

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7314238号  
(P7314238)

(45)発行日 令和5年7月25日(2023.7.25)

(24)登録日 令和5年7月14日(2023.7.14)

(51)国際特許分類		F I	
G 0 6 F	3/16 (2006.01)	G 0 6 F	3/16 6 5 0
H 0 4 M	3/42 (2006.01)	H 0 4 M	3/42 P
G 1 0 L	17/00 (2013.01)	G 1 0 L	17/00 2 0 0 C
G 1 0 L	15/00 (2013.01)	G 1 0 L	15/00 2 0 0 A

請求項の数 17 外国語出願 (全41頁)

(21)出願番号	特願2021-179423(P2021-179423)	(73)特許権者	502208397
(22)出願日	令和3年11月2日(2021.11.2)		グーグル エルエルシー
(62)分割の表示	特願2019-545937(P2019-545937)		Google LLC
	)の分割		アメリカ合衆国 カリフォルニア州 9 4
原出願日	平成30年5月16日(2018.5.16)		0 4 3 マウンテン ビュー アンフィシ
(65)公開番号	特開2022-19745(P2022-19745A)		アター パークウェイ 1 6 0 0
(43)公開日	令和4年1月27日(2022.1.27)		1 6 0 0 Amphitheatre P
審査請求日	令和3年12月1日(2021.12.1)		arkway 9 4 0 4 3 Mounta
(31)優先権主張番号	62/506,805		in View, CA U.S.A.
(32)優先日	平成29年5月16日(2017.5.16)	(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
		(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】 共有された音声対応デバイスにおける電話の処理

(57)【特許請求の範囲】

【請求項1】

1つまたは複数のプロセッサによって実装される方法であって、

デジタルアシスタントデバイスに関連付けられた電話元ユーザから発話を受け取るステップであって、前記発話が、前記デジタルアシスタントデバイスが、前記発話によってまた識別された電話先ユーザへの音声電話を開始することを要求する、ステップと、

少なくとも前記発話に基づいて、複数の異なるユーザから前記電話元ユーザを特定するステップと、

前記電話元ユーザを特定することに基づいて、前記電話元ユーザの異なるデバイスに関連付けられたボイス番号が前記電話元ユーザに関連付けられた前記デジタルアシスタントデバイスに利用可能であると判定するステップと、

前記発話によって特定された前記電話先ユーザに基づいて電話すべき受話者ボイス番号を決定するステップと、

前記デジタルアシスタントデバイスに、前記受話者ボイス番号に前記音声電話を開始させるステップであって、前記受話者ボイス番号への前記音声電話が、前記デジタルアシスタントデバイスに関連付けられた前記ボイス番号の代わりに前記音声電話の元の番号として、前記異なるデバイスに関連付けられた前記ボイス番号を示す、ステップと、

前記デジタルアシスタントデバイスが追加の音声電話を開始することを要求する発話を受信するステップと、

前記複数の異なるユーザのいずれによっても話されていないとして追加の発話を分類する

ステップと、

前記複数の異なるユーザのいずれによっても話されていないとして第2の発話を分類したことに応答して、前記デジタルアシスタントデバイスに、前記複数の異なるユーザのいずれの連絡先エントリにアクセスすることなく前記追加の音声電話を開始させるステップとを備える、方法。

【請求項2】

前記発話によって特定された前記電話先ユーザに基づいて電話すべき受話者ボイス番号を決定するステップが、

少なくとも前記発話に基づいて、複数の連絡先エントリの中から特定の連絡先エントリを特定するステップと、

前記受話者ボイス番号として、前記特定の連絡先エントリに関連付けられた特定のボイス番号を決定するステップと

を備える、請求項1に記載の方法。

【請求項3】

前記複数の連絡先エントリの中から前記特定の連絡先エントリを特定するステップが、前記発話の文字起こしを生成するステップと、

前記特定の連絡先エントリに関連付けられた名前を含む前記文字起こしを決定するステップと

を備える、請求項2に記載の方法。

【請求項4】

少なくとも前記発話に基づいて前記複数の異なるユーザから前記電話元ユーザを特定するステップが、

アシスタントサーバから、前記発話に含まれる語りが前記電話元ユーザに対応する語りと一致すると判定されたことの指示を取得するステップ

を備える、請求項1に記載の方法。

【請求項5】

少なくとも前記発話に基づいて前記複数の異なるユーザから前記電話元ユーザを特定するステップが、

前記アシスタントサーバによって、前記発話に含まれる語りが前記電話元ユーザに一致すると判定するステップ

を備える、請求項4に記載の方法。

【請求項6】

前記デジタルアシスタントデバイスに、前記受話者ボイス番号に前記音声電話を開始させるステップが、

前記デジタルアシスタントデバイスへ、前記受話者ボイス番号および前記受話者ボイス番号に前記音声電話を開始する命令を提供するステップ

を備える、請求項1に記載の方法。

【請求項7】

システムであって、

少なくとも1つのプロセッサと、

命令を備えたメモリであって、

1つまたは複数のプロセッサによって実装される方法であって、実行されたとき、前記少なくとも1つのプロセッサに、

デジタルアシスタントデバイスに関連付けられた電話元ユーザから発話を受け取ることであって、前記発話が、前記デジタルアシスタントデバイスが、前記発話によってまた識別された電話先ユーザへの音声電話を開始することを要求する、受け取ることと、

少なくとも前記発話に基づいて、複数の異なるユーザから前記電話元ユーザを特定することと、

前記電話元ユーザを特定することに基づいて、前記電話元ユーザの異なるデバイスに関連付けられたボイス番号が前記電話元ユーザに関連付けられた前記デジタルアシスタ

10

20

30

40

50

ントデバイスに利用可能であると判定することと、

前記発話によって特定された前記電話先ユーザに基づいて電話すべき受話者ボイス番号を決定することと、

前記デジタルアシスタントデバイスに、前記受話者ボイス番号に前記音声電話を開始させることであって、前記受話者ボイス番号への前記音声電話が、前記デジタルアシスタントデバイスに関連付けられた前記ボイス番号の代わりに前記音声電話の元の番号として、前記異なるデバイスに関連付けられた前記ボイス番号を示す、開始させることと、前記デジタルアシスタントデバイスが追加の音声電話を開始することを要求する発話を受信することと、

前記複数の異なるユーザのいずれによっても話されていないとして追加の発話を分類することと、

10

前記複数の異なるユーザのいずれによっても話されていないとして第2の発話を分類したことに応答して、前記デジタルアシスタントデバイスに、前記複数の異なるユーザのいずれの連絡先エントリにアクセスすることなく前記追加の音声電話を開始させることと

を行わせる命令を備える、メモリと

を備える、システム。

#### 【請求項 8】

前記発話によって特定された前記電話先ユーザに基づいて電話すべき受話者ボイス番号を決定するための命令が、

少なくとも前記発話に基づいて、複数の連絡先エントリの中から特定の連絡先エントリを特定することと、

20

前記受話者ボイス番号として、前記特定の連絡先エントリに関連付けられた特定のボイス番号を決定することと

を行わせるための命令を備える、請求項7に記載のシステム。

#### 【請求項 9】

前記複数の連絡先エントリの中から前記特定の連絡先エントリを特定するための命令が、前記発話の文字起こしを生成することと、

前記特定の連絡先エントリに関連付けられた名前を含む前記文字起こしを決定することとを行わせるための命令を備える、請求項8に記載のシステム。

#### 【請求項 10】

30

少なくとも前記発話に基づいて前記複数の異なるユーザから前記電話元ユーザを特定するための命令が、

アシスタントサーバから、前記発話に含まれる語りが前記電話元ユーザに対応する語りと一致すると判定されたことの指示を取得すること

を行わせるための命令を備える、請求項7に記載のシステム。

#### 【請求項 11】

少なくとも前記発話に基づいて前記複数の異なるユーザから前記電話元ユーザを特定するための命令が、

前記アシスタントサーバによって、前記発話に含まれる語りが前記電話元ユーザに一致すると判定すること

40

を行わせるための命令を備える、請求項10に記載のシステム。

#### 【請求項 12】

前記デジタルアシスタントデバイスに、前記受話者ボイス番号に前記音声電話を開始させるための命令が、

前記デジタルアシスタントデバイスへ、前記受話者ボイス番号および前記受話者ボイス番号に前記音声電話を開始する命令を提供すること

を行わせるための命令を備える、請求項7に記載のシステム。

#### 【請求項 13】

命令を記憶した非一時的コンピュータ可読記憶媒体であって、実行されたとき、少なくとも1つのプロセッサに、

50

デジタルアシスタントデバイスに関連付けられた電話元ユーザから発話を受け取ること  
 であって、前記発話が、前記デジタルアシスタントデバイスが、前記発話によってまた識  
 別された電話先ユーザへの音声電話を開始することを要求する、受け取ることと、

少なくとも前記発話に基づいて、複数の異なるユーザから前記電話元ユーザを特定する  
 ことと、

前記電話元ユーザを特定することに基づいて、前記電話元ユーザの異なるデバイスに関  
 連付けられたボイス番号が前記電話元ユーザに関連付けられた前記デジタルアシスタント  
 デバイスに利用可能であると判定することと、

前記発話によって特定された前記電話先ユーザに基づいて電話すべき受話者ボイス番  
 号を決定することと、

前記デジタルアシスタントデバイスに、前記受話者ボイス番号に前記音声電話を開始  
 させることであって、前記受話者ボイス番号への前記音声電話が、前記デジタルアシスタ  
 ントデバイスに関連付けられた前記ボイス番号の代わりに前記音声電話の元の番号として  
 、前記異なるデバイスに関連付けられた前記ボイス番号を示す、開始させることと、

前記デジタルアシスタントデバイスが追加の音声電話を開始することを要求する発話を受  
 信することと、

前記複数の異なるユーザのいずれによっても話されていないとして追加の発話を分類する  
 ことと、

前記複数の異なるユーザのいずれによっても話されていないとして第2の発話を分類した  
 ことに応答して、前記デジタルアシスタントデバイスに、前記複数の異なるユーザのい  
 ずれの連絡先エントリにアクセスすることなく前記追加の音声電話を開始させることと

を備える動作を実施させる、非一時的コンピュータ可読記憶媒体。

#### 【請求項 14】

前記発話によって特定された前記電話先ユーザに基づいて電話すべき受話者ボイス番号  
 を決定することが、

少なくとも前記発話に基づいて、複数の連絡先エントリの中から特定の連絡先エントリ  
 を特定することと、

前記受話者ボイス番号として、前記特定の連絡先エントリに関連付けられた特定のボイ  
 ス番号を決定することと

を備える、請求項13に記載の非一時的コンピュータ可読記憶媒体。

#### 【請求項 15】

前記複数の連絡先エントリの中から前記特定の連絡先エントリを特定することが、

前記発話の文字起こしを生成することと、

前記特定の連絡先エントリに関連付けられた名前を含む前記文字起こしを決定することと

を備える、請求項14に記載の非一時的コンピュータ可読記憶媒体。

#### 【請求項 16】

少なくとも前記発話に基づいて前記複数の異なるユーザから前記電話元ユーザを特定す  
 ることが、

アシスタントサーバによって、前記発話に含まれる語りが前記電話元ユーザに一致する  
 と判定すること

を備える、請求項13に記載の非一時的コンピュータ可読記憶媒体。

#### 【請求項 17】

前記デジタルアシスタントデバイスに、前記受話者ボイス番号に前記音声電話を開始さ  
 せることが、

前記デジタルアシスタントデバイスへ、前記受話者ボイス番号および前記受話者ボイス  
 番号に前記音声電話を開始する命令を提供すること

を備える、請求項13に記載の非一時的コンピュータ可読記憶媒体。

#### 【発明の詳細な説明】

#### 【技術分野】

【0001】

10

20

30

40

50

## 関連出願の相互参照

本出願は、2017年5月16日に出願した、「HANDLING PERSONAL TELEPHONE CALLS USING VOICE CONTROL」と題した米国特許仮出願第62/506,805号の利益を主張するものであり、この米国特許仮出願は、参照によりその全体が本明細書に組み込まれる。

### 【0002】

概して、本明細書は、自然言語処理に関する。

### 【背景技術】

### 【0003】

音声対応デバイスは、ユーザからの発話に応答してアクションを実行する可能性がある。たとえば、ユーザが、「オーケー、コンピュータ。今日は雨が降る?」という可能性があり、音声対応デバイスが、「一日中晴れます」と聞こえるように応答する可能性がある。音声対応デバイスを使用することの恩恵は、音声対応デバイスとインタラクションすることが概してハンズフリーであり得ることである。たとえば、ユーザが質問をするとき、音声対応デバイスは、ユーザがユーザの手を使用していかなる物と物理的にインタラクションすることも必要とせず、聞き取ることができる答えを提供する可能性がある。しかし、よくある音声対応デバイスは、サポートされるインタラクションの種類を限定される。

### 【発明の概要】

### 【課題を解決するための手段】

### 【0004】

音声対応デバイスが、音声電話を掛けるために使用される可能性がある。たとえば、ジョン・ドゥが、音声対応デバイスに電話番号(555) 555-5555への電話を掛けさせるために「オーケー、コンピュータ。(555) 555-5555に電話を掛けて」という可能性がある。通常、アウトバウンドコール(outbound call)は、発信者を特定するために使用され得る発信者番号に関連付けられる。たとえば、ジョン・ドゥが自分の電話を使用して(555) 555-5555に電話を掛けるとき、電話を受ける電話は、電話がジョン・ドゥの電話に関連する電話番号から来ていることを示す可能性がある。

### 【0005】

発信者番号を電話と関連付けることは、電話の受話者が電話に答えるべきかどうかを判断するために発信者番号を使用し、さらに、受話者が折り返し電話をはける必要がある場合に発信者番号を使用する可能性があるため有用である可能性がある。しかし、通常の電話とは異なり、一部の音声対応デバイスは、電話のための発信者番号として使用され得る電話番号に関連付けられない可能性がある。

### 【0006】

電話を掛けるときに発信者番号を提供するために、音声対応デバイスは、発話者のパーソナルボイス番号(personal voice number)を発信者番号として使用しようと試みる可能性がある。パーソナルボイス番号は、ユーザに電話を掛けるために使用される番号である可能性がある。たとえば、ジョンが「オーケー、コンピュータ。(555) 555-5555に電話を掛けて」と言うとき、音声対応デバイスは、ジョン・ドゥの電話番号(555) 999-9999を発信者番号として使用する可能性がある。音声対応デバイスが発話者のパーソナルボイス番号を決定することができない場合、音声対応デバイスは、その代わりに、電話が折り返し電話を掛けるために使用され得るボイス番号(voice number)に関連付けられないように非通知で電話を掛ける可能性がある。たとえば、そのような電話は、「不明な番号(Unknown Number)」または「非公開番号(Private Number)」を発信者番号として示す可能性がある。

### 【0007】

場合によっては、電話が緊急サービスへのものである場合、電話は、音声対応デバイスに折り返し電話を掛けるために受話者が使用することができる仮番号を使用して掛けられる可能性がある。たとえば、そのような電話は、音声対応デバイスに折り返し電話を掛けるためにその後2時間使用される可能性がある電話番号(555) 888-8888を示す可能性がある。

10

20

30

40

50

## 【0008】

追加的にまたは代替的に、音声対応デバイスは、電話を掛けるべきボイス番号を決定するために発話者のアイデンティティ(identity)を使用する可能性がある。たとえば、ジョンが「オーケー、コンピュータ。父さんに電話を掛けて」と言うとき、音声対応デバイスは、ジョンを認識するかまたはそうでなければ認証し、それから、「父さん」に関する電話番号を決定するためにジョンの連絡先のレコードにアクセスする可能性がある。別の例において、ジェーンが「オーケー、コンピュータ。父さんに電話を掛けて」と言う場合、音声対応デバイスは、音声認識またはその他の認証技術によってジェーンをジョンと区別し、その後、「父さん」に関する電話番号を決定するためにジェーンの連絡先のレコードにアクセスする可能性がある。さらに別の例において、ゲストが「オーケー、コンピュータ。父さんに電話を掛けて」と言うとき、音声対応デバイスは、音声(またはその他の認証技術)によってゲストを認識せず、「父さん」に関する電話番号を決定するためにいかなるユーザの連絡先のレコードにもアクセスしない可能性がある。したがって、これらの3つの例に見られるように、「オーケー、コンピュータ。父さんに電話を掛けて」は、発話者のアイデンティティに基づいて異なる結果を有する可能性がある。

10

## 【0009】

追加的にまたは代替的に、音声対応デバイスは、音声対応デバイスによって掛けられた音声電話中にユーザからの発話に回答する可能性がある。たとえば、電話中に、音声対応デバイスは、「オーケー、コンピュータ。電話を切って」、「オーケー、コンピュータ。スピーカーのボリュームを上げて」、「オーケー、コンピュータ。今日の天気はどう」というコマンドに回答する可能性がある。音声電話中の発話に回答して、音声対応デバイスは、受話者からの発話の少なくとも一部をブロックする可能性がある。たとえば、ユーザが「オーケー、コンピュータ。スピーカーのボリュームを上げて」と言うとき、音声対応デバイスは、スピーカーのボリュームを上げ、受話者が「オーケー、コンピュータ」のみを聞くように「スピーカーのボリュームを上げて」をブロックする可能性がある。別の例において、音声対応デバイスは、受話者に音声を提供する際にレイテンシーがある可能性があり、したがって、発話が「オーケー、コンピュータ」で始まるときに発話全体が受話者によって聞かれることをブロックする可能性がある。

20

## 【0010】

したがって、一部の実装において、利点は、複数のユーザによって共有された音声対応デバイスが、ユーザが電話を掛けることを依然として可能にし、ユーザのモバイルコンピューティングデバイスのボイス番号である、受話者の電話上で発呼番号(calling number)に見える番号を有する可能性があることであり得る。概して人々は認識されない番号からの電話を取らない可能性があるため、これは、音声対応デバイスを使用して掛けられた電話が答えられる見込みを高める可能性がある。加えて、電話は、電話を掛けられている人がユーザに関連するボイス番号の使用に基づいて誰が掛けてきているのかを既に知っている可能性があるとき、より効率的である可能性がある。同時に、音声対応デバイスが発話者の語りに一致するボイス番号を使用するので、ユーザが音声対応デバイスのいかなるその他のユーザのボイス番号も使用しない可能性があるという点で、セキュリティが提供される可能性がある。

30

40

## 【0011】

一部の実装における別の利点は、ユーザがボイス番号の数字を言う代わりに連絡先の名前を迅速に言うことができる可能性があるため、音声対応デバイス上の連絡先の使用を可能にすることが、ユーザが電話をより迅速に掛けることを可能にする可能性があることであり得る。音声対応デバイスは、複数のユーザの間で連絡先の曖昧さを取り除くこともできる可能性がある。たとえば、異なるユーザは、異なる電話番号に関連付けられている同じ名前「母さん」を有するそれぞれの連絡先のエントリーを有する可能性がある。使用される連絡先が発話者の語りに一致する連絡先であることを音声対応デバイスが保証する可能性があるため、ユーザが音声対応デバイスのその他のユーザの連絡先を使用しない可能性があるという点で、セキュリティがやはり提供される可能性がある。

50

## 【 0 0 1 2 】

一部の実装におけるさらに別の利点は、音声電話中の問い合わせの処理を可能にすることが、電話に関してより優れたハンズフリー体験を可能にする可能性があることであり得る。たとえば、ユーザは、発信者が特定の番号の押下によって応答するように要求する自動音声応答機能(automated attendant)に応答して仮想的に数字を押すことができる可能性がある。問い合わせが処理されている間は両方向とも保留にし、問い合わせが解決されると両方向の保留を自動的に終了させるという点で、セキュリティがやはり提供される可能性がある。加えて、両方向の保留は、音声対応仮想アシスタントからの問い合わせに対する応答が相手からの音声によって不明瞭にされないことを保証する可能性がある。たとえば、両方向とも保留にしないと、相手が、音声対応仮想アシスタントからの応答が出力されるのと同時にしゃべる可能性がある。

10

## 【 0 0 1 3 】

一部の態様において、本明細書に記載の対象は、音声電話を要求する発話を受け取るアクションと、発話を特定の知られているユーザによって言われたものとして分類するアクションと、特定の知られているユーザがパーソナルボイス番号に関連付けられているかどうかを判定するアクションと、特定の知られているユーザがパーソナルボイス番号に関連付けられているとの判定に応じて、パーソナルボイス番号を用いて音声電話を開始するアクションとを含む可能性がある方法に具現化される可能性がある。

## 【 0 0 1 4 】

一部の実装において、発話を特定の知られているユーザによって言われたものとして分類するアクションは、発話の中の語りが特定の知られているユーザに対応する語りと一致するかどうかを判定することを含む。特定の实装において、発話を特定の知られているユーザによって言われたものとして分類するアクションは、発話者の少なくとも一部の視覚的画像が特定の知られているユーザに対応する視覚的情報と一致するかどうかを判定することを含む。一部の实装において、特定の知られているユーザがパーソナルボイス番号に関連付けられているかどうかを判定するアクションは、特定の知られているユーザのアカウント情報にアクセスすることと、ユーザのアカウント情報が特定の知られているユーザに関するボイス番号を記憶するかどうかを判定することを含む。

20

## 【 0 0 1 5 】

特定の实装において、特定の知られているユーザがパーソナルボイス番号に関連付けられているかどうかを判定するアクションは、特定の知られているユーザの指示および発話の表現をサーバに提供することと、特定の知られているユーザのパーソナルボイス番号、電話を掛けるべきボイス番号、および音声電話を掛ける命令をサーバから受け取ることとを含む。一部の实装において、特定の知られているユーザがパーソナルボイス番号に関連付けられているかどうかを判定するアクションは、特定の知られているユーザのアカウントにアクセスすることと、ユーザのアカウントが電話を示すかどうかを判定することと、電話が音声対応デバイスに接続されていると判定することとを含む。

30

## 【 0 0 1 6 】

特定の实装において、パーソナルボイス番号を用いて音声電話を開始するアクションは、音声対応デバイスに接続された電話によって音声電話を開始することを含む。一部の实装においては、特定の知られているユーザがパーソナルボイス番号に関連付けられているとの判定に応じて、パーソナルボイス番号を用いて音声電話を開始するアクションは、ボイスオーバーインターネットプロトコル電話プロバイダ(Voice over Internet Protocol call provider)を通じて音声電話を開始することを含む。

40

## 【 0 0 1 7 】

一部の態様において、本明細書に記載の対象は、音声電話を要求する発話を受け取るアクションと、発話を特定の知られているユーザによって言われたものとして分類するアクションと、発話を特定の知られているユーザによって言われたものとする分類に応じて、特定の知られているユーザに関する連絡先に基づいて電話を掛けるべき受話者のボイス番号を決定するアクションと、受話者のボイス番号への音声電話を開始するアクションとを

50

含む可能性がある方法に具現化される可能性がある。

【0018】

一部の実装においては、発話を特定の知られているユーザによって言われたものとする分類に応じて、特定の知られているユーザによって作成された連絡先のエントリーを取得することは、発話を特定の知られているユーザによって言われたものとする分類に応じて、特定の知られているユーザの連絡先のエントリーが利用可能であると判定することと、特定の知られているユーザの連絡先のエントリーが利用可能であるとの判定に応じて、特定の知られているユーザによって作成された連絡先のエントリーを取得することを含む。特定の  
10 実装においては、発話を特定の知られているユーザによって言われたものとする分類に応じて、特定の知られているユーザに関する音声連絡先(voice contact)に基づいて電話を掛けるべき受話者のボイス番号を決定するアクションは、発話を特定の知られているユーザによって言われたものとする分類に応じて、特定の知られているユーザによって作成された連絡先のエントリーを取得することと、連絡先のエントリーの中から特定の連絡先のエントリーを特定することと、特定の連絡先のエントリーが、発話に一致する名前を含む、特定することと、特定の連絡先のエントリーによって示されたボイス番号を受話者のボイス番号として決定することを含む。

【0019】

一部の実装において、連絡先のエントリーの中から特定の連絡先のエントリーを特定することと、特定の連絡先のエントリーが、発話に一致する名前を含む、特定することは、  
20 発話の文字に起こされたものを生成することと、文字に起こされたものが名前を含むと判定することとを含む。特定の  
実装において、発話を特定の知られているユーザによって言われたものとして分類するアクションは、発話の中の語りが特定の知られているユーザに対応する語りに一致すると音声対応デバイスによって判定されたという指示を取得することを含む。一部の実装において、発話を特定の知られているユーザによって言われたものとして分類するアクションは、発話の中の語りが特定の知られているユーザに対応する語りと一致するかどうかを判定することを含む。特定の  
実装において、受話者のボイス番号への音声電話を開始するアクションは、受話者のボイス番号と、受話者のボイス番号への音声電話を開始する命令とを音声対応デバイスに与えることを含む。

【0020】

一部の実装において、アクションは、第2の音声電話を要求する第2の発話を受け取ることと、第2の発話を音声対応デバイスのいかなる知られているユーザによっても言われていないものとして分類することと、第2の発話を音声対応デバイスのいかなる知られているユーザによっても言われていないものとする分類に応じて、音声対応デバイスのいかなる知られているユーザに関する音声連絡先にもアクセスせずに第2の音声電話を開始することを含む。  
30

【0021】

一部の態様において、本明細書に記載の対象は、第一者(first party)が第一者と第二者(second party)との間の音声電話中に音声対応仮想アシスタントへの問い合わせを言ったと判定するアクションと、第一者が第一者と第二者との間の音声電話中に音声対応仮想アシスタントへの問い合わせを言ったとの判定に応じて、第一者と第二者との間の音声電話を保留にするアクションと、音声対応仮想アシスタントが問い合わせを解決したと判定するアクションと、音声対応仮想アシスタントが問い合わせを処理したとの判定に応じて、第一者と第二者との間の音声電話の保留を解除するアクションとを含む可能性がある方法に具現化される可能性がある。  
40

【0022】

一部の実装において、第一者が第一者と第二者との間の音声電話中に音声対応仮想アシスタントへの問い合わせを言ったと判定するアクションは、音声電話中に第一者によってホットワード(hotword)が言われたと音声対応デバイスによって判定することを含む。特定の  
50 実装において、第一者と第二者との間の音声電話を保留にするアクションは、音声電話を保留にする命令を音声コールプロバイダに与えることを含む。一部の実装において、



第一者と第二者との間の音声電話を保留にするアクションは、マイクロフォンからの音声をボイスサーバではなく音声対応仮想アシスタントにルーティングすることと、ボイスサーバからの音声中ではなく音声対応仮想アシスタントからの音声をスピーカーにルーティングすることを含む。

【0023】

特定の実装において、音声対応仮想アシスタントが問い合わせを解決したと判定するアクションは、問い合わせ、および音声電話が音声対応デバイス上で継続中であるという指示を音声対応仮想アシスタントに与えることと、問い合わせへの応答および問い合わせが解決されるという指示を音声対応仮想アシスタントから受け取ることとを含む。一部の実装において、問い合わせへの応答および問い合わせが解決されるという指示を音声対応仮想アシスタントから受け取るとは、問い合わせに対する応答として出力されるべき音声、および問い合わせが解決されるかどうかを示す値を有する2進フラグを受け取ることを含む。特定の実装において、音声対応仮想アシスタントは、問い合わせに対応するコマンドを特定し、コマンドが音声電話中に実行され得ると判定し、コマンドが音声電話中に実行され得るとの判定に応じて、コマンドに対する答えを示すための応答を決定するように構成される。

10

【0024】

一部の実装において、音声対応仮想アシスタントは、問い合わせに対応するコマンドを特定し、コマンドが音声電話中に実行され得ないと判定し、コマンドが音声電話中に実行され得ないと判定することとを含む。特定の実装において、コマンドが音声電話中に実行され得ないと判定することは、音声電話中に通常実行され得るコマンドのリストを取得することと、特定されたコマンドがコマンドのリストに載っていないと判定することとを含む。一部の実装において、コマンドが音声電話中に実行され得ないと判定することは、音声電話中に通常実行され得ないコマンドのリストを取得することと、特定されたコマンドがコマンドのリストに載っていると判定することとを含む。

20

【0025】

特定の実装においては、音声対応仮想アシスタントが問い合わせを処理したとの判定に応じて、第一者と第二者との間の音声電話の保留を解除するアクションは、音声電話の保留を解除する命令を音声コールプロバイダに与えることとを含む。一部の実装において、音声対応仮想アシスタントが問い合わせを処理したとの判定に応じて、第一者と第二者との間の音声電話の保留を解除するアクションは、マイクロフォンからの音声を音声対応仮想アシスタントではなくボイスサーバにルーティングすることと、音声対応仮想アシスタントからの音声中ではなくボイスサーバからの音声をスピーカーにルーティングすることとを含む。特定の実装においては、音声対応仮想アシスタントが問い合わせを処理したとの判定に応じて、第一者と第二者との間の音声電話の保留を解除するアクションは、デュアルトーンマルチ周波数信号を生成する命令を音声対応仮想アシスタントから受け取ることと、デュアルトーンマルチ周波数信号を生成する命令を音声対応仮想アシスタントから受け取ったことに応じて、音声電話の保留を解除する命令を音声コールプロバイダに与えた後、デュアルトーンマルチ周波数信号を生成する第2の命令を音声コールプロバイダに与えることとを含む。一部の実装において、音声対応アシスタントサーバは、問い合わせが1つまたは複数のデュアルトーンマルチ周波数信号を生成するコマンドおよび1つまたは複数のデュアルトーンマルチ周波数信号に対応する1つまたは複数の数を示すと判定するように構成される。

30

40

【0026】

この態様およびその他の態様のその他の実装は、コンピュータ記憶装置上に符号化された方法のアクションを実行するように構成された対応するシステム、装置、およびコンピュータプログラムを含む。1つまたは複数のコンピュータのシステムが、動作中にシステムにアクションを実行させるシステムにインストールされたソフトウェア、ファームウェア、ハードウェア、またはそれらの組合せによってそのように構成され得る。1つまたは

50

複数の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きに装置にアクションを実行させる命令を有することによってそのように構成され得る。

【0027】

本明細書に記載の対象の1つまたは複数の実装の詳細が、添付の図面および以下の説明に記載されている。対象のその他の特徴、態様、および潜在的な利点は、説明、図面、および請求項から明らかになるであろう。

【図面の簡単な説明】

【0028】

【図1A】電話を掛ける音声対応デバイスとの例示的なインタラクションを示すブロック図である。

10

【図1B】電話を掛ける音声対応デバイスとの例示的なインタラクションを示すブロック図である。

【図1C】電話を掛ける音声対応デバイスとの例示的なインタラクションを示すブロック図である。

【図1D】電話を掛ける音声対応デバイスとの例示的なインタラクションを示すブロック図である。

【図2】電話を掛けるためのプロセスの例を示す流れ図である。

【図3】電話を掛けるべきボイス番号を決定するためのプロセスの例を示す流れ図である。

【図4】電話中の音声対応デバイスとの例示的なインタラクションを示すブロック図である。

20

【図5】電話を掛ける音声対応デバイスとインタラクションするためのシステムの例を示すブロック図である。

【図6】発信者番号を決定するためのプロセスの例を示す流れ図である。

【図7】電話を掛けるべき受話者の番号を決定するためのプロセスの例を示す流れ図である。

【図8】音声電話中に問い合わせを処理するためのプロセスの例を示す流れ図である。

【図9】コンピューティングデバイスの例の図である。

【発明を実施するための形態】

【0029】

様々な図面における同様の参照番号および参照指示は、同様の要素を示す。

30

【0030】

図1A～図1Dは、システム100における異なる例示的なインタラクションを示すブロック図である。システム100は、ユーザ110にタッチによってシステム100と物理的にインタラクションをさせることなく、受話者155に電話を掛けるためにユーザ110によって使用され得る音声対応デバイス125を含む。

【0031】

一部の実装において、音声対応デバイス125は、音声対応デバイス125に呼びかけるためにユーザが言うホットワードとも呼ばれる予め決められた語句を含む発話の検出に応じてアクションを実行する可能性がある。たとえば、ホットワードは、ユーザが音声対応デバイス125に言う任意の要求の直前にユーザが言わなければならない「オーケー、コンピュータ」またはその他の語句である可能性がある。

40

【0032】

発信者番号を用いて電話を掛けるために、音声対応デバイス125は、発話を特定の知られているユーザによって言われたものとして分類し、特定の知られているユーザの発信者番号を用いて電話を掛ける可能性がある。知られているユーザは、システム100のユーザとして登録されているユーザである可能性があり、ゲストユーザは、システム100のユーザとして登録されていないユーザである可能性がある。たとえば、「母さん」が、音声対応デバイス125の知られているユーザとして登録する可能性があり、音声対応デバイス125は、発話が知られているユーザ「母さん」によって言われるかどうかを後で分類する可能性がある。

50

## 【 0 0 3 3 】

たとえば、図1Aは、音声対応デバイス125が発話「オーケー、コンピュータ。店Xに電話を掛けて」を受け取り、発話者を知られている発話者「マット」として分類し、「マット」に関する記憶された電話番号を用いて店Xに電話を掛けることを示す。別の例において、図1Bは、音声対応デバイス125が発話「オーケー、コンピュータ。店Xに電話を掛けて」を受け取り、発話者を知られている発話者「父さん」として分類し、店Xに非通知電話を掛けることを示す。さらに別の例において、図1Cは、音声対応デバイス125が発話「オーケー、コンピュータ。店Xに電話を掛けて」を受け取り、発話者をゲスト発話者として分類し、店Xに非通知電話を掛けることを示す。

## 【 0 0 3 4 】

さらに別の例において、図1Dは、発話者をゲスト発話者として分類し、仮番号を用いて緊急サービスに電話を掛ける発話「オーケー、コンピュータ。緊急電話を掛けて」を音声対応デバイス125が受け取ることを示す。仮番号は、緊急サービスが少なくとも特定の継続時間、たとえば、1時間、2時間、24時間などの間に音声対応デバイス125に折り返し電話を掛けるために使用することができるボイス番号である可能性がある。仮番号は、仮番号が緊急事態の間に折り返し電話を掛けるために緊急サービスによって使用されることのみ可能であるように発話者には知られていない可能性がある。

## 【 0 0 3 5 】

より詳細には、音声対応デバイス125は、1つまたは複数のマイクロフォンおよび1つまたは複数のスピーカーを含む可能性がある。音声対応デバイス125は、1つまたは複数のマイクロフォンを使用して発話を受け取り、1つまたは複数のスピーカーを通じて発話に対する聞き取ることができる応答を出力する可能性がある。

## 【 0 0 3 6 】

音声対応デバイス125は、音声対応デバイス125のそれぞれの知られているユーザに関するユーザアカウント情報を記憶する可能性がある。たとえば、音声対応デバイス125は、知られているユーザ「母さん」に関するユーザアカウント情報の第1の組132と、知られているユーザ「父さん」に関するユーザアカウント情報の第2の組134と、知られているユーザ「マット」に関するユーザアカウント情報の第3の組136とを記憶する可能性がある。

## 【 0 0 3 7 】

ユーザのユーザアカウント情報は、ユーザが電話を掛けるときに発信者番号として使用される可能性があるボイス番号を示す可能性がある。たとえば、「母さん」に関するユーザアカウント情報の第1の組132は、(555) 111-1111である第1の電話番号140を記憶する可能性があり、「父さん」に関するユーザアカウント情報の第2の組134は、空白である可能性があり(つまり、記憶された電話番号はない)、「マット」に関するユーザアカウント情報の第3の組136は、(555) 222-2222である第2の電話番号142を記憶する可能性がある。特定の実施形態において、ユーザに関するユーザアカウント情報は、「自宅」、「職場」、「モバイル」などの複数の番号を記憶する可能性がある。

## 【 0 0 3 8 】

ユーザのユーザアカウント情報は、発話者がユーザであるかどうかを認識するために使用される可能性がある発話者識別の特徴を示す可能性がある。たとえば、「母さん」に関するユーザアカウント情報の第1の組132は、事前にホットワードを複数回言うユーザ「母さん」を表す特徴ベクトルを集合的に形成し得るメル周波数ケプストラム係数(MFCC)の特徴を記憶する可能性がある。

## 【 0 0 3 9 】

一部の実装において、ユーザは、モバイルコンピューティングデバイス上の付随アプリケーションを通じて知られているユーザとして登録する可能性があり、モバイルコンピューティングデバイスは、ローカルワイヤレス接続を介して音声対応デバイス125と通信する。たとえば、ユーザ「母さん」は、ユーザの電話上の付随アプリケーションを通じてユーザのアカウントにログインし、そして、ユーザが音声対応デバイス125の知られている

10

20

30

40

50

ユーザとして登録したいことを付随アプリケーション内で示し、それから、ユーザの電話にパスワードを複数回言う可能性がある。

【0040】

登録の一部として、またはその後で、ユーザは、音声対応デバイス125を使用してユーザが掛ける電話のための発信者番号として使用するためのボイス番号をユーザが関連付けたいかどうかを示す可能性がある。たとえば、ユーザ「母さん」は、ユーザが音声対応デバイス125によってユーザの電話が掛けられるようにしたいことを示し、発信者番号がユーザの電話の電話番号であることを示す可能性がある。別の例において、ユーザ「母さん」は、ユーザの電話が、たとえば、Bluetooth(登録商標)接続を通じて音声対応デバイス125に接続されるときに、音声対応デバイス125によって掛けられたユーザの電話がユーザの電話を通過させられるようにしたいことを示す可能性がある。

10

【0041】

音声対応デバイス125は、複数の種類のコールプロバイダを通じて電話を掛ける可能性がある。たとえば、音声対応デバイス125は、インターネット接続を有し、ボイスオーバーインターネットプロトコル(VoIP)を使用して電話を掛ける可能性がある。別の例において、音声対応デバイス125は、セルラーネットワークと通信し、セルラーネットワークを使用して電話を掛ける可能性がある。さらに別の例において、音声対応デバイス125は、セルラー(または固定)電話と通信し、電話を通じて電話を掛ける可能性があり、したがって、ユーザは、音声対応デバイス125に話し、音声対応デバイス125を聞くが、電話は、電話を通じて確立される。

20

【0042】

一部の実装において、ユーザは、ユーザが使用したいコールプロバイダを選択することに基づいて、音声対応デバイス125を使用してユーザが掛ける電話のための発信者番号として使用するボイス番号を示す可能性がある。たとえば、「母さん」は、「母さん」が電話番号(555) 111-1111を使用して電話を受け取ることも可能である第1のコールプロバイダ、たとえば、セルラーネットワークプロバイダを通じて「母さん」の電話が掛けられるようにしたいことを示し、その後、代わりに、「母さん」が電話番号(555) 111-2222を使用して電話を受け取ることが可能である第2のコールプロバイダ、たとえば、VoIPプロバイダを通じて「母さん」の電話が掛けられるようにしたいことを示す可能性がある。

【0043】

一部の実装において、音声対応デバイス125は、コンテキスト情報(contextual information)に基づいて発話を特定のユーザによって言われたものとして分類する可能性がある。コンテキスト情報は、聴覚的、視覚的、またはその他の情報のうちの1つまたは複数を含む可能性がある。聴覚的情報に関して、音声対応デバイス125は、知られているユーザの1つまたは複数の発話の発話者識別の特徴(たとえば、特徴ベクトルを集合的に形成し得るメル周波数ケプストラム係数(MFCC)の特徴)に基づいて発話を分類する可能性がある。たとえば、音声対応デバイス125は、「オーケー、コンピュータ」と言う知られているユーザの各々に関する発話者識別の特徴を記憶する可能性がある。今受け取られた発話の発話者識別の特徴の「オーケー、コンピュータ」と言う知られているユーザ「父さん」の記憶された発話者識別の特徴との十分な一致に応じて、音声対応デバイス125は、発話を知られているユーザ「父さん」によって言われたものとして分類する可能性がある。

30

40

【0044】

別の例において、音声対応デバイス125は、発話の音声全体に基づいて発話を分類する可能性がある。たとえば、音声対応デバイス125は、受け取られた発話全体の中の語りが知られているユーザ「父さん」に対応する語りに一致するかどうかを判定する可能性がある。

【0045】

視覚的情報に関して、音声対応デバイス125は、発話者の少なくとも一部の1つまたは複数の画像を受け取り、1つまたは複数の画像に基づいて発話者を認識しようと試みる可能性がある。たとえば、音声対応デバイス125は、カメラを含み、カメラの視野内の発話

50

者が音声対応デバイス125が知られているユーザ「父さん」に対応する顔に一致すると分類する顔を有すると判定する可能性がある。その他の例において、音声対応デバイス125は、発話者の指紋、網膜スキャン、顔認識、姿勢、別のデバイスの共存、または別のデバイスもしくはソフトウェアの要素からのアイデンティティの確認のうちの1つまたは複数

【0046】

音声対応デバイス125は、遠隔のサーバと協力して電話を掛けるローカルのフロントエンドデバイスである可能性がある。たとえば、音声対応デバイス125が発話「オーケー、コンピュータ。店Xに電話を掛けて」を受け取る時、音声対応デバイス125は、発話者がホットワード「オーケー、コンピュータ」と言うときを検出し、発話「オーケー、コンピュータ」の中の発話者識別の特徴に基づいてユーザを「母さん」として分類し、「店Xに電話を掛けて」という表現と、発話者が「母さん」であるという指示とをサーバに与える可能性がある。サーバは、それから、「店Xに電話を掛けて」を文字に起こし、テキスト「店Xに電話を掛けて」が電話を掛けるアクションに対応し、店Xが電話番号(555) 999-9999を有し、発信者番号(555) 111-1111を有する「母さん」のVoIPアカウントを通じて「母さん」の電話が掛けられるべきであることを「母さん」が示したと判定する可能性がある。そして、サーバは、「VoIPアカウント(555) 111-1111を用いて(555) 999-9999に電話を掛ける」命令を音声対応デバイス125に送信する可能性がある。その他の実装において、音声対応デバイス125は、遠隔のサーバとは独立して、遠隔のサーバによって示されたアクションを実行する可能性がある。

【0047】

一部の実装において、音声対応デバイス125は、聴覚的情報および視覚的情報に加えてその他の情報に基づいて発話を分類する可能性がある。特に、音声対応デバイス125は、話したユーザのアイデンティティを認めるために、発話者識別の特徴およびユーザからの確認に基づいて発話を分類する可能性がある。加えて、音声対応デバイス125は、話したユーザのアイデンティティを認めるために、発話者の少なくとも一部の1つまたは複数の受け取られた画像およびユーザからの確認に基づいて発話を分類する可能性がある。たとえば、上述のように、音声対応デバイス125は、話したユーザからの1つまたは複数の発話を受け取る可能性がある。音声対応デバイス125は、1つまたは複数の受け取られた発話の中の発話者識別の特徴が「オーケー、コンピュータ」という知られているユーザ「父さん」の記憶された発話者識別の特徴と十分に一致すると判定する可能性がある。それに応じて、音声対応デバイス125は、「父さんが話していますか?」とユーザに尋ねることによって、話しているユーザが「父さん」であるという判定を確認する可能性がある。発話者は、音声対応デバイス125の確認を認めるために「はい」または「いいえ」と答えることによって応答することができる。発話者が「いいえ」と答えたならば、音声対応デバイス125は、「話している方の名前は何ですか?」などの追加の質問をして、名前が音声対応デバイス125に記憶された知られているユーザ名と一致するかどうかを判定する可能性がある。

【0048】

図2は、電話を掛けるためのプロセス200の例を示す流れ図である。プロセス200の動作は、図1A~図1Dのシステム100などの1つまたは複数のコンピューティングシステムによって実行される可能性がある。

【0049】

プロセス200は、発話を受け取ること(210)を含む。たとえば、音声対応デバイス125は、「オーケー、コンピュータ。(555) 999-9999に電話を掛けて」という発話を受け取る可能性がある。

【0050】

プロセス200は、電話が緊急サービスへのものであるかどうかを判定すること(212)を含む。たとえば、音声対応デバイス125は、(555) 999-9999がいかなる緊急サービスとも関連付けられていないので、番号への電話が緊急サービスへの電話ではないと判定する

可能性がある。別の例において、音声対応デバイス125は、番号「911」が緊急サービスに関連付けられているので、番号「911」への電話が緊急電話であると判定する可能性がある。

【0051】

プロセス200が電話が緊急サービスへのものであると判定する場合、プロセス200は、仮番号を用いて電話を開始すること(214)を含む。たとえば、音声対応デバイス125は、コールプロバイダが音声対応デバイスに折り返し電話をするために24時間使用され得る電話番号を生成し、それから、緊急サービスへの電話を開始し、仮番号を発信者番号として示すことを要求する可能性がある。

【0052】

プロセス200が電話が緊急サービスへのものでないと判定する場合、プロセス200は、発話の発話者が知られているユーザであるかどうかを判定すること(216)を含む。たとえば、音声対応デバイス125は、発話者を知られているユーザ「マット」とする分類に応じて、「オーケー、コンピュータ。(555) 999-9999に電話を掛けて」の発話者が知られているユーザであると判定する可能性がある。別の例において、音声対応デバイス125は、発話者を知られているユーザ「父さん」とする分類に応じて、発話者が知られているユーザであると判定する可能性がある。さらに別の例において、音声対応デバイス125は、発話者をゲストユーザとする分類に応じて、発話者が知られているユーザでないと判定する可能性がある。

【0053】

一部の実装において、発話の発話者が知られているユーザであるかどうかを判定することは、発話の中の語りが特定の知られているユーザに対応する語りと一致するかどうかを判定することを含む。たとえば、音声対応デバイス125は、発話者が「オーケー、コンピュータ」と言った言い方が知られているユーザ「マット」が「オーケー、コンピュータ」と言う言い方と一致すると判定し、それに応じて、発話者を知られているユーザ「マット」として分類する可能性がある。別の例において、音声対応デバイス125は、発話者が「オーケー、コンピュータ」と言った言い方が知られているユーザ「父さん」が「オーケー、コンピュータ」と言う言い方と一致すると判定し、それに応じて、発話者を知られているユーザ「父さん」として分類する可能性がある。追加的にまたは代替的に、発話の発話者が知られているユーザであるかどうかを判定することは、発話者の少なくとも一部の視覚的画像が特定の知られているユーザに対応する視覚的情報と一致するかどうかを判定することを含む。

【0054】

プロセス200が発話の発話者が知られているユーザであると判定する場合、プロセス200は、知られているユーザがパーソナルボイス番号に関連付けられているかどうかを判定すること(218)を含む。たとえば、音声対応デバイス125は、音声対応デバイス125を通じて電話を掛けるときに知られているユーザが使用したいコールプロバイダを示すアカウント情報を知られているユーザ「マット」が有すると判定し、それに応じて、知られているユーザが個人電話番号(personal phone number)に関連付けられていると判定する可能性がある。別の例において、音声対応デバイス125は、音声対応デバイス125を通じて電話を掛けるときに知られているユーザが使用したいコールプロバイダを示すアカウント情報を知られているユーザ「父さん」が持たないと判定し、それに応じて、知られているユーザが個人電話番号に関連付けられていないと判定する可能性がある。

【0055】

プロセス200が知られているユーザがパーソナルボイス番号に関連付けられていると判定する場合、プロセス200は、パーソナルボイス番号を用いて電話を開始すること(220)を含む。たとえば、音声対応デバイス125は、「マット」のアカウント情報によって示されるコールプロバイダに連絡し、「マット」のために電話番号(555) 999-9999に電話が掛けられることを要求する可能性がある。

【0056】

10

20

30

40

50

218に戻って、プロセス200が知られているユーザがパーソナルボイス番号に関連付けられていないと判定する場合、プロセスは、非通知電話を開始すること(222)を含む。たとえば、音声対応デバイス125は、コールプロバイダが(555) 999-9999に非通知電話を掛けることを要求する可能性がある。

【0057】

216に戻って、プロセス200が発話の発話者が知られているユーザでないと判定する場合、プロセス200は、222のために上述のように非通知電話を開始すること(222)を含む。

【0058】

電話が緊急サービスへのものであるかどうかを判定すること(212)がプロセス200に最初に示されるが、プロセス200は、異なる可能性がある。たとえば、プロセス200は、その代わりに最初に、(216)において上述されたように発話者が知られているユーザであると判定し、それから、(218)において上述されたように知られているユーザがパーソナルボイス番号に関連付けられていると判定し、次に、(212)において上述されたように電話が緊急サービスへのものであると判定し、そして、知られているユーザのパーソナルボイス番号を使用する可能性がある。音声対応デバイス125に関する仮番号の代わりに知られているユーザのパーソナルボイス番号を緊急時対応要員(emergency responder)に提供する1つの理由は、そのとき、知られているユーザが音声対応デバイス125の近くにいるか否かにかかわらず緊急時対応要員が知られているユーザに連絡することができることである。

【0059】

図3は、電話を掛けるべきボイス番号を決定するためのプロセス300の例を示す流れ図である。プロセス300の動作は、図1A~図1Dのシステム100などの1つまたは複数のコンピューティングシステムによって実行される可能性がある。

【0060】

プロセス300は、電話を要求する発話を受け取ること(310)を含む。たとえば、音声対応デバイス125は、「オーケー、コンピュータ。おばあちゃんに電話を掛けて」というような電話を要求するユーザ110に関する発話を受け取る可能性がある。

【0061】

プロセス300は、発話の発話者が知られているユーザであるかどうかを判定すること(312)を含む。たとえば、音声対応デバイス125は、発話者を知られているユーザ「母さん」として分類する可能性がある。

【0062】

プロセス300が発話の発話者が知られているユーザであると判定する場合、プロセス300は、知られているユーザに関して個人の連絡先が利用可能であるかどうかを判定すること(314)を含む。たとえば、音声対応デバイス125は、音声対応デバイス125が知られているユーザ「母さん」に関する連絡先のレコードにアクセスすることができると判定することに基づいて、知られているユーザ「母さん」に関して個人の連絡先が利用可能であると判定する可能性がある。知られているユーザに関する個人の連絡先は、知られているユーザに関して作成された電話連絡先のエントリを指す可能性がある。たとえば、知られているユーザは、新しい電話連絡先のエントリを作成するためのインターフェースを開き、電話番号「(123) 456-7890」および連絡先名「ジョン・ドウ」を打ち込み、それから、名前「ジョン・ドウ」によってラベル付けされ、電話番号「(123) 456-7890」を示す電話のエントリを作成することを選択することによって知られているユーザに関する電話連絡先のエントリを作成する可能性がある。知られているユーザの連絡先リストが、知られているユーザに関するすべての個人の連絡先によって形成される可能性がある。たとえば、知られているユーザに関する連絡先リストは、「ジョン・ドウ」に関する連絡先のエントリと、知られているユーザによって作成されたその他の連絡先のエントリとを含む可能性がある。

【0063】

プロセス300が知られているユーザに関して個人の連絡先が利用可能であると判定する

場合、プロセス300は、個人の連絡先を使用して受話者に関連する番号を決定すること(316)を含む。たとえば、音声対応デバイス125は、知られているユーザ「母さん」の連絡先のレコードから受話者「おばあちゃん」に関する個人の連絡先リストをスキャンし、「おばあちゃん」に関連する番号を取り出す。

【0064】

314に戻って、プロセス300が、そうではなく、知られているユーザに関する個人の連絡先が利用可能でないと判定する場合、プロセス300は、知られているユーザに関連する個人の連絡先なしに受話者の番号を決定すること(318)を含む。たとえば、音声対応デバイス125は、受話者の番号に関してインターネットを検索する可能性がある。この例において、音声対応デバイス125は、地理的位置特定サービスを使用して知られているユーザの近くににいる可能性がある「おばあちゃん」に対応する受話者の番号に関してインターネットを検索し、受話者の番号を特定することができず、「連絡先の番号が見つかりません」という音声メッセージを知られているユーザに与える可能性がある。受話者の番号が見つからない場合、音声対応デバイス125は、電話を掛けるべきボイス番号を言うように発話者に促し、そして、その番号に電話を掛ける可能性がある。

【0065】

312に戻って、プロセス300が、そうではなく、発話の発話者が知られているユーザではないと判定する場合、プロセス300は、上述のように個人の連絡先なしに受話者の番号を決定すること(318)を含む。

【0066】

図4は、電話中の音声対応デバイスとの例示的なインタラクションを示すブロック図である。図4は、示された順序または別の順序で実行され得る段階(A)から(C)までの様々な動作を示す。

【0067】

一部の実装において、音声対応デバイス125は、電話中に音声対応デバイス125に呼びかけるためにユーザが言うホットワードのような予め決められた語句を含む発話の検出に応じてアクションを実行する可能性がある。たとえば、図4は、音声対応デバイス125が発話「オーケー、コンピュータ。店Xに電話を掛けて」を受け取り、発話者を知られている発話者「マット」として分類し、「マット」に関する記憶された電話番号を用いて店Xに電話を掛けることを示す。加えて、発話者「マット」は、受話者155に聞こえないコマンドを電話中に音声対応デバイス125に伝える可能性がある。電話中のコマンドに回答して、音声対応デバイス125は、受話者からの発話の少なくとも一部をブロックすることができる。

【0068】

段階(A)の間に、音声対応デバイス125は、発話120「オーケー、コンピュータ。店Xに電話を掛けて」を受け取る。発話120を受け取ったことに応じて、音声対応デバイス125は、上述の方法のうちの1つを使用して発話者を知られている発話者「マット」として分類し、「あなたの番号を使って店Xに電話を掛けています」という応答を「マット」に返す。応答は、発話者を分類し、コマンドに関連するアクションを行い、「マット」に関連する番号を使用することによって音声対応デバイス125が発話を理解したことをユーザ110に示す。段階(B)の間に、音声対応デバイス125は、受話者155、たとえば、店Xへの電話を開始する。たとえば、音声対応デバイス125は、ユーザ110と受話者155との間の電話を開始する。音声対応デバイス125は、ユーザ110に折り返し電話を掛けるために受話者155によって使用され得るユーザ110の番号を使用して受話者155に電話を掛ける。受話者155は、「はい」ということによって電話に答える。それに応じて、ユーザ110は、音声対応デバイス125を介して受話者155に「もしもし、お店屋さん。開いていますか?」と言う。受話者155は、「ええ。午後10時閉店です」と応答する。

【0069】

段階(B)の間に、音声対応デバイス125は、受話者155との電話中にユーザ110からのコマンドからホットワードを検出する。たとえば、音声対応デバイス125は、「オーケー、



コンピュータ。今何時」と言うコマンドをユーザ110から得る。電話中に受け取られた発話に応じて、音声対応デバイス125は、ユーザ110が言うホットワード「オーケー、コンピュータ」を送信するが、それから、ホットワードの後のコマンドはブロックし、したがって、受話者155は、「オーケー、コンピュータ」を聞くが、「今何時」は聞かない。音声対応デバイス125は、「午後9時です」と言ってユーザ110にだけ応答し、したがって、受話者155は、応答を聞かない。代替的に、ホットワードを電話の一部として受話者にブロードキャストする前に音声対応デバイス125がホットワードを検出することを可能にするために、通信にある量のレイテンシーが導入され得る。このようにして、ホットワードに関連する命令だけでなくホットワード自体が、電話の一部として受話者に配信されることをブロックされ得る。

10

【0070】

一部の実装において、音声対応デバイス125は、ユーザ110がホットワードを言うことを検出した後、ユーザ110と受話者155との間を両方向とも保留(2-way hold)にすることによって、受話者155がユーザ110と音声対応デバイス125との間のコミュニケーションを聞くことを防止する可能性がある。両方向とも保留にしている間、受話者155およびユーザ110は、互いに聞くことができない可能性がある。たとえば、発話「オーケー、コンピュータ。今何時」を受け取ったことに応じて、音声対応デバイス125は、「オーケー、コンピュータ」の直後および「今何時」の前に両方向とも保留を開始する可能性があり、したがって、店Xの受話者155は、「オーケー、コンピュータ」のみを聞く。

【0071】

20

音声対応デバイス125は、音声対応デバイス125がユーザからのコマンドが解決されたと判定すると両方向とも保留することやめる可能性がある。たとえば、音声対応デバイス125は、応答「午後9時です」がユーザの質問「今何時」に答えると判定し、それに応じて、両方向とも保留することをやめる可能性がある。別の例において、音声対応デバイス125は、「オーケー、コンピュータ。アラームを午後7時にセットして」と言うユーザ110に応じて、「何曜日の午後7時にアラームをセットしますか」と応答し、ユーザ110が曜日を与えるために両方向とも保留し続ける可能性がある。その他の実施形態において、ユーザ110は、たとえば、「オーケー、コンピュータ。電話を保留にして」と言うことによって音声対応デバイス125に電話を保留するように要求する可能性がある。音声対応デバイス125は、たとえば、「オーケー、コンピュータ。電話を再開して」と言うことによってユーザが保留をやめるように要求するまで電話を保留し続ける可能性がある。

30

【0072】

一部の实装において、音声対応デバイス125は、ユーザ110と長いインタラクションをするコマンドをブロックする可能性がある。たとえば、音声対応デバイス125は、音楽、ニュース、またはポッドキャストなどのメディアを再生すること、日報(daily brief)を再生すること、第三者の会話アクション、追加の電話を掛けること、および雑学クイズなどのゲームをすることに関連する特徴をブロックする可能性がある。音声対応デバイス125は、これらの機能をブロックするときにエラーを出し、たとえば、「すみません。電話中に音楽は再生できません」と出力するか、またはこれらのタスクのうちの1つに関連するいずれのコマンドも無視し、電話を続ける可能性がある。

40

【0073】

段階(C)の間に、音声対応デバイス125は、店Xの受話者155との電話中にユーザ110からの別のコマンドからホットワードを検出する。たとえば、音声対応デバイス125は、「オーケー、コンピュータ。電話を切って」と言うコマンドをユーザ110から得る。電話中に受け取られた発話に応じて、音声対応デバイス125は、「電話が終了しました」と言って、または非言語的な音声の合図を出して、ユーザ110に応答する。加えて、音声対応デバイス125は、店Xの受話者155に応答「電話が終了しました」または非言語的な音声の合図を送信しない。

【0074】

図5は、電話を掛ける音声対応デバイスとインタラクションするためのシステム500の

50

例を示すブロック図である。システム500は、音声対応デバイス125、アシスタントサーバ502、連絡先データベース504、ボイスサーバ506、クライアントデバイス510、ネットワーク508、ならびに通信回線512および514を含む。

【0075】

一部の実装において、音声対応デバイス125は、1つまたは複数のコンピュータを含むことができ、複数の地理的位置に分散されたコンピュータを含む可能性がある。音声対応デバイス125は、1つまたは複数のクライアントデバイス510、アシスタントサーバ502、およびボイスサーバ506と通信する。

【0076】

一部の実装において、アシスタントサーバ502およびボイスサーバ506は、それぞれ、1つまたは複数のコンピュータを含むことができ、複数の地理的位置に分散されたコンピュータを含む可能性がある。アシスタントサーバ502は、音声対応デバイス125および連絡先データベース504と通信する。ボイスサーバ506は、音声対応デバイス125および店Xなどの1つまたは複数の受話者と通信する。

【0077】

クライアントデバイス510は、たとえば、デスクトップコンピュータ、ラップトップコンピュータ、タブレットコンピュータ、ウェアラブルコンピュータ、セルラー電話、スマートフォン、音楽プレーヤー、電子ブックリーダー、ナビゲーションシステム、または任意のその他の適切なコンピューティングデバイスであることが可能である。ネットワーク508は、有線、またはワイヤレス、またはこれら両方の組合せであることが可能であり、インターネットを含むことができる。

【0078】

一部の実装において、音声対応デバイス125は、Bluetooth(登録商標)、WiFi、またはその他の近距離通信プロトコルなどの近距離通信プロトコルを使用して通信回線512を介してクライアントデバイス510に接続する可能性がある。たとえば、音声対応デバイス125は、それぞれ関連する通信回線512によって最大7つの異なるクライアントデバイス510をペアリングし、接続する可能性がある。一部の実装において、音声対応デバイス125は、任意の所与の時間にクライアントデバイス510のうちの1つからの音声ルーティングする可能性がある。

【0079】

一部の実装において、音声対応デバイス125は、ユーザ110から発話「オーケー、コンピュータ。店Xに電話を掛けて」を受け取る可能性がある。音声対応デバイス125は、さらに、発話者(ユーザ110)を知られている発話者「マット」として分類する可能性がある。たとえば、音声対応デバイス125は、「マット」に関連するユーザアカウント情報に含まれる発話者識別の特徴をユーザ110によって言われた受け取られたホットワードと比較する可能性がある。音声対応デバイス125は、比較に応じてユーザ110が「マット」であると判定する可能性がある。一部の実装において、音声対応デバイス125は、それから、発話の音声表現を、さらに処理するためにアシスタントサーバ502に問い合わせとして送信する可能性がある。

【0080】

一部の実装において、音声対応デバイス125は、ユーザ110が電話を掛けることを要求するときに様々なイベントを停止する可能性がある。たとえば、音声対応デバイス125は、ユーザが「オーケー、コンピュータ。店Xに電話を掛けて」と言うと音楽の再生またはアラームを停止する可能性がある。ユーザ110が電話を掛けることを要求するときに様々なイベントを停止するために、音声対応デバイス125は、ユーザが電話を掛けることを要求しているときに停止されるべき特定の種類のイベントを記憶し、ユーザが電話を掛けていることの検出に応じて、それらの記憶された特定の種類のイベントを終了する可能性がある。たとえば、音声対応デバイス125は、ユーザが電話を掛けるときに停止されるべき音楽再生およびアラームのイベントを記憶し、ユーザが電話を掛けていることの検出に応じて、音楽再生およびアラームのいずれのイベントも終了するが、その他のイベントは継

10

20

30

40

50

続する可能性がある。

【0081】

一部の実装において、音声対応デバイス125は、電話を掛ける前に任意のイベントを無効化するようにユーザ110に要求する可能性がある。たとえば、音声対応デバイス125は、現在、音楽を再生しているか、またはアラームもしくはタイマーが原因で鳴っている可能性がある。音声対応デバイス125は、ユーザ110が音楽またはアラームもしくはタイマーの音をとめるまでユーザ110がいかなる電話を掛けることも許さない可能性がある。一部の实装において、ユーザ110は、それぞれ、「オーケー、コンピュータ。音楽をとめて」または「オーケー、コンピュータ。アラームをとめて」ということによって音楽またはアラームもしくはタイマーの音を無効化する可能性がある。その他の実装において、ユーザ110は、音声対応デバイス125上のインタラクティブボタンをタップすることによって音楽またはアラームもしくはタイマーの音を無効化する可能性がある。たとえば、音声対応デバイス125は、ユーザが電話を掛けることを要求するときに無効化するためのユーザインタラクションを必要とする特定のイベントを記憶する可能性がある。ユーザが電話を掛けることを要求し、特定のイベントのうち少なくとも1つが発生していることの検出に応じて、音声対応デバイス125は、「電話を掛ける前にイベントを無効化してください」という警告メッセージをユーザに言い、電話を掛ける要求を無視する可能性がある。ユーザが音声対応デバイス125に音声コマンドを送信するかまたは音声対応デバイス125上のインタラクティブボタンをタップすることによって特定のイベントを無効化するように音声対応デバイス125に命じると、ユーザは、電話を掛けるように音声対応デバイス125に要求する可能性がある。

10

20

【0082】

一部の実装において、音声対応デバイス125は、電話を掛けるコマンドをユーザ110から受け取ったことに応じてユーザ110にやがて起こるアラームを警告する可能性がある。たとえば、ユーザ110は、アラームを午後6:30に音声対応デバイス125上で鳴るように設定する可能性がある。ユーザ110は、午後6:29に音声対応デバイス125に発話「オーケー、コンピュータ。店Xに電話を掛けて」という可能性がある。発話を受け取ったことに応じて、音声対応デバイス125は、ユーザに対して出力し、「電話を掛ける前にアラームを無効化してください」または「アラームが1分後の午後6:30に設定されています。電話を掛ける前にこのアラームを無効化したいですか?」という可能性がある。その後、ユーザ110は、音声対応デバイス125を用いて電話を掛ける前にアラームを無効化するかまたはアラームを聞き流す可能性がある。

30

【0083】

一部の実装において、音声対応デバイス125は、アラームが掛けられている電話の予め決められた長さの時間、たとえば、1分、5分、15分、または何らかのその他の長さの時間以内に鳴るように設定されているかどうかを判定することに基づいて、ユーザ110にやがて起こるアラームを警告する可能性がある。たとえば、音声対応デバイス125は、午後6:29に電話を掛ける要求を受け取り、午後6:29から5分以内の、午後6:30にアラームが設定されていると判定し、アラームが午後6:29から5分以内に設定されているとの判定に応じて、ユーザ110にやがて起こるアラームの警告を与える可能性がある。

40

【0084】

一部の実装において、アシスタントサーバ502は、要求516を取得する。たとえば、音声対応デバイス125は、ユーザ110から受け取られた発話の音声表現を示す検索要求を含むデータを送信する可能性がある。データは、特定された知られている発話者「マット」、発話「オーケー、コンピュータ。店Xに電話を掛けて」120の音声表現、音声対応デバイス125に関連する一意ID、および特定された知られている発話者「マット」に関連する個人結果ビット(personal results bit)を示す可能性がある。音声対応デバイス125に関連する一意IDは、応答をどこに送信すべきかをアシスタントサーバ502に示す。たとえば、一意IDは、音声対応デバイス125に関連するIPアドレス、URL、またはMACアドレスである可能性がある。

50

## 【0085】

一部の実装において、アシスタントサーバ502は、取得された要求516を処理する。特に、アシスタントサーバ502は、発話に関連するコマンドを決定するために取得された要求516を解析する。たとえば、アシスタントサーバ502は、発話の音声表現を発話のテキスト表現に変換することによって取得された要求516を処理する可能性がある。変換に応じて、アシスタントサーバ502は、ホットワードに続くコマンド「店Xに電話を掛けて」に関するテキスト表現を解析する。一部の实装において、アシスタントサーバ502は、テキストのコマンドに関連するアクションを決定する。たとえば、アシスタントサーバ502は、テキストのアクション「電話を掛ける」を記憶されたテキストのアクションと比較することによって取得された要求516からのアクションが「店Xに電話を掛ける」ことであると決定する。

10

## 【0086】

加えて、アシスタントサーバ502は、連絡先データベース504にアクセスすることによって受話者「店X」に関する番号を解決する。一部の实装において、アシスタントサーバ502は、知られているユーザに関連する連絡先を取り出すために連絡先データベース504にアクセスする。連絡先データベース504は、連絡先に関連する知られているユーザ名によって連絡先をインデックス付けすることによって連絡先を記憶する。たとえば、連絡先データベース504は、「マット」に関連する個人の連絡先をさらに含む「マット」に関するエントリを含む。個人の連絡先は、「母さん」-(555)111-1111、「父さん」-(555)222-2222、および「店X」-(555)333-3333などの名前および関連する番号を含む。

20

## 【0087】

加えて、アシスタントサーバ502は、取得された要求516内で受け取られた個人結果ビットが有効化されるときにのみ受話者に関する番号を解決する可能性がある。個人結果ビットが有効化されていないまたは「0」である場合、アシスタントサーバ502は、「コンピュータに個人の連絡先へのアクセスを許可してください」というメッセージをユーザ110に中継するように音声対応デバイス125に示すためにアクションメッセージ518内で識別子を送信する。個人結果ビットが有効化されているまたは「1」である場合、アシスタントサーバ502は、特定された知られている発話者の個人の連絡先に関して連絡先データベース504にアクセスする。一部の实装において、アシスタントサーバ502は、特定された知られている発話者の個人の連絡先の中の受話者に関連する番号を取り出す。この例において、アシスタントサーバ502は、店Xに関する番号(555) 333-3333を取り出す。その他の実装において、受話者の番号は、ホットワードに続くコマンドに関するテキスト表現に含まれる可能性がある。たとえば、コマンドは、「オーケー、コンピュータ。555-333-3333に電話を掛けて」を含む可能性がある。

30

## 【0088】

一部の实装において、アシスタントサーバ502は、連絡先データベース504内の特定された知られている発話者の個人の連絡先の中に見つからない取得された要求516内の受話者を特定する可能性がある。たとえば、アシスタントサーバ502は、「おばあちゃんに電話を掛けて」を含む取得された要求516からのホットワードに続くコマンドのテキスト表現を決定する可能性がある。しかし、「マット」に関連する連絡先データベース504からの個人の連絡先は、「おばあちゃん」に関するエントリを含まない。もっと正確に言えば、連絡先は、「母さん」、「父さん」、および「店X」を含む。受話者「おばあちゃん」に関する番号を解決するために、アシスタントサーバ502は、「おばあちゃん」に関する番号を見つけるためにその他のデータベースおよび/またはインターネットを検索する可能性がある。

40

## 【0089】

その他のデータベースおよび/またはインターネットを検索する際、アシスタントサーバ502は、ナレッジグラフ内を検索する可能性がある。たとえば、アシスタントサーバ502は、「X社カスタマーサービス」をユーザの個人の連絡先の中のいかなるレコードともマッチングせず、そして、名前「X社カスタマーサービス」を有するエンティティに関して

50

ナレッジグラフを検索し、そのエンティティに関してナレッジグラフに記憶された電話番号を特定する可能性がある。

【0090】

一部の実装において、コマンドは、音声対応デバイス125と地理的に近い会社に電話を掛けることを含む可能性がある。アシスタントサーバ502は、音声対応デバイス125に最も近い会社に関連するボイス番号に関してインターネットを検索する可能性がある。しかし、アシスタントサーバ502が要求された受話者に関連する番号を発見しないならば、アシスタントサーバ502は、「連絡先が見つかりません」というメッセージをユーザ110に中継するように音声対応デバイス125に示すためにアクションメッセージ518内で識別子を送信する可能性がある。たとえば、アシスタントサーバ502は、個人の連絡先のレコードまたはナレッジグラフ内で「店X」に関する電話番号を見つけることができない場合、「店X」の名前を有する近くの地元の会社に関して地図データベース内を検索する可能性がある。

10

【0091】

一部の実装において、アシスタントサーバ502は、コマンドに含まれる番号がサポートされないボイス番号である可能性があると判定する可能性がある。たとえば、番号が、123-4567のように7桁しか含まない可能性がある。それに応じて、アシスタントサーバ502は、「サポートされない電話番号です」というメッセージをユーザ110に中継するように音声対応デバイス125に示すためにアクションメッセージ518内で識別子を送信する可能性がある。

20

【0092】

受話者に関連する連絡先の番号の決定に応じて、アシスタントサーバ502は、音声対応デバイス125へのアクションメッセージ518を生成する。特に、アクションメッセージ518は、連絡先の番号および電話をトリガするためのアクションを含む可能性がある。たとえば、アクションメッセージ518は、「店X」に関する電話番号を555-333-3333として含み、「店X」にすぐに電話を掛けるように音声対応デバイス125に命じるアクションを含む可能性がある。一部の実装において、アシスタントサーバ502は、コマンドのコンテキスト(context)に基づいて使用すべき発信番号(outbound number)をアクションメッセージ518に含める可能性がある。たとえば、コマンドが緊急サービスへの電話を含む場合、アシスタントサーバ502は、受話者155が音声対応デバイス125に折り返し電話するために時間の特定の期間使用することができる番号をアクションメッセージ518に含める可能性がある。たとえば、電話番号(555) 888-8888が、音声対応デバイス125に折り返し電話を掛けるためにその後2時間使用される可能性がある。

30

【0093】

一部の実装において、音声対応デバイス125は、アシスタントサーバ502からアクションメッセージ518を得る。アクションメッセージ518の取得に応じて、音声対応デバイス125は、アクションメッセージ518に対するアクションを起こす。たとえば、アクションメッセージは、示された電話番号555-333-3333を使用して「店X」に電話を掛けるように音声対応デバイス125に示す。

【0094】

一部の実装において、音声対応デバイス125は、ユーザ110のプリファレンスに基づいてボイスサーバ506または関連するクライアントデバイス510を使用して、アシスタントサーバ502によって示されたように受話者に電話を掛ける可能性がある。特に、ユーザ110のプリファレンスは、音声対応デバイス125に記憶される可能性がある。たとえば、音声対応デバイス125は、ユーザ110のプリファレンスが任意のアウトバウンドコールのためにボイスサーバ506またはボイスオーバIP(VoIP)を使用することであると判定する可能性がある。したがって、音声対応デバイス125は、受話者に電話を掛けるためにボイスサーバ506に指示を送信する。一部の実装において、ボイスサーバ506は、アウトバウンドコールのために関連する番号を使用する可能性がある。一部の実装において、音声対応デバイス125は、ユーザが複数の異なるVoIPプロバイダの中からVoIPプロバイダを使用す

40

50

ることを選択することを可能にし、そして、そのユーザが今後電話を開始するときそのVoIPプロバイダを使用する可能性がある。

【0095】

一部の実装において、音声対応デバイス125は、ユーザ110が音声対応デバイス125の近くに在るとの判定に応じて、緊急サービスに電話を掛けるためにボイスサーバ506に関連する番号を使用する可能性がある。たとえば、音声対応デバイス125は、クライアントデバイス510のうちの1つが音声対応デバイス125に接続されているとの判定に応じて、ボイスサーバ506に関連する番号を使用して緊急サービスに電話を掛ける可能性がある。クライアントデバイス510と音声対応デバイス125との間の接続を保証することにより、音声対応デバイス125は、ユーザ110が音声対応デバイス125の近くに在ることを保証し得る。

10

【0096】

代替的に、音声対応デバイス125は、ユーザ110の補助的なプリファレンスが受話者にアウトバウンドコールをするために既存のクライアントデバイス510を使用することであると判定する可能性がある。ユーザ110の補助的なプリファレンスが関連するクライアントデバイス510を使用して受話者に電話を掛けることであると音声対応デバイス125が判定する場合、音声対応デバイス125は、クライアントデバイス510への通信回線512を検証する。たとえば、音声対応デバイス125は、クライアントデバイス510へのBluetooth(登録商標)接続を検証する可能性がある。音声対応デバイス125がクライアントデバイス510へのBluetooth(登録商標)接続をすることができない場合、音声対応デバイス125は、「あなたのBluetooth(登録商標)接続が有効になっているかどうか確かめてください」と言うメッセージをユーザ110に中継する可能性がある。Bluetooth(登録商標)接続が確立されると、音声対応デバイス125は、受話者に電話を掛けるためにクライアントデバイス510に指示を送信する。その他の実施形態において、音声対応デバイス125が近距離通信プロトコルのいずれの手段によってもクライアントデバイス510を発見することができないならば、音声対応デバイス125は、受話者に対してボイスサーバ506によって非公開番号を使用して受話者に電話を掛ける可能性がある。

20

【0097】

一部の実装において、音声対応デバイス125は、受話者の電話への接続に応じて、ユーザ110が聞くための聞き取ることができる音声を再生する可能性がある。たとえば、音声対応デバイス125は、受話者の電話が答えるために利用可能である場合、聞き取ることができる呼び出し音を再生する可能性がある。別の例において、音声対応デバイス125は、受話者の電話が答えるために利用可能でない場合、話中信号音を再生する可能性がある。別の例において、音声対応デバイス125は、受話者の電話番号が無効である場合、「サポートされない電話番号です」などの音声メッセージをユーザに与える可能性がある。その他の実施形態において、ユーザ110は、受話者の電話に電話をつなごうと試みる間に、受話者の電話への電話を切断するために音声対応デバイス125上のインタラクティブボタンをタップする可能性がある。

30

【0098】

一部の実装において、音声対応デバイス125は、ユーザ110によって掛けられた一番最近の電話をリダイヤルする可能性がある。たとえば、ユーザ110は、番号を言わずに「オーケー、コンピュータ。リダイヤルして」と言うことができ、音声対応デバイス125は、電話を掛けられた最後の受話者の番号をリダイヤルする。一部の実装においては、音声対応デバイス125が一番最近の電話をリダイヤルするために、音声対応デバイス125は、それぞれの電話の後に一番最近の電話に関連する設定をメモリに記憶する。メモリ内の一番最近の電話に関連する設定は、電話を掛けるユーザ、電話を掛けるために使用された番号、および受話者の番号を含む。

40

【0099】

一部の実装において、音声対応デバイス125は、音声自動応答システムをナビゲートするためのデュアルトーンマルチ周波数(Dual Tone Multiple Frequencies)(DTMF)トーン

50

ンを受け取る可能性がある。たとえば、ユーザ110は、「オーケー、コンピュータ。Nを押しして」と言うことができ、Nは、\*キー、#キー、または0から9までの数字である。それに応じて、音声対応デバイス125は、「オーケー、コンピュータ」を検出した後、両方向とも保留にし、受話者155に送信される数字Nに関するダイヤルトーンを生成し、両方向とも保留することをやめる可能性がある。

**【0100】**

一部の実装において、音声対応デバイス125は、ユーザ110にステータスライトを提供する可能性がある。たとえば、ステータスライトは、音声対応デバイス125のステータスを示すためのLEDライトであることが可能である。ステータスライトは、電話をつなげること、電話がつながったこと、電話が終了したこと、ユーザから音声コマンドを受け取ること、およびユーザ110にメッセージを与えることを示すための色、点滅の継続時間、または明るさを変化させる可能性がある。

10

**【0101】**

一部の実装において、ユーザ110は、特定の音声コマンドによって電話を終了させる可能性がある。たとえば、ユーザ110は、「オーケー、コンピュータ。電話をやめて」、「オーケー、コンピュータ。電話を切って」、または「オーケー、コンピュータ。電話を切断して」と言うことができる。一部の実装においては、受話者が、電話を終了する可能性がある。電話が終了された後、音声対応デバイス125は、聞き取ることができる話中音を再生し、音声対応デバイス125を電話をつなぐ前の以前の状態に戻す可能性がある。たとえば、音声対応デバイス125を以前の状態に戻すことは、電話が開始されたときに歌などのメディアが停止した時点のメディアの再生を続けることを含む可能性がある。

20

**【0102】**

一部の実装において、音声対応デバイス125は、着信を受け取られるときを示す可能性がある。たとえば、音声対応デバイス125は、音声対応デバイス125が電話を受け取っていることを示すためにLEDを光らせるか、鳴り響く音を聞こえるよう出力するか、または「着信です」と聞こえるよう出力する可能性がある。それに応じて、ユーザ110は、着信に対してアクションを起こす可能性がある。たとえば、ユーザ110は、いくつか例を挙げるとすれば、以下、すなわち、「オーケー、コンピュータ。電話を取って」、「オーケー、コンピュータ。答えて」、「オーケー、コンピュータ。受け取って」、または「オーケー、コンピュータ。いいよ」のうちの1つを言うことによって電話に答える可能性がある。別の例において、ユーザ110は、いくつか例を挙げるとすれば、以下、すなわち、「オーケー、コンピュータ。ダメだ」、「オーケー、コンピュータ。受け取らないで」、または「オーケー、コンピュータ。電話を切って」のうちの1つを言うことによって電話を拒否し、接続の試みを切断する可能性がある。

30

**【0103】**

一部の実装において、音声対応デバイス125は、仮番号によって掛けられた着信のみを受け入れる可能性がある。特に、音声対応デバイス125は、緊急サービスへの発信をするために使用された仮番号への電話から着信を受け取られるときにのみ鳴る可能性がある。たとえば、音声対応デバイス125は、緊急サービスにダイヤルするためにアウトバウンドコール用の仮番号として番号(555) 555-5555を使用する可能性があり、番号(555) 555-5555への着信のみを受け入れる可能性がある。

40

**【0104】**

一部の実装において、ユーザ110は、スピーカーフォンとして使用するために、別のデバイス上の着信を音声対応デバイス125に転送する可能性がある。ユーザ110は、電話が鳴っている間にまたは電話中に電話を転送する可能性がある。たとえば、ユーザ110は、「オーケー、コンピュータ。私の電話からあなたに電話を転送して」と言う可能性がある。一部の実装において、音声対応デバイス125は、電話を転送するために近距離通信プロトコルを使用してその別のデバイスと通信する可能性がある。たとえば、音声対応デバイス125は、たとえば、音声対応デバイス125のスピーカーに現在の電話をルーティングするように別のデバイスに命じるためにBluetooth(登録商標)またはWiFiを使用して別のデ

50

バイスに接続する可能性がある。

【0105】

一部の実装において、ユーザ110は、音声対応デバイス125からクライアントデバイス510に電話を転送する可能性がある。特に、ユーザ110は、電話が鳴っている間にまたは電話中に電話を転送する可能性がある。これは、クライアントデバイス510がBluetooth(登録商標)などの近距離通信プロトコルのうちの少なくとも1つを使用して音声対応デバイス125に接続される場合に実行される可能性がある。たとえば、ユーザ110は、「オーケー、コンピュータ。私の電話に電話を転送して」という可能性がある。さらに、ユーザ110は、ある音声対応デバイス125から別室に置かれた別の音声対応デバイス125に電話を転送する可能性がある。たとえば、ユーザ110は、「オーケー、コンピュータ。寝室のコンピュータに電話を転送して」という可能性がある。クライアントデバイス510または別の音声対応デバイス125が電源が入っていないかまたは音声対応デバイス125に接続されていない場合、音声対応デバイス125は、「接続を確立するためにデバイスをオンにしてください」という可能性がある。

10

【0106】

図6は、発信者番号を決定するためのプロセス600の例を示す流れ図である。プロセス600の動作は、システム500などの1つまたは複数のコンピューティングシステムによって実行される可能性がある。

【0107】

プロセス600は、音声電話を要求する発話を受け取ること(610)を含む。たとえば、音声対応デバイス125は、ユーザが「オーケー、コンピュータ。(123) 456-7890に電話を掛けて」と言うときに発話を受け取る可能性があり、そのとき、音声対応デバイス125のマイクロフォンが、発話に対応する音声データを生成する。一部の実装において、音声電話は、音声のみを含む電話を指す可能性がある。その他の実装において、音声電話は、音声のみを含むのではない電話、たとえば、音声と映像との両方を含むテレビ会議電話を指す可能性がある。

20

【0108】

プロセス600は、発話を特定の知られているユーザによって言われたものとして分類すること(620)を含む。たとえば、音声対応デバイス125は、発話「オーケー、コンピュータ。(123) 456-7890に電話を掛けて」を特定の知られているユーザ「マット」によって言われたものとして分類する可能性がある。別の例において、音声対応デバイス125は、発話「オーケー、コンピュータ。(123) 456-7890に電話を掛けて」を音声対応デバイスに知られていないユーザによって言われたものとして分類する可能性がある。

30

【0109】

発話を特定の知られているユーザによって言われたものとして分類することは、発話の中の語りが特定の知られているユーザに対応する語りと一致するかどうかを判定することを含む可能性がある。たとえば、上述のように、音声対応デバイス125は、事前にホットワード「オーケー、コンピュータ」を言う知られているユーザ「マット」に対応するMFCCを記憶し、今受け取られた発話の中のホットワード「オーケー、コンピュータ」からのMFCCを決定し、それから、発話からのMFCCが知られているユーザ「マット」に関して記憶されたMFCCに一致すると判定し、それに応じて、発話を知られているユーザ「マット」によって言われたものとして分類する可能性がある。別の例において、音声対応デバイス125は、事前にホットワード「オーケー、コンピュータ」を言う知られているユーザ「マット」に対応するMFCCを記憶し、今受け取られた発話の中のホットワード「オーケー、コンピュータ」からのMFCCを決定し、それから、発話からのMFCCが知られているユーザ「マット」に関して記憶されたMFCCに一致しないと判定し、それに応じて、発話を知られているユーザ「マット」によって言われたものとして分類しない可能性がある。

40

【0110】

発話を特定の知られているユーザによって言われたものとして分類することは、発話者の少なくとも一部の視覚的画像が特定の知られているユーザに対応する視覚的情報と一致

50



するかどうかを判定することを含む可能性がある。たとえば、上述のように、音声対応デバイス125は、カメラを含み、カメラによって撮影された発話者の顔の画像を取得し、画像内の発話者の顔が知られているユーザ「マット」の顔を示す情報と一致すると判定し、その判定に応じて、発話者を知られているユーザ「マット」として分類する可能性がある。別の例において、音声対応デバイス125は、カメラを含み、カメラによって撮影された発話者の顔の画像を取得し、画像内の発話者の顔が知られているユーザ「マット」の顔を示す情報と一致しないと判定し、その判定に応じて、発話者を知られているユーザ「マット」ではないものとして分類する可能性がある。一部の実装においては、発話が特定の知られているユーザによって言われたかどうかを分類するために、視覚的画像および語りが、組み合わせて考慮される可能性がある。

10

**【0111】**

プロセス600は、特定の知られているユーザがパーソナルボイス番号に関連付けられているかどうかを判定すること(630)を含む。たとえば、音声対応デバイス125は、知られているユーザ「マット」が個人電話番号(555) 222-2222に関連付けられていると判定する可能性がある。別の例において、音声対応デバイス125は、特定の知られているユーザ「父さん」が個人の番号に関連付けられていないと判定する可能性がある。

**【0112】**

特定の知られているユーザがパーソナルボイス番号に関連付けられているかどうかを判定することは、特定の知られているユーザのアカウント情報にアクセスすることと、ユーザのアカウント情報が特定の知られているユーザに関するボイス番号を記憶するかどうかを判定することを含む可能性がある。たとえば、音声対応デバイス125は、音声対応デバイス125に記憶された知られているユーザ「マット」のアカウント情報にアクセスし、アカウント情報が個人電話番号(555) 222-2222を含むと判定し、それに応じて、知られているユーザ「マット」が個人の番号に関連付けられていると判定する可能性がある。別の例において、音声対応デバイス125は、音声対応デバイス125に記憶された知られているユーザ「父さん」のアカウント情報にアクセスし、アカウント情報が個人電話番号を含まないと判定し、それに応じて、知られているユーザ「父さん」が個人の番号に関連付けられていないと判定する可能性がある。

20

**【0113】**

追加的にまたは代替的に、特定の知られているユーザがパーソナルボイス番号に関連付けられているかどうかを判定することは、特定の知られているユーザの指示および発話の表現をサーバに提供することと、特定の知られているユーザのパーソナルボイス番号、電話を掛けるべきボイス番号、および音声電話を掛ける命令をサーバから受け取ることを含む可能性がある。たとえば、一部の実装において、音声対応デバイス125は、個人電話番号を記憶しない可能性があり、アシスタントサーバ502が、個人電話番号を記憶する可能性がある。したがって、音声対応デバイス125は、発話「オーケー、コンピュータ。(123)456-7890に電話を掛けて」の音声表現を、発話者が知られているユーザ「マット」であるという指示と一緒にアシスタントサーバ502に提供する可能性がある。そして、アシスタントサーバ502は、発話を文字に起こし、文字に起こされたものの中の「電話を掛けて」から、発話が電話を開始することを要求していると判定し、文字に起こされたものから、「(123) 456-7890」が電話を掛けるべき番号であると判定し、発話が電話を要求しているとの判定に応じて、知られているユーザ「マット」に関する記憶されたアカウント情報にアクセスし、知られているユーザ「マット」に関する記憶されたアカウントがパーソナルボイス番号(555) 222-2222を含むと判定し、それに応じて、番号(123) 456-7890に電話を掛ける命令を音声対応デバイス125に与え、(555) 222-2222を電話を開始している電話番号として示す可能性がある。

30

40

**【0114】**

特定の知られているユーザがパーソナルボイス番号に関連付けられているかどうかを判定することは、特定の知られているユーザのアカウントにアクセスすることと、ユーザのアカウントが電話を示すかどうかを判定することと、電話が音声対応デバイスに接続され

50

ていると判定することを含む可能性がある。たとえば、音声対応デバイス125が発話を知られているユーザ「マット」によって話されたものとして分類した後、音声対応デバイス125は、特定の電話が知られているユーザ「マット」に関連するものとして示されるかどうかを判定するために記憶されたアカウント情報にアクセスし、アカウントが特定の電話を示すとの判定に応じて、特定の電話が、たとえば、Bluetooth(登録商標)(登録商標)によって接続されているかどうかを判定し、特定の電話が接続されているとの判定に応じて、特定の電話によって電話を開始する可能性がある。

#### 【0115】

プロセス600は、パーソナルボイス番号を用いて音声電話を開始すること(640)を含む。たとえば、音声対応デバイス125は、個人の番号「(555) 222-2222」を使用して「(123) 456-7890」への電話を開始する命令をボイスサーバ506に与える可能性がある。一部の実装において、パーソナルボイス番号を用いて電話を開始することは、VoIPプロバイダを通じて電話を開始することを含む可能性がある。たとえば、ボイスサーバ506が、VoIPプロバイダである可能性があり、音声対応デバイス125が、ボイスサーバ506が電話を開始することを要求する可能性がある。別の例において、音声対応デバイス125は、音声対応デバイスに接続されていると判定された知られているユーザ「マット」に関連する電話に電話を開始する命令を与える可能性がある。

#### 【0116】

図7は、電話を掛けるべき受話者の番号を決定するためのプロセスの例を示す流れ図である。プロセス600の動作は、システム500などの1つまたは複数のコンピューティングシステムによって実行される可能性がある。

#### 【0117】

プロセス700は、音声電話を要求する発話を受け取ること(710)を含む。たとえば、アシスタントサーバ502は、「おばあちゃんに電話を掛けて」という発話の表現と、発話が知られているユーザ「マット」によって言われたものとして音声対応デバイス125によって判定されたという指示とを音声対応デバイス125から受け取る可能性がある。指示は、その他のユーザのアカウントからマットのアカウントを一意に特定する英数字の値、または発話の発話者が英数字の値によって特定されたアカウントに関連付けられているかどうかを示す英数字の値に関連する2進値の包含である可能性がある。

#### 【0118】

プロセスは、発話を特定の知られているユーザによって言われたものとして分類すること(720)を含む。たとえば、アシスタントサーバ502は、発話を知られているユーザ「マット」によって話されたものとして分類する可能性がある。発話を特定の知られているユーザによって言われたものとして分類することは、発話の中の語りが特定の知られているユーザに対応する語りに一致すると音声対応デバイスによって判定されたという指示を取得することを含む可能性がある。たとえば、アシスタントサーバ502は、発話「おばあちゃんに電話を掛けて」の発話者に一致するものとして知られているユーザ「マット」のアカウントを一意に特定する値「854978」を音声対応デバイス125が提供したと判定し、それに応じて、発話を知られているユーザ「マット」によって言われたものとして分類する可能性がある。

#### 【0119】

追加的にまたは代替的に、発話を特定の知られているユーザによって言われたものとして分類することは、発話の中の語りが特定の知られているユーザに対応する語りと一致するかどうかを判定することを含む可能性がある。たとえば、アシスタントサーバ502は、発話の音声表現からMFCCを生成し、発話からのMFCCが知られているユーザ「マット」に関して記憶されたMFCCに一致するかどうかを判定し、MFCCが一致するとの判定に応じて、発話を知られているユーザ「マット」によって言われたものとして分類する可能性がある。

#### 【0120】

プロセス700は、発話を特定の知られているユーザによって言われたものとする分類に

10

20

30

40

50

応じて、特定の知られているユーザに関する連絡先に基づいて電話を掛けるべき受話者のボイス番号を決定すること(730)を含む。たとえば、「おばあちゃんに電話を掛けて」を知られているユーザ「マット」によって言われたものとする分類に応じて、アシスタントサーバ502は、知られているユーザ「マット」に関して記憶された電話連絡先に基づいて電話を掛けるべき受話者の番号「(987) 654-3210」を決定する可能性がある。別の例においては、「おばあちゃんに電話を掛けて」を知られているユーザ「父さん」によって言われたものとする分類に応じて、アシスタントサーバ502は、知られているユーザ「父さん」に関して記憶された電話連絡先に基づいて電話を掛けるべき受話者の番号「(876) 543-2109」を決定する可能性がある。

**【0121】**

特定の知られているユーザによって作成された連絡先のエントリを取得することは、発話を特定の知られているユーザによって言われたものとする分類に応じて、特定の知られているユーザの連絡先のエントリが利用可能であると判定することと、特定の知られているユーザの連絡先のエントリが利用可能であるとの判定に応じて、特定の知られているユーザによって作成された連絡先のエントリを取得することを含む可能性がある。たとえば、発話を知られているユーザ「マット」によって言われたものとする分類に応じて、アシスタントサーバ502は、知られているユーザ「マット」に関する電話連絡先のエントリが利用可能であると判定し、それに応じて、知られているユーザ「マット」の電話連絡先のエントリにアクセスする可能性がある。

**【0122】**

特定の知られているユーザの連絡先のエントリが利用可能であると判定することは、特定の知られているユーザがパーソナライズされた結果が欲しいことを特定の知られているユーザが事前に示したかどうかを判定することを含む可能性がある。たとえば、アシスタントサーバ502は、発話と一緒に、音声対応デバイス125からパーソナライズされた結果ビットを受け取り、パーソナライズされた結果ビットが知られているユーザ「マット」がパーソナライズされた結果が欲しいことを示す値に設定されていると判定し、それに応じて、知られているユーザ「マット」の電話連絡先のエントリが利用可能であると判定する可能性がある。別の例において、アシスタントサーバ502は、発話と一緒に、音声対応デバイス125からパーソナライズされた結果ビットを受け取り、パーソナライズされた結果ビットが知られているユーザ「父さん」がパーソナライズされた結果が欲しくないことを示す値に設定されていると判定し、それに応じて、知られているユーザ「父さん」の電話連絡先のエントリが利用可能でないと判定する可能性がある。

**【0123】**

発話を特定の知られているユーザによって言われたものとする分類に応じて、特定の知られているユーザに関する連絡先に基づいて電話を掛けるべき受話者のボイス番号を決定することは、発話を特定の知られているユーザによって言われたものとする分類に応じて、特定の知られているユーザによって作成された連絡先のエントリを取得することと、連絡先のエントリの中から特定の連絡先のエントリを特定することと、特定の連絡先のエントリが、発話に一致する名前を含む、特定することと、特定の連絡先のエントリによって示されたボイス番号を受話者のボイス番号として決定することを含む可能性がある。たとえば、発話「おばあちゃんに電話を掛けて」を知られているユーザ「マット」によって言われたものとする分類に応じて、アシスタントサーバ502は、知られているユーザ「マット」によって作成された電話連絡先のエントリを取得し、電話連絡先のエントリのうちの1つが発話の中の「おばあちゃん」に一致する「おばあちゃん」と名付けられ、番号「(987) 654-3210」を有することを特定し、受話者の電話番号が番号「(987) 654-3210」であると決定する可能性がある。

**【0124】**

連絡先のエントリの中から特定の連絡先のエントリを特定することと、特定の連絡先のエントリが、発話に一致する名前を含む、特定することは、発話の文字に起こされたものを生成することと、文字に起こされたものが名前を含むと判定することを含む可

10

20

30

40

50

能性がある。たとえば、アシスタントサーバ502は、発話「おばあちゃんに電話を掛けて」の文字に起こされたものを生成し、文字に起こされたものからの「おばあちゃん」が知られているユーザ「マット」の電話連絡先のエントリーに関する名前「おばあちゃん」と同一であると判定し、それに応じて、「おばあちゃん」と名付けられた連絡先のエントリーを特定する可能性がある。

**【0125】**

プロセス700は、受話者のボイス番号への音声電話を開始すること(740)を含む。たとえば、アシスタントサーバ502は、「おばあちゃん」と名付けられた知られているユーザの電話連絡先から取得された受話者の電話番号「(987) 654-3210」への電話を開始する可能性がある。受話者のボイス番号への音声電話を開始することは、受話者のボイス番号と、受話者のボイス番号への音声電話を開始する命令とを音声対応デバイスに与えることを含む可能性がある。たとえば、アシスタントサーバ502は、番号(555) 222-2222を用いて番号(987) 654-3210への電話を開始する命令を音声対応デバイス125に与える可能性がある。

10

**【0126】**

一部の実装において、プロセス700は、第2の音声電話を要求する第2の発話を受け取ることと、第2の発話を音声対応デバイス125のいかなる知られているユーザによっても言われていないものとして分類することと、第2の発話を音声対応デバイスのいかなる知られているユーザによっても言われていないものとする分類に応じて、音声対応デバイスのいかなる知られているユーザに関する連絡先にもアクセスせずに第2の音声電話を開始することを含む可能性がある。たとえば、アシスタントサーバ502は、第2の発話「店Xに電話を掛けて」を受け取り、第2の発話を音声対応デバイス125のいかなる知られているユーザによっても言われていないものとして分類し、発話の中の「店X」が電話番号ではないと判定し、第2の発話を音声対応デバイスのいかなる知られているユーザによっても言われていないものとして分類し、発話の中の「店X」が電話番号ではないとの分類に応じて、「店X」の名前を有する近くの地元の会社に関して地図データベースを検索し、名前「店X」および電話番号「(765) 432-1098」を有する単一の近くの地元の会社を特定し、それに応じて、音声対応デバイスのいかなる知られているユーザに関する電話連絡先にもアクセスせずに(765) 432-1098への第2の電話を開始する可能性がある。

20

**【0127】**

図8は、音声電話中に問い合わせを処理するためのプロセスの例を示す流れ図である。プロセス800の動作は、システム500などの1つまたは複数のコンピューティングシステムによって実行される可能性がある。

30

**【0128】**

プロセス800は、第一者が第一者と第二者との間の音声電話中に音声対応仮想アシスタントへの問い合わせを言ったと判定すること(810)を含む。たとえば、音声対応デバイス125は、ユーザがユーザと別の人の間の電話中にアシスタントサーバ502への問い合わせを言ったと判定する可能性がある。第一者が第一者と第二者との間の電話中に音声対応仮想アシスタントへの問い合わせを言ったと判定することは、電話中に第一者によってキーワードが言われたと音声対応デバイスによって判定することを含む可能性がある。たとえば、音声対応デバイス125は、電話が音声対応デバイス125によって継続中である間にキーワード「オーケー、コンピュータ」が言われたと判定する可能性がある。電話は、別の人へのユーザからの語りを拾い、ユーザへの別の人の語りを出力するために音声対応デバイス125のマイクロフォンおよびスピーカが使用されているときに、音声対応デバイス125によって継続中であると考えられる可能性がある。

40

**【0129】**

プロセス800は、第一者が第一者と第二者との間の電話中に音声対応仮想アシスタントへの問い合わせを言ったとの判定に応じて、第一者と第二者との間の音声電話を保留にすること(810)を含む。たとえば、第一者が第一者と第二者との間の電話中に音声対応仮想アシスタントへの問い合わせ「オーケー、コンピュータ。私の次の予定は何?」を言ったと

50

の判定に応じて、音声対応デバイス125は、電話を両方向とも保留にする可能性がある。音声電話は、別の人がユーザから音声対応仮想アシスタントへの問い合わせを聞かない可能性があり、問い合わせに対する音声対応仮想アシスタントからの応答を聞かない可能性があるように両方向とも保留にされる可能性がある。

【0130】

プロセス800は、音声電話を保留にすること(820)を含む。たとえば、音声対応デバイス125は、電話を両方向とも保留にする可能性がある。第一者と第二者との間の音声電話を保留にすることは、音声電話を保留にする命令を音声コールプロバイダに与えることを含む可能性がある。たとえば、音声対応デバイス125は、継続中の電話を保留するようにボイスサーバ506に命じる可能性がある。追加的にまたは代替的に、第一者と第二者との間の音声電話を保留にすることは、マイクロフォンからの音声をボイスサーバではなく音声対応仮想アシスタントにルーティングすることと、ボイスサーバからの音声をスピーカーにルーティングすることとを含む可能性がある。たとえば、音声対応デバイス125は、音声対応デバイス125のマイクロフォンからの音声をボイスサーバ506ではなくアシスタントサーバ502にルーティングし、ボイスサーバ506からの音声をスピーカーにルーティングする可能性がある。

【0131】

プロセス800は、音声対応仮想アシスタントが問い合わせを解決したと判定すること(830)を含む。たとえば、音声対応デバイス125は、アシスタントサーバ502が問い合わせ「オーケー、コンピュータ。私の次の予定は何」を解決したと判定する可能性がある。音声対応仮想アシスタントが問い合わせを解決したと判定することは、問い合わせ、および音声電話が音声対応デバイス上で継続中であるという指示を音声対応仮想アシスタントに与えることと、問い合わせへの応答および問い合わせが解決されるという指示を音声対応仮想アシスタントから受け取ることとを含む可能性がある。たとえば、音声対応デバイス125は、問い合わせ「オーケー、コンピュータ。私の次の予定は何」の表現および「継続中の電話=True(Ongoing call=True)」という指示を与え、それに応じて、「あなたの次の予定は午後3:30の『コーヒープレイク』です」という合成された語りの表現を問い合わせに対する応答として受け取り、「問い合わせ解決済み=True(Query resolved=True)」という指示を受け取る。

【0132】

一部の実装において、音声対応仮想アシスタントは、問い合わせに対応するコマンドを特定し、コマンドが音声電話中に実行され得ると判定し、コマンドが音声電話中に実行され得るとの判定に応じて、コマンドに対する答えを示すための応答を決定するように構成される可能性がある。たとえば、アシスタントサーバ502は、発話「オーケー、コンピュータ。私の次の予定は何」の表現を受け取り、発話の表現から文字に起こされたものからコマンド「次の予定を特定して」を特定し、コマンド「次の予定を特定して」が電話中に実行され得ると判定し、コマンドが電話中に実行され得るとの判定に応じて、「あなたの次の予定は午後3:30の『コーヒープレイク』です」という答えを示すための応答を決定する可能性がある。

【0133】

一部の実装において、音声対応仮想アシスタントは、問い合わせに対応するコマンドを特定し、コマンドが音声電話中に実行され得ないと判定し、コマンドが音声電話中に実行され得ないと判定に応じて、コマンドが実行され得ないことを示すための応答を決定するように構成される可能性がある。たとえば、アシスタントサーバ502は、発話「オーケー、コンピュータ。何か音楽をかけて」の表現を受け取り、発話の表現から文字に起こされたものからコマンド「音楽をかけて」を特定し、コマンド「音楽をかけて」が電話中に実行され得ないと判定し、コマンドが電話中に実行され得ないと判定に応じて、「すみません。電話中は音楽を再生することができません」という答えを示すための応答を決定する可能性がある。

10

20

30

40

50

## 【0134】

一部の実装において、コマンドが音声電話中に実行され得ないと判定することは、音声電話中に通常実行され得るコマンドのリストを取得することと、特定されたコマンドがコマンドのリストに載っていないと判定することを含む。たとえば、アシスタントサーバ502は、「次の予定を特定して」を含み、「音楽をかけて」を含まない、実行され得るコマンドのリストを取得し、コマンド「音楽をかけて」がリスト内で特定されないと判定し、それに応じて、コマンド「音楽をかけて」が電話中に通常実行され得ないと判定する可能性がある。

## 【0135】

一部の実装において、コマンドが音声電話中に実行され得ないと判定することは、音声電話中に通常実行され得ないコマンドのリストを取得することと、特定されたコマンドがコマンドのリストに載っていると判定することを含む。たとえば、アシスタントサーバ502は、「音楽をかけて」を含み、「次の予定を特定して」を含まない、実行され得ないコマンドのリストを取得し、コマンド「音楽をかけて」がリスト内で特定されると判定し、それに応じて、コマンド「音楽をかけて」が電話中に通常実行され得ないと判定する可能性がある。

10

## 【0136】

プロセス800は、音声対応仮想アシスタントが問い合わせを処理したとの判定に応じて、第一者と第二者との間の音声電話の保留を解除すること(840)を含む。たとえば、音声対応デバイス125は、電話を再開する可能性がある。音声対応仮想アシスタントが問い合わせを処理したとの判定に応じて、第一者と第二者との間の音声電話の保留を解除することは、音声電話の保留を解除する命令を音声コールプロバイダに与えることを含む可能性がある。たとえば、音声対応デバイス125は、電話の保留を解除する命令をボイスサーバ506に与える可能性がある。

20

## 【0137】

追加的にまたは代替的に、音声対応仮想アシスタントが問い合わせを処理したとの判定に応じて、第一者と第二者との間の音声電話の保留を解除することは、マイクロフォンからの音声を音声対応仮想アシスタントではなくボイスサーバにルーティングすることと、音声対応仮想アシスタントからの音声中ではなくボイスサーバからの音声をスピーカーにルーティングすることを含む可能性がある。たとえば、音声対応デバイス125は、マイクロフォンからの音声をアシスタントサーバ502ではなくボイスサーバ506にルーティングする可能性があり、アシスタントサーバ502からの音声中ではなくボイスサーバ506からの音声をスピーカーにルーティングする可能性がある。

30

## 【0138】

一部の実装において、音声対応仮想アシスタントが問い合わせを処理したとの判定に応じて、第一者と第二者との間の音声電話の保留を解除することは、デュアルトーンマルチ周波数信号を生成する命令を音声対応仮想アシスタントから受け取ることと、デュアルトーンマルチ周波数信号を生成する命令を音声対応仮想アシスタントから受け取ったことに応じて、音声電話の保留を解除する命令を音声コールプロバイダに与えた後、デュアルトーンマルチ周波数信号を生成する第2の命令を音声コールプロバイダに与えることとを含む可能性がある。たとえば、音声対応デバイス125は、「1のDTMFを生成して」という命令を受け取り、それに応じて、「1」キーの押下を表すDTMFを生成するようにボイスサーバ506に命じる可能性がある。

40

## 【0139】

一部の実装において、音声対応アシスタントサーバは、問い合わせが1つまたは複数のデュアルトーンマルチ周波数信号を生成するコマンドおよび1つまたは複数のデュアルトーンマルチ周波数信号に対応する1つまたは複数の数を示すと判定するように構成される。たとえば、アシスタントサーバ502は、発話「オーケー、コンピュータ。1を押して」の表現を受け取り、文字に起こされたものから、「1を押して」が文字に起こされたものの中の「1」によって表された数に関するDTMF信号を生成することを示すと判定し、それ

50

に応じて、「1」に関するDTMFを生成するようにボイスサーバ506に命じるように音声対応デバイス125に命じる命令を音声対応デバイス125に与える可能性がある。追加的にまたは代替的に、一部の実装においては、音声対応デバイス125が、DTMFを生成する可能性がある。たとえば、音声対応デバイス125は、「1」に関するDTMFを生成する命令をアシスタントサーバ502から受け取り、それに応答して、「1」に関するDTMFトーンを生成し、それらのトーンをボイスサーバ506に送信する可能性がある。

#### 【0140】

上の説明に加えて、ユーザは、本明細書に記載のシステム、プログラム、または特徴がユーザ情報(たとえば、ユーザのソーシャルネットワーク、社会的行為もしくは活動、職業、ユーザのプリファレンス、またはユーザの現在位置についての情報)の収集を可能にしてもよいかどうかといつ可能にしてもよいかの両方、およびユーザがコンテンツまたは通信をサーバから送信されるかどうかについての選択をユーザが行うことを可能にするコントロールを提供される可能性がある。さらに、特定のデータが、個人を特定できる情報が削除されるように、記憶されるかまたは使用される前に1つまたは複数の方法で処理される可能性がある。たとえば、ユーザのアイデンティティが、個人を特定できる情報がユーザに関して決定され得ないか、または位置情報が取得される場合にユーザの地理的位置が(都市、郵便番号、もしくは州のレベルまでになど)一般化される可能性があり、したがって、ユーザの特定の位置が決定され得ないように処理される可能性がある。したがって、ユーザは、どの情報がユーザについて収集されるか、その情報がどのように使用されるのか、およびどの情報がユーザに提供されるのかをコントロールすることができる可能性がある。

#### 【0141】

システム100の異なる構成が、使用される可能性があり、音声対応デバイス125、アシスタントサーバ502、およびボイスサーバ506の機能が、組み合わされるか、さらに分割されるか、分散されるか、または入れ替えられる可能性がある。たとえば、アシスタントサーバ502が文字に起こすために問い合わせに発話の音声表現を含める代わりに、音声対応デバイス125が、発話を文字に起こし、アシスタントサーバ502への問い合わせに文字に起こされたものを含める可能性がある。

#### 【0142】

図9は、本明細書に記載の技術を実装するために使用され得るコンピューティングデバイス900およびモバイルコンピューティングデバイス950の例を示す。コンピューティングデバイス900は、ラップトップ、デスクトップ、ワークステーション、携帯情報端末、サーバ、ブレードサーバ、メインフレーム、およびその他の適切なコンピュータなどの様々な形態のデジタルコンピュータを表すように意図される。モバイルコンピューティングデバイス950は、携帯情報端末、セルラー電話、スマートフォン、およびその他の同様のコンピューティングデバイスなどの様々な形態のモバイルデバイスを表すように意図される。本明細書において示された構成要素、それらの接続および関係、ならびにそれらの機能は、例であるように意図されているに過ぎず、限定であるように意図されていない。

#### 【0143】

コンピューティングデバイス900は、プロセッサ902、メモリ904、記憶装置906、メモリ904および複数の高速拡張ポート910に接続する高速インターフェース908、ならびに低速拡張ポート914および記憶装置906に接続する低速インターフェース912を含む。プロセッサ902、メモリ904、記憶装置906、高速インターフェース908、高速拡張ポート910、および低速インターフェース912の各々は、様々なバスを使用して相互に接続され、共通のマザーボードに搭載されるか、または適宜その他の方法で搭載される可能性がある。プロセッサ902は、メモリ904内または記憶装置906上に記憶された命令を含む、コンピューティングデバイス900内で実行するための命令を処理して、高速インターフェース908に結合されたディスプレイ916などの外部入力/出力デバイス上のグラフィカルユーザインターフェース(GUI)のためのグラフィカルな情報を表示することができる。その他の実装においては、複数のプロセッサおよび/または複数のバスが、複数のメモリおよび

10

20

30

40

50

複数の種類のメモリと一緒に適宜使用される可能性がある。また、複数のコンピューティングデバイスが、各デバイスが必要な動作の一部を提供するようにして(たとえば、サーババンク、一群のブレードサーバ、またはマルチプロセッサシステムとして)接続される可能性がある。

【0144】

メモリ904は、コンピューティングデバイス900内で情報を記憶する。一部の実装において、メモリ904は、1つの揮発性メモリユニットまたは複数の揮発性メモリユニットである。一部の実装において、メモリ904は、1つの不揮発性メモリユニットまたは複数の不揮発性メモリユニットである。メモリ904は、磁気または光ディスクなどの別の形態のコンピュータ可読媒体である可能性もある。

10

【0145】

記憶装置906は、コンピューティングデバイス900に大容量記憶を提供することができる。一部の実装において、記憶装置906は、フロッピーディスクデバイス、ハードディスクデバイス、光ディスクデバイス、またはテープデバイス、フラッシュメモリもしくはその他の同様のソリッドステートメモリデバイス、またはストレージエリアネットワークもしくはその他の構成内のデバイスを含むデバイスのアレイなどのコンピュータ可読媒体であるか、またはそのようなコンピュータ可読媒体を含む可能性がある。命令は、情報担体に記憶され得る。命令は、1つまたは複数の処理デバイス(たとえば、プロセッサ902)によって実行されるとき、上述の方法などの1つまたは複数の方法を実行する。命令は、コンピュータ可読媒体または機械可読媒体(たとえば、メモリ904、記憶装置906、またはプロセッサ902上のメモリ)などの1つまたは複数の記憶装置によって記憶されることも可能である。

20

【0146】

高速インターフェース908が、コンピューティングデバイス900に関する帯域を大量に消費する動作を管理する一方、低速インターフェース912は、帯域をそれほど消費しない動作を管理する。そのような機能の割り振りは、例であるに過ぎない。一部の实装において、高速インターフェース908は、メモリ904に、(たとえば、グラフィックスプロセッサまたはアクセラレータを通じて)ディスプレイ916に、および様々な拡張カード(図示せず)を受け入れる可能性がある高速拡張ポート910に結合される。実装において、低速インターフェース912は、記憶装置906および低速拡張ポート914に結合される。様々な通信ポート(たとえば、USB、Bluetooth(登録商標)、イーサネット(登録商標)、ワイヤレスイーサネット(登録商標))を含む可能性がある低速拡張ポート914は、キーボード、ポインティングデバイス、スキャナなどの1つもしくは複数の入力/出力デバイスと結合される可能性があり、またはたとえばネットワークアダプタを介してスイッチもしくはルータなどのネットワークングデバイスと結合される可能性がある。

30

【0147】

コンピューティングデバイス900は、図に示されるように、多くの異なる形態で実装される可能性がある。たとえば、コンピューティングデバイス900は、1つの標準的なサーバ920として実装されるか、または一群のそのようなサーバ内で複数回実装される可能性がある。さらに、コンピューティングデバイス900は、ラップトップコンピュータ922などのパーソナルコンピュータ内で実装される可能性がある。コンピューティングデバイス900は、ラックサーバシステム924の一部として実装される可能性もある。代替的に、コンピューティングデバイス900の構成要素は、モバイルコンピューティングデバイス950などのモバイルデバイスのその他の構成要素(図示せず)と組み合わせられる可能性がある。そのようなデバイスの各々は、コンピューティングデバイス900およびモバイルコンピューティングデバイス950のうちの1つまたは複数を含む可能性があり、システム全体が、互いに通信する複数のコンピューティングデバイスによって構成される可能性がある。

40

【0148】

モバイルコンピューティングデバイス950は、構成要素の中でもとりわけ、プロセッサ952、メモリ964、ディスプレイ954などの入力/出力デバイス、通信インターフェース9

50



66、およびトランシーバ968を含む。モバイルコンピューティングデバイス950は、追加的なストレージを提供するために、マイクロドライブまたはその他のデバイスなどの記憶装置を備える可能性もある。プロセッサ952、メモリ964、ディスプレイ954、通信インターフェース966、およびトランシーバ968の各々は、様々なバスを使用して相互に接続されており、構成要素のうちのいくつかは、通常のマザーボードに搭載されるか、または適宜その他の方法で搭載される可能性がある。

【0149】

プロセッサ952は、メモリ964に記憶された命令を含むモバイルコンピューティングデバイス950内の命令を実行することができる。プロセッサ952は、別々の複数のアナログおよびデジタルプロセッサを含むチップのチップセットとして実装される可能性がある。プロセッサ952は、たとえば、ユーザインターフェース、モバイルコンピューティングデバイス950によって実行されるアプリケーション、およびモバイルコンピューティングデバイス950によるワイヤレス通信の制御などの、モバイルコンピューティングデバイス950のその他の構成要素の調整を行う可能性がある。

10

【0150】

プロセッサ952は、ディスプレイ954に結合された制御インターフェース958およびディスプレイインターフェース956を通じてユーザとコミュニケーションする可能性がある。ディスプレイ954は、たとえば、TFT(薄膜トランジスタ液晶ディスプレイ)ディスプレイもしくはOLED(有機発光ダイオード)ディスプレイ、またはその他の適切なディスプレイテクノロジーである可能性がある。ディスプレイインターフェース956は、ユーザに対してグラフィカルな情報およびその他の情報を提示するようにディスプレイ954を駆動するための適切な回路を含む可能性がある。制御インターフェース958は、ユーザからコマンドを受け取り、それらのコマンドを、プロセッサ952に送るために変換する可能性がある。加えて、外部インターフェース962が、その他のデバイスとのモバイルコンピューティングデバイス950の近い地域の通信を可能にするために、プロセッサ952との通信を提供する可能性がある。外部インターフェース962は、たとえば、一部の実装においては有線通信を、またはその他の実装においてはワイヤレス通信を提供する可能性があり、複数のインターフェースが使用される可能性もある。

20

【0151】

メモリ964は、モバイルコンピューティングデバイス950内で情報を記憶する。メモリ964は、1つのコンピュータ可読媒体もしくは複数のコンピュータ可読媒体、1つの揮発性メモリユニットもしくは複数の揮発性メモリユニット、または1つの不揮発性メモリユニットもしくは複数の不揮発性メモリユニットのうちの1つまたは複数として実装され得る。また、拡張メモリ974が設けられ、たとえば、SIMM(シングルインラインメモリモジュール(Single In Line Memory Module))カードインターフェースを含む可能性がある拡張インターフェース972を通じてモバイルコンピューティングデバイス950に接続される可能性がある。拡張メモリ974は、モバイルコンピューティングデバイス950に追加的なストレージ空間を提供する可能性があり、またはモバイルコンピューティングデバイス950に関するアプリケーションまたはその他の情報を記憶する可能性もある。特に、拡張メモリ974は、上述のプロセスを実行または補足する命令を含む可能性があり、安全な情報を含む可能性もある。したがって、たとえば、拡張メモリ974は、モバイルコンピューティングデバイス950のセキュリティモジュールとして設けられる可能性があり、モバイルコンピューティングデバイス950の安全な使用を可能にする命令を用いてプログラムされる可能性がある。さらに、ハッキングすることができない方法でSIMMカードに識別情報を置くなど、追加的な情報とともに、安全なアプリケーションがSIMMカードによって提供される可能性がある。

30

40

【0152】

メモリは、たとえば、以下で検討されるように、フラッシュメモリおよび/またはNVRAMメモリ(不揮発性ランダムアクセスメモリ)を含む可能性がある。一部の実装において、1つまたは複数の処理デバイス(たとえば、プロセッサ952)によって実行されるとき、上述

50

の方法などの1つまたは複数の方法を実行する命令が、情報担体に記憶される。命令は、1つまたは複数のコンピュータ可読媒体または機械可読媒体(たとえば、メモリ964、拡張メモリ974、またはプロセッサ952上のメモリ)などの1つまたは複数の記憶装置によって記憶されることも可能である。一部の実装において、命令は、たとえば、トランシーバ968または外部インターフェース962を介して伝播信号内で受け取られ得る。

【0153】

モバイルコンピューティングデバイス950は、必要に応じてデジタル信号処理回路を含む可能性がある通信インターフェース966を通じてワイヤレスで通信する可能性がある。通信インターフェース966は、とりわけ、GSM(登録商標)音声電話(移動体通信用グローバルシステム(Global System for Mobile communications))、SMS(ショートメッセージサービス)、EMS(拡張メッセージングサービス(Enhanced Messaging Service))、またはMMSメッセージング(マルチメディアメッセージングサービス)、CDMA(登録商標)(符号分割多元接続)、TDMA(時分割多元接続)、PDC(パーソナルデジタルセルラー)、WCDMA(登録商標)(広帯域符号分割多元接続)、CDMA2000、またはGPRS(汎用パケット無線サービス)などの様々なモードまたはプロトコルの下で通信を提供する可能性がある。そのような通信は、たとえば、無線周波数を用いるトランシーバ968を通じて行われる可能性がある。さらに、近距離通信が、Bluetooth(登録商標)、WiFi、またはその他のそのようなトランシーバ(図示せず)を用いるなどして行われる可能性がある。加えて、GPS(全地球測位システム)受信機モジュール970が、モバイルコンピューティングデバイス950で実行されるアプリケーションによって適宜使用される可能性があるさらなるナビゲーションおよび位置に関連するワイヤレスデータをモバイルコンピューティングデバイス950に提供する可能性がある。

【0154】

モバイルコンピューティングデバイス950は、ユーザから話された情報を受け取りし、その情報を使用可能なデジタル情報に変換する可能性がある音声コーデック960を使用して音声通信する可能性もある。同じく、音声コーデック960は、たとえば、モバイルコンピューティングデバイス950のハンドセットのスピーカーを介するなどして、ユーザのための聞き取ることができる音声を生成する可能性がある。そのような音声は、音声電話の音声を含む可能性があり、記録された音声(たとえば、音声メッセージ、音楽ファイルなど)を含む可能性があり、モバイルコンピューティングデバイス950上で動作するアプリケーションによって生成された音声も含む可能性がある。

【0155】

モバイルコンピューティングデバイス950は、図に示されるように、多くの異なる形態で実装される可能性がある。たとえば、モバイルコンピューティングデバイス950は、セルラー電話980として実装される可能性がある。また、モバイルコンピューティングデバイス950は、スマートフォン982、携帯情報端末、またはその他の同様のモバイルデバイスの一部として実装される可能性がある。

【0156】

本明細書に記載のシステムおよび技術の様々な実装は、デジタル電子回路、集積回路、特別に設計されたASIC、コンピュータハードウェア、ファームウェア、ソフトウェア、および/またはこれらの組合せ内に実現され得る。これらの様々な実装は、ストレージシステム、少なくとも1つの入力デバイス、および少なくとも1つの出力デバイスからデータおよび命令を受け取られ、それらにデータおよび命令を送信するために結合された、専用または汎用である可能性がある少なくとも1つのプログラミング可能なプロセッサを含むプログラミング可能なシステム上の、実行可能および/または解釈可能な1つまたは複数のコンピュータプログラムへの実装を含み得る。

【0157】

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、またはコードとしても知られるこれらのコンピュータプログラムは、プログラミング可能なプロセッサ用の機械命令を含み、高級手続き型プログラミング言語および/もしくはオブジェクト指向プログラミ

10

20

30

40

50

ング言語、ならびに/またはアセンブリ/機械言語で実装され得る。プログラムは、その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、マークアップ言語のドキュメントに記憶された1つもしくは複数のスクリプト、問題にしているプログラムに専用の単一のファイル、または複数の組織されたファイル、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、もしくはコードの一部を記憶するファイルに記憶され得る。コンピュータプログラムは、1つのコンピュータ上で、または1つの場所に置かれるか、もしくは複数の場所に分散され、通信ネットワークによって相互に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。

**【0158】**

本明細書で使用されるとき、用語「機械可読媒体」、「コンピュータ可読媒体」とは、機械命令を機械可読信号として受け取る機械可読媒体を含む、プログラミング可能なプロセッサに機械命令および/またはデータを提供するために使用される任意のコンピュータプログラム製品、装置、および/またはデバイス、たとえば、磁気ディスク、光ディスク、メモリ、プログラマブルロジックデバイス(PLD)を指す。用語「機械可読信号」とは、プログラミング可能なプロセッサに機械命令および/またはデータを提供するために使用される任意の信号を指す。

10

**【0159】**

ユーザとのインタラクションを行うために、本明細書に記載のシステムおよび技術は、ユーザに対して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(ブラウン管)またはLCD(液晶ディスプレイ)モニタ、ならびに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を有するコンピュータ上に実装され得る。その他の種類のデバイスが、ユーザとのインタラクションを行うためにやはり使用されることが可能であり、たとえば、ユーザに提供さ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ることが可能であり、ユーザからの入力、音響、発話、または触覚による入力を含む任意の形態で受け取られることが可能である。

20

**【0160】**

本明細書に記載のシステムおよび技術は、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、またはアプリケーションサーバなどのミドルウェア構成要素を含むか、またはユーザが本明細書に記載のシステムおよび技術の実装とインタラクション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などのフロントエンド構成要素を含むか、または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実装され得る。システムの構成要素は、通信ネットワークなどのデジタルデータ通信の任意の形態または媒体によって相互に接続され得る。通信ネットワークの例は、ローカルエリアネットワーク(「LAN」)、広域ネットワーク(「WAN」)、およびインターネットを含む。

30

**【0161】**

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、概して互いに離れており、通常は通信ネットワークを通じてインタラクションする。クライアントとサーバとの関係は、それぞれのコンピュータ上で実行されており、互いにクライアント-サーバの関係にあるコンピュータプログラムによって生じる。

40

**【0162】**

上の説明に加えて、ユーザは、本明細書に記載のシステム、プログラム、または特徴がユーザ情報(たとえば、ユーザのソーシャルネットワーク、社会的行為もしくは活動、職業、ユーザのプリファレンス、またはユーザの現在位置についての情報)の収集を可能にしてもよいかどうかといつ可能にしてもよいかの両方、およびユーザがコンテンツまたは通信をサーバから送信されるかどうかについての選択をユーザが行うことを可能にするコントロールを提供される可能性がある。さらに、特定のデータが、個人を特定できる情報が

50

削除されるように、記憶されるかまたは使用される前に1つまたは複数の方法で処理される可能性がある。

【0163】

たとえば、一部の実施形態においては、ユーザのアイデンティティが、個人を特定できる情報がユーザに関して決定され得ないか、または位置情報が取得される場合にユーザの地理的位置が(都市、郵便番号、もしくは州のレベルまでになど)一般化される可能性があり、したがって、ユーザの特定の位置が決定され得ないように処理される可能性がある。したがって、ユーザは、どの情報がユーザについて収集されるか、その情報がどのように使用されるのか、およびどの情報がユーザに提供されるのかをコントロールすることができる可能性がある。

10

【0164】

いくつかの実施形態が、説明された。しかしながら、本発明の範囲から逸脱することなく様々な修正がなされる可能性があることが理解されるであろう。たとえば、上で示された様々な形式のフローは、ステップの順序を変えるか、ステップを追加するか、またはステップを削除して使用される可能性がある。また、システムおよび方法のいくつかの応用が説明されたが、多くのその他の応用が考えられることを認識されたい。したがって、その他の実施形態は、添付の請求項の範囲内にある。

【0165】

対象の特定の実施形態が、説明された。その他の実施形態は、添付の請求項の範囲内にある。たとえば、請求項に挙げられたアクションは、異なる順序で実行され、それでも所望の結果を達成することができる。一例として、添付の図面に示されたプロセスは、所望の結果を達成するために、必ずしも、示された特定の順序、または逐次的順序である必要はない。場合によっては、マルチタスクおよび並列処理が有利である可能性がある。

20

【符号の説明】

【0166】

- 100 システム
- 110 ユーザ
- 120 発話
- 125 音声対応デバイス
- 132 ユーザアカウント情報の第1の組
- 134 ユーザアカウント情報の第2の組
- 136 ユーザアカウント情報の第3の組
- 140 第1の電話番号
- 142 第2の電話番号
- 155 受話者
- 200 プロセス
- 300 プロセス
- 500 システム
- 502 アシスタントサーバ
- 504 連絡先データベース
- 506 ボイスサーバ
- 508 ネットワーク
- 510 クライアントデバイス
- 512 通信回線
- 514 通信回線
- 516 要求
- 518 アクションメッセージ
- 600 プロセス
- 700 プロセス
- 800 プロセス

30

40

50

900	コンピューティングデバイス	
902	プロセッサ	
904	メモリ	
906	記憶装置	
908	高速インターフェース	
910	高速拡張ポート	
912	低速インターフェース	
914	低速拡張ポート	
916	ディスプレイ	
920	サーバ	10
922	ラップトップコンピュータ	
924	ラックサーバシステム	
950	モバイルコンピューティングデバイス	
952	プロセッサ	
954	ディスプレイ	
956	ディスプレイインターフェース	
958	制御インターフェース	
960	音声コーデック	
962	外部インターフェース	
964	メモリ	20
966	通信インターフェース	
968	トランシーバ	
970	GPS受信機モジュール	
972	拡張インターフェース	
974	拡張メモリ	
980	セルラー電話	
982	スマートフォン	

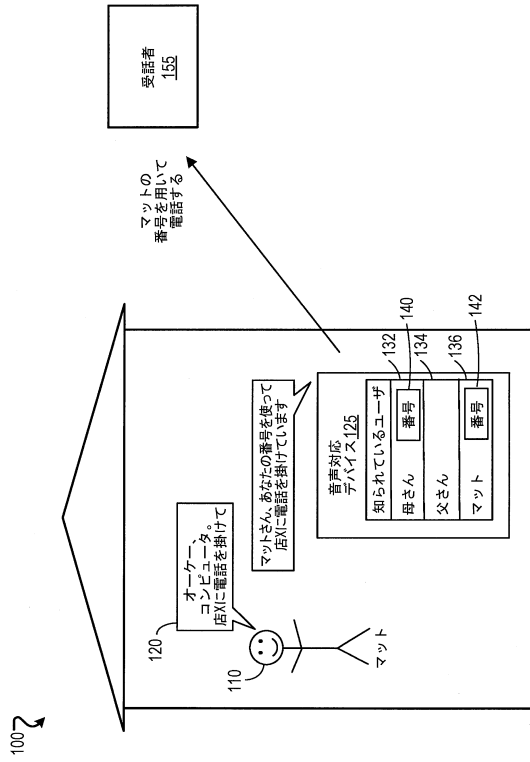
30

40

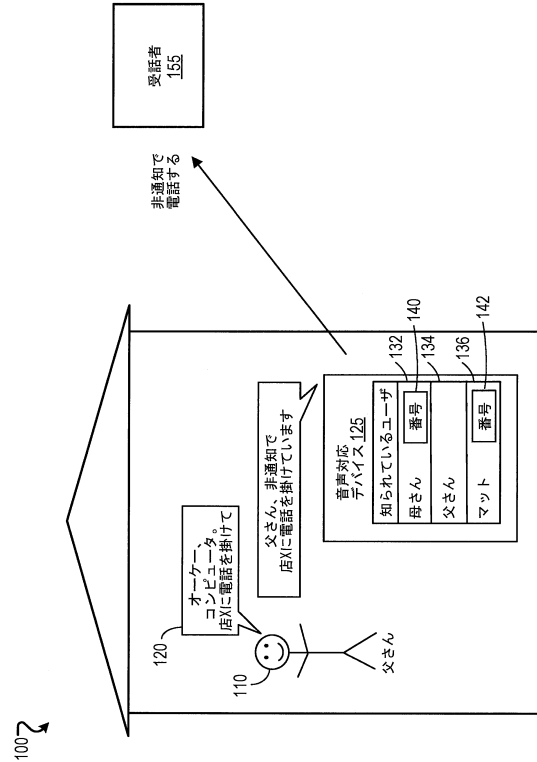
50

【図面】

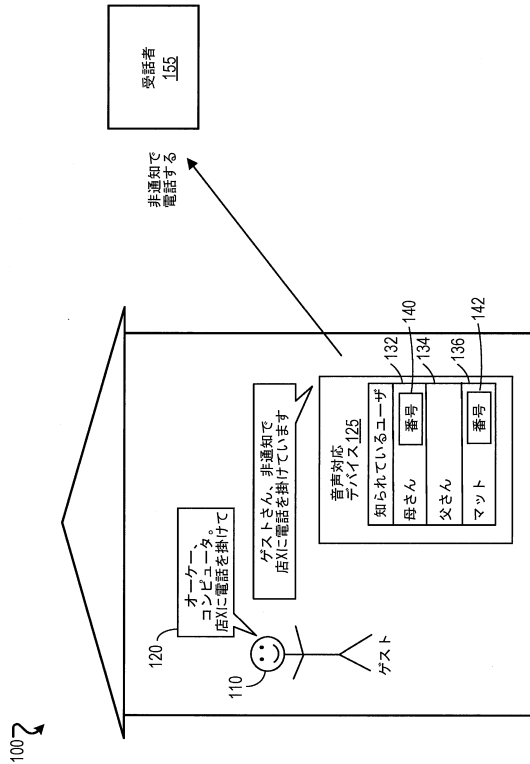
【図 1 A】



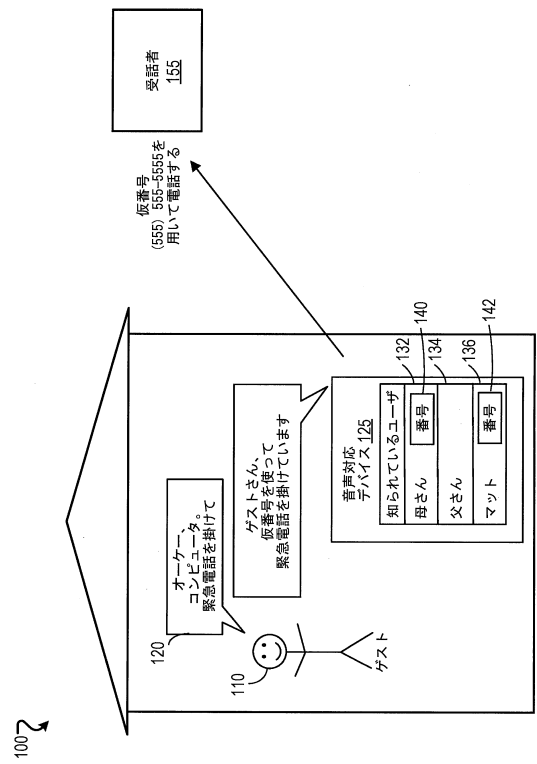
【図 1 B】



【図 1 C】



【図 1 D】



10

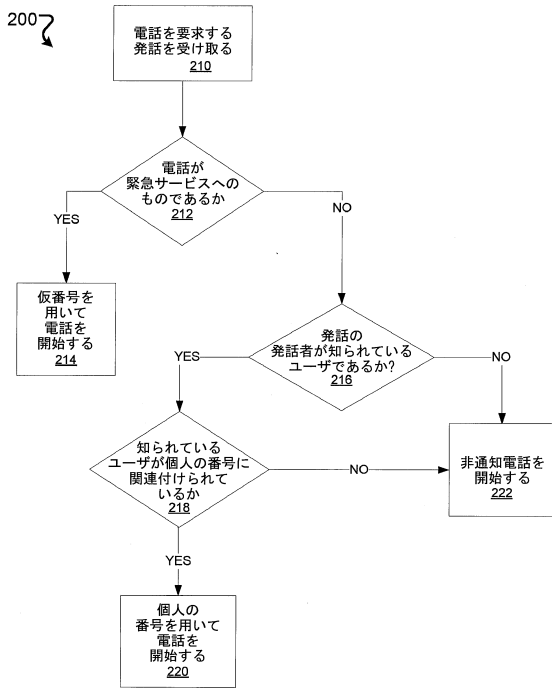
20

30

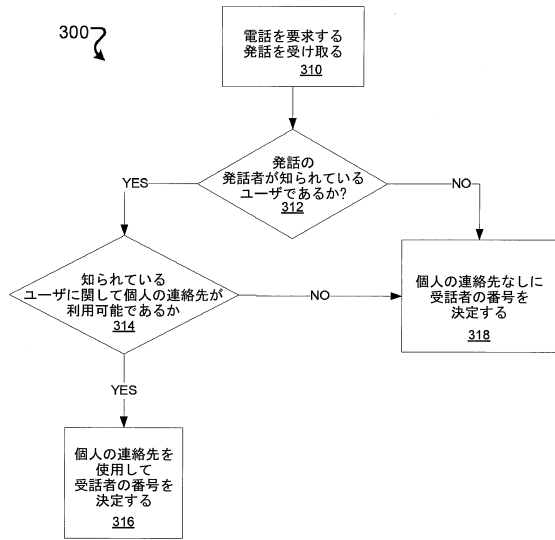
40

50

【図2】



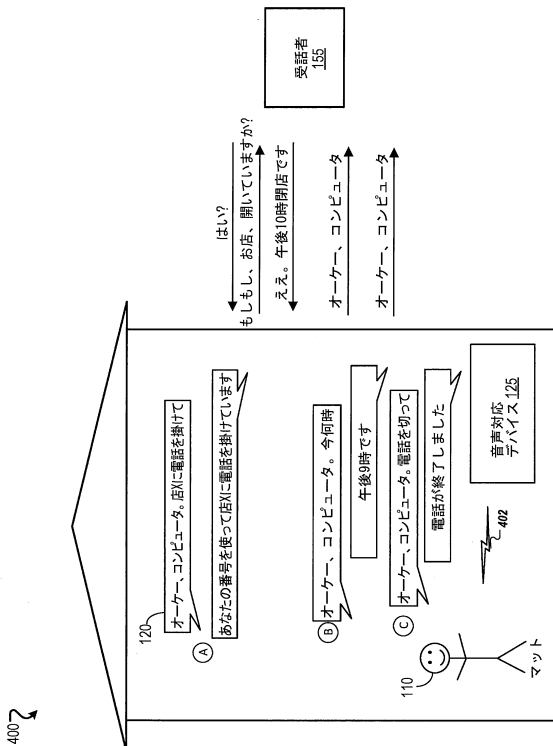
【図3】



10

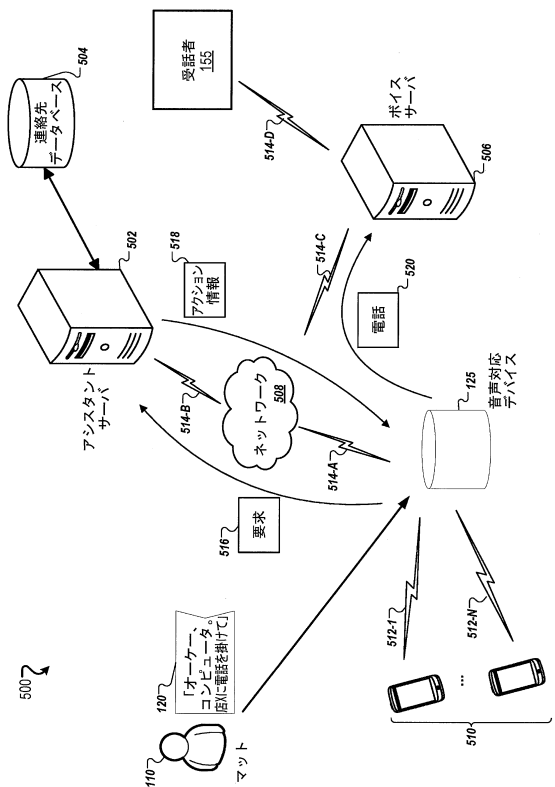
20

【図4】



400

【図5】



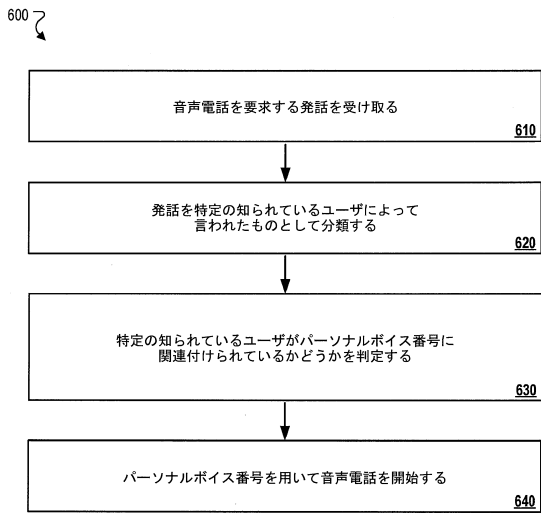
500

30

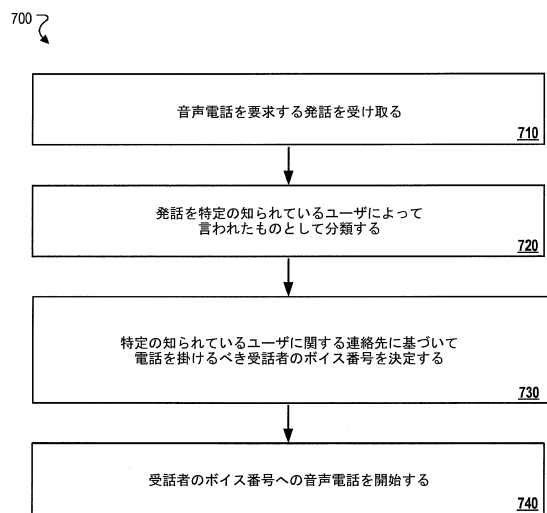
40

50

【図 6】

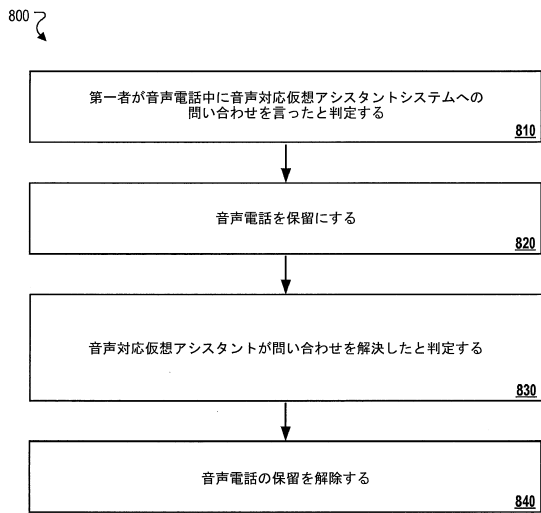


【図 7】

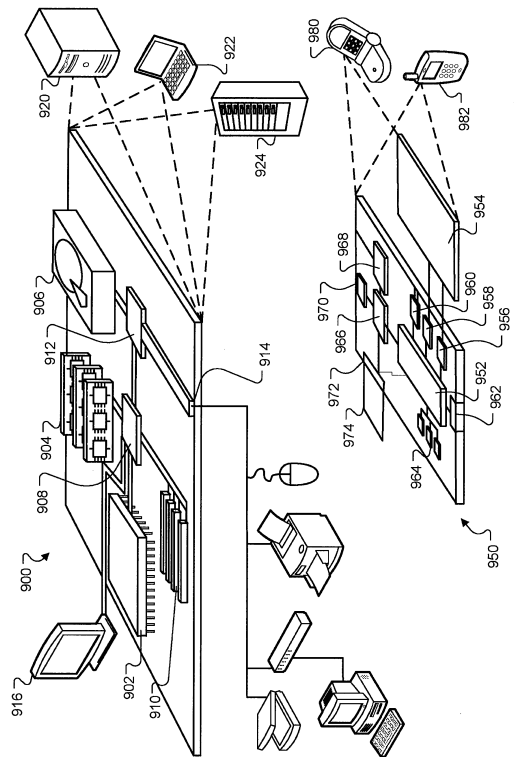


10

【図 8】



【図 9】



20

30

40

50



## フロントページの続き

- (72)発明者 ヴィン・クオック・リ  
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク  
ウェイ・1600
- (72)発明者 ラウナック・シャー  
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク  
ウェイ・1600
- (72)発明者 オカン・コーラック  
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク  
ウェイ・1600
- (72)発明者 デニズ・ピネイ  
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク  
ウェイ・1600
- (72)発明者 ティエンユ・ワン  
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク  
ウェイ・1600
- 審査官 滝谷 亮一
- (56)参考文献 特開2017-034386(JP,A)  
特開2003-218996(JP,A)
- (58)調査した分野 (Int.Cl., DB名)
- |      |       |
|------|-------|
| G06F | 3/16  |
| H04M | 3/42  |
| G10L | 17/00 |
| G10L | 15/00 |