理装置による実行のために、またはデータ処理装置の動作を制御するために、有形の非一時的プログラムキャリア上に符号化されたコンピュータプログラム命令の1つまたは複数のモジュールとして実装され得る。代替的にまたは追加として、プログラム命令は、データ処理装置による実行のために、好適な受信機装置への送信のために情報を符号化するために生成される、人工的に生成された伝搬信号、たとえば、マシンにより生成された電気信号、光信号、または電磁信号上で符号化され得る。コンピュータ記憶媒体は、マシン可読記憶デバイス、マシン可読記憶基板、ランダムもしくは順次アクセスメモリデバイス、またはそれらのうちの1つもしくは複数の組合せであり得る。

30

40

【0110】

「データ処理装置」という用語は、例として、プログラマブルプロセッサ、コンピュータ、または複数のプロセッサもしくはコンピュータを含む、データを処理するためのすべての種類の装置、デバイス、およびマシンを包含する。装置は、専用論理回路、たとえば、FPGA(フィールドプログラマブルゲートアレイ)、またはASIC(特定用途向け集積回路)を含み得る。装置はまた、ハードウェアに加えて、問題のコンピュータプログラムのための実行環境を作成するコード、たとえば、プロセッサファームウェア、プロトコルスタック、データベース管理システム、オペレーティングシステム、またはそれらのうちの1つも

50

しくは複数の組合せを構成するコードも含み得る。

#### 【0111】

コンピュータプログラム(プログラム、ソフトウェア、ソフトウェアアプリケーション、モジュール、ソフトウェアモジュール、スクリプト、またはコードと呼ばれるか、またはそのように記述されることもある)は、コンパイラ型言語もしくはインタープリタ型言語、または宣言型言語もしくは手続き型言語を含む、任意の形式のプログラミング言語において記述され得、スタンドアロンプログラムとして、またはモジュール、構成要素、サブルーチン、もしくはコンピューティング環境において使用するのに好適な他のユニットとしてを含む、任意の形態において展開され得る。コンピュータプログラムは、ファイルシステム内のファイルに対応し得るが、そうである必要はない。プログラムは、他のプログラムもしくはデータを保持するファイルの一部、たとえば、マークアップ言語ドキュメント中に記憶された1つもしくは複数のスクリプトにおいて、問題のプログラム専用の単一のファイルにおいて、または複数の協調されたファイル、たとえば、1つもしくは複数のモジュール、サブプログラム、もしくはコードの部分を記憶するファイルにおいて記憶され得る。コンピュータプログラムは、1つのコンピュータ上、または、1つのサイトに位置するか、もしくは複数のサイトにわたって分散され、通信ネットワークによって相互接続される、複数のコンピュータ上で実行されるように展開され得る。

10

#### 【0112】

本明細書で説明するプロセスおよび論理フローは、入力データにおいて動作すること、および出力を生成することによって、機能を実施するために、1つまたは複数のプログラマブルコンピュータが1つまたは複数のコンピュータプログラムを実行することによって実施され得る。プロセスおよび論理フローはまた、専用論理回路、たとえば、FPGA(フィールドプログラマブルゲートアレイ)、ASIC(特定用途向け集積回路)、またはGPGPU(汎用グラフィックス処理ユニット)によって実施され得、装置もまた専用論理回路として実装され得る。

20

#### 【0113】

コンピュータプログラムの実行に好適なコンピュータは、汎用もしくは専用のマイクロプロセッサ、または両方、あるいは任意の他の種類の中央処理ユニットを含み、例として、それらに基づき得る。一般に、中央処理ユニットは、読取り専用メモリまたはランダムアクセスメモリまたは両方から、命令およびデータを受信することになる。コンピュータの本質的な要素は、命令を実施または実行するための中央処理ユニットと、命令およびデータを記憶するための1つまたは複数のメモリデバイスとである。一般に、コンピュータはまた、データを記憶するための1つまたは複数の大容量記憶デバイス、たとえば、磁気ディスク、光磁気ディスク、または光ディスクを含むか、あるいはそれからデータを受信するため、またはそれにデータを転送するため、またはその両方のために動作可能に結合されることになる。ただし、コンピュータは、そのようなデバイスを有する必要はない。その上、コンピュータは、別のデバイス、たとえば、ほんのいくつかの例を挙げれば、モバイル電話、携帯情報端末(PDA)、モバイルオーディオもしくはビデオプレーヤ、ゲームコンソール、全地球測位システム(GPS)受信機、またはポータブル記憶デバイス、たとえば、ユニバーサルシリアルバス(USB)フラッシュドライブに埋め込まれ得る。

30

40

#### 【0114】

コンピュータプログラム命令およびデータを記憶するのに好適なコンピュータ可読媒体には、例として、半導体メモリデバイス、たとえば、EPROM、EEPROM、およびフラッシュメモリデバイス、磁気ディスク、たとえば、内蔵ハードディスクまたはリムーバブルディスク、光磁気ディスク、ならびにCD-ROMおよびDVD-ROMディスクを含む、すべての形態の不揮発性メモリ、媒体、およびメモリデバイスが含まれる。プロセッサおよびメモリは、専用論理回路によって補助されるか、または専用論理回路に組み込まれ得る。

#### 【0115】

ユーザとの対話を提供するために、本明細書で説明する主題の実施形態は、ユーザに情報を表示するためのディスプレイデバイス、たとえば、CRT(陰極線管)またはLCD(液晶デ

50

イスプレイ)モニタと、それによってユーザがコンピュータに入力を与えることができるキーボードおよびポインティングデバイス、たとえば、マウスまたはトラックボールとを有する、コンピュータ上で実装され得る。他の種類のデバイスが、ユーザとの対話を提供するために同様に使用され得、たとえば、ユーザに与えられるフィードバックは、任意の形態の感覚フィードバック、たとえば、視覚フィードバック、聴覚フィードバック、または触覚フィードバックであり得、ユーザからの入力、音響入力、スピーチ入力、または触覚入力を含む、任意の形態で受信され得る。追加として、コンピュータは、ユーザによって使用されるデバイスにドキュメントを送ること、およびそのデバイスからドキュメントを受信することによって、たとえば、ユーザのクライアントデバイス上のウェブブラウザから受信された要求に応答して、ウェブブラウザにウェブページを送ることによって、ユーザと対話することができる。

10

#### 【0116】

本明細書で説明する主題の実施形態は、バックエンド構成要素を、たとえば、データサーバとして含むか、あるいは、ミドルウェア構成要素、たとえば、アプリケーションサーバを含むか、あるいは、フロントエンド構成要素、たとえば、それを通してユーザが、本明細書で説明する主題の一実装形態と対話することができる、グラフィカルユーザインターフェースもしくはウェブブラウザを有するクライアントコンピュータ、または、1つもしくは複数のそのようなバックエンド構成要素、ミドルウェア構成要素、もしくはフロントエンド構成要素の任意の組合せを含む、コンピューティングシステムにおいて実装され得る。システムの構成要素は、任意の形態または任意の媒体のデジタルデータ通信、たとえば、通信ネットワークによって相互接続され得る。通信ネットワークの例には、ローカルエリアネットワーク(「LAN」)、およびワイドエリアネットワーク(「WAN」)、たとえば、インターネットが含まれる。

20

#### 【0117】

コンピューティングシステムは、クライアントおよびサーバを含み得る。クライアントおよびサーバは、一般に互いからリモートであり、典型的には、通信ネットワークを通して対話する。クライアントおよびサーバの関係は、それぞれのコンピュータ上で稼働しており、互いにクライアントサーバ関係を有する、コンピュータプログラムによって生じる。

#### 【0118】

本明細書は、多数の特定の実装詳細を含むが、これらは、いずれかの発明の、または請求され得るものの範囲の限定として解釈されるべきではなく、むしろ、特定の発明の特定の実施形態に固有であり得る特徴の説明として解釈されるべきである。別個の実施形態との関連で本明細書で説明するいくつかの特徴もまた、単一の実施形態において組み合わせて実装され得る。逆に、単一の実施形態との関連で説明する様々な特徴もまた、複数の実施形態において別個に、または任意の好適な部分組合せにおいて実装され得る。その上、特徴は、いくつかの組合せにおいて作用するとして上記で説明されることがあり、最初にそのようなものとして請求されることさえあるが、請求される組合せからの1つまたは複数の特徴は、場合によっては、その組合せから削除され得、請求される組合せは、部分組合せ、または部分組合せの変形形態を対象とし得る。

30

40

#### 【0119】

同様に、動作は、特定の順序で図面において図示されるが、これは、望ましい結果を達成するために、そのような動作が図示された特定の順序で、もしくは順番に実施されること、またはすべての図示された動作が実施されることを必要とするものとして理解されるべきではない。いくつかの状況では、マルチタスキングおよび並列処理が有利であり得る。その上、上記で説明した実施形態における様々なシステムモジュールおよび構成要素の分離は、すべての実施形態においてそのような分離を必要とするものとして理解されるべきではなく、説明したプログラム構成要素およびシステムは、一般に単一のソフトウェア製品とともに統合され得るか、または複数のソフトウェア製品にパッケージ化され得ることを理解されたい。

50

## 【0120】

主題の特定の実施形態について説明した。他の実施形態は、以下の特許請求の範囲の範囲内である。たとえば、特許請求の範囲に記載されているアクションは、異なる順序で実施され、なお、望ましい結果を達成することが可能である。一例として、添付の図面に図示されたプロセスは、望ましい結果を達成するために、必ずしも図示された特定の順序、または順番を必要とするとは限らない。いくつかの実装形態では、マルチタスキングおよび並列処理が有利であり得る。

## 【符号の説明】

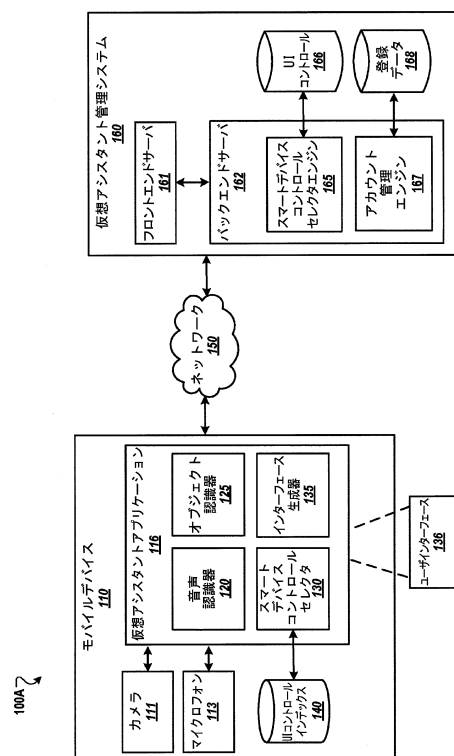
## 【0121】

|                     |                                            |    |
|---------------------|--------------------------------------------|----|
| 100A                | 例示的な環境                                     | 10 |
| 110                 | モバイルデバイス                                   |    |
| 111                 | カメラ                                        |    |
| 113                 | マイクロフォン                                    |    |
| 115                 | オブジェクト                                     |    |
| 116                 | 仮想アシスタントアプリケーション                           |    |
| 119                 | インターフェースコントロール                             |    |
| 120                 | 音声認識器                                      |    |
| 123                 | 画像データ                                      |    |
| 124                 | ロケーションデータ                                  |    |
| 125                 | オブジェクト認識器                                  | 20 |
| 126                 | 識別されたオブジェクトデータ                             |    |
| 130                 | スマートデバイスコントロールセクタ                          |    |
| 131                 | インターフェースコントロール識別データ                        |    |
| 135                 | ユーザインターフェース生成器、インターフェース生成器                 |    |
| 136、212、222、322、332 | ユーザインターフェース                                |    |
| 138                 | 提示データ                                      |    |
| 140                 | ユーザインターフェースコントロールインデックス、コントロールインデックス       |    |
| 150                 | データ通信ネットワーク、ネットワーク                         |    |
| 160                 | 仮想アシスタント管理システム                             | 30 |
| 161                 | フロントエンドサーバ                                 |    |
| 162                 | バックエンドサーバ                                  |    |
| 165                 | スマートデバイスコントロール選択エンジン、スマートデバイスコントロールセクタエンジン |    |
| 166                 | スマートデバイスコントロール記憶ユニット                       |    |
| 167                 | アカウント管理エンジン                                |    |
| 168                 | 登録データ                                      |    |
| 210、310             | スクリーンショット、第1のスクリーンショット                     |    |
| 215                 | 音声アシスタントデバイス                               |    |
| 220、320             | スクリーンショット、第2のスクリーンショット                     | 40 |
| 224                 | ユーザインターフェースコントロール                          |    |
| 226                 | 音量トグルバー                                    |    |
| 228                 | 音楽制御ボタン                                    |    |
| 312                 | インターフェース                                   |    |
| 315                 | 物理的コントロール                                  |    |
| 324                 | ユーザインターフェースコントロール、インターフェースコントロール           |    |
| 325                 | 「ON」トグルボタン                                 |    |
| 326                 | 調光トグルバー                                    |    |
| 330                 | スクリーンショット、第3のスクリーンショット                     |    |
| 334                 | 「OFF」トグルボタン                                | 50 |

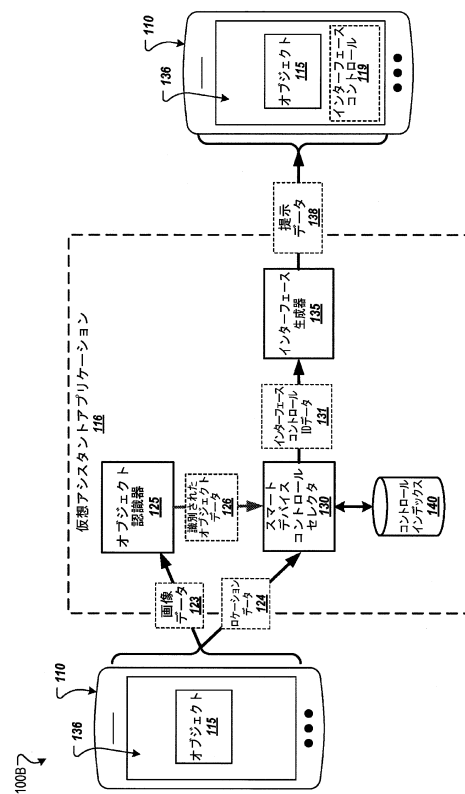


- 500 コンピュータシステム、システム  
 510 プロセッサ、構成要素  
 520 メモリ、構成要素  
 530 記憶デバイス、構成要素  
 540 入出力デバイス、構成要素  
 550 システムバス  
 560 他の入出力デバイス、たとえば、キーボード、プリンタ、およびディスプレイ  
 デバイス

【図 1 A】



【図 1 B】



【図 2】

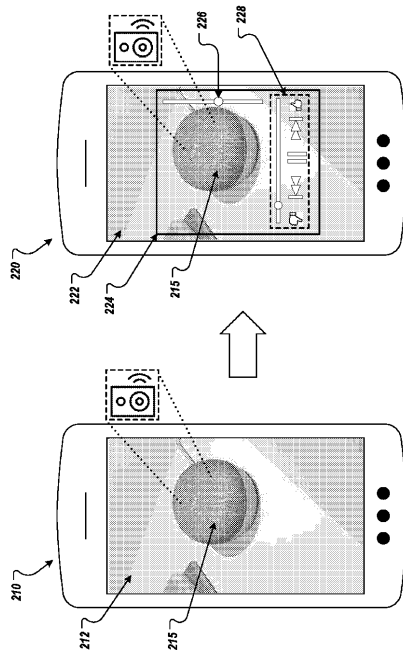


FIG. 2

【図 3】

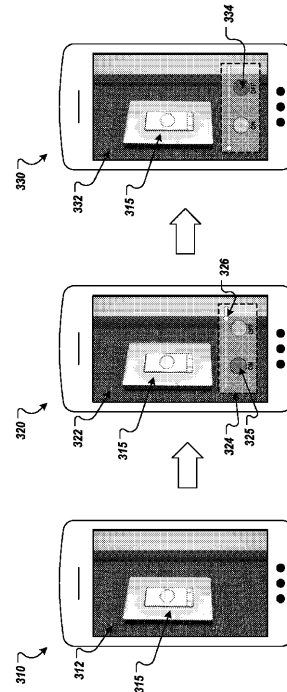
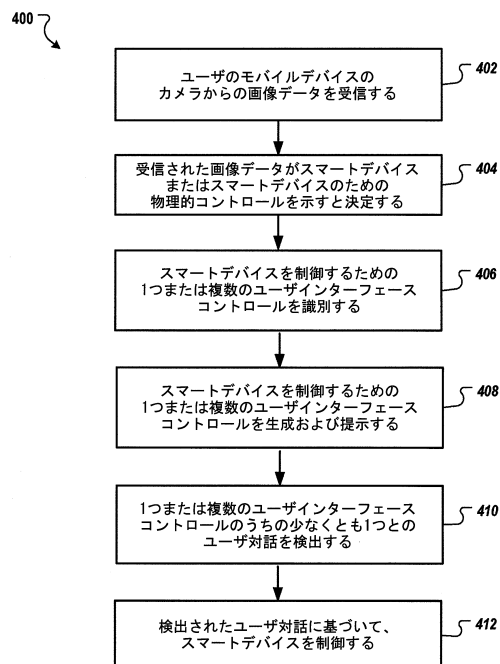
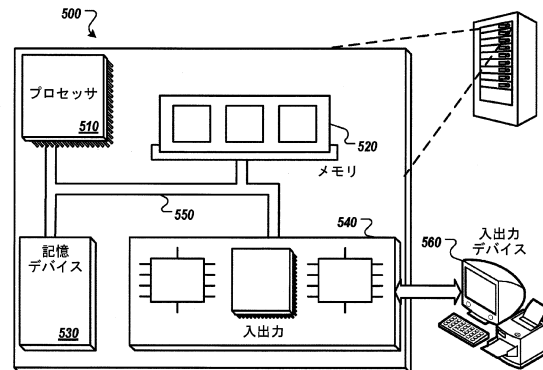


FIG. 3

【図 4】



【図 5】



---

フロントページの続き

- (72)発明者 イブラヒム・バドル  
スイス・８００２・チューリッヒ・ブラントシェンケシュトラーセ・１１０
- (72)発明者 ゴクハン・エイチ・バキル  
スイス・８００２・チューリッヒ・ブラントシェンケシュトラーセ・１１０
- (72)発明者 ローラント・ペーター・ケール  
スイス・８００２・チューリッヒ・ブラントシェンケシュトラーセ・１１０
- (72)発明者 ニルス・グリムスモ  
スイス・８００２・チューリッヒ・ブラントシェンケシュトラーセ・１１０

審査官 鈴木 順三

- (56)参考文献 特表２０１７－５３５９７７（ＪＰ，Ａ）  
米国特許出願公開第２０１８／００６３５６９（ＵＳ，Ａ１）  
特開２０１５－０３８７４０（ＪＰ，Ａ）

- (58)調査した分野(Int.Cl., DB名)
- |         |           |   |             |
|---------|-----------|---|-------------|
| G 0 6 F | 3 / 0 1   |   |             |
| G 0 6 F | 3 / 0 4 8 | — | 3 / 0 4 8 9 |
| H 0 4 N | 5 / 7 6   | — | 5 / 9 5 6   |
| H 0 4 N | 2 1 / 0 0 | — | 2 1 / 8 5 8 |