

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2018-205751
(P2018-205751A)

(43) 公開日 平成30年12月27日(2018.12.27)

(5) Int.Cl.	F I	テーマコード (参考)
G 1 0 L 13/06 (2013.01)	G 1 0 L 13/06 1 2 0 Z	
G 1 0 L 19/12 (2013.01)	G 1 0 L 19/12	
G 0 6 F 3/16 (2006.01)	G 0 6 F 3/16 6 9 0	

審査請求 有 請求項の数 34 O L 外国語出願 (全 69 頁)

(21) 出願番号	特願2018-134862 (P2018-134862)	(71) 出願人	507364838
(22) 出願日	平成30年7月18日 (2018. 7. 18)		クアルコム, インコーポレイテッド
(62) 分割の表示	特願2016-564593 (P2016-564593) の分割		アメリカ合衆国 カリフォルニア 9 2 1 2 1 サン ディエゴ モアハウス ドラ イブ 5 7 7 5
原出願日	平成27年4月30日 (2015. 4. 30)	(74) 代理人	100108453
(31) 優先権主張番号	61/986, 701		弁理士 村山 靖彦
(32) 優先日	平成26年4月30日 (2014. 4. 30)	(74) 代理人	100163522
(33) 優先権主張国	米国 (US)		弁理士 黒田 晋平
(31) 優先権主張番号	14/700, 009	(72) 発明者	シャラス・マンジュナト
(32) 優先日	平成27年4月29日 (2015. 4. 29)		アメリカ合衆国・コロラド・80602・ ゾートン・イースト・ワンハンドレッド トゥエンティエイス・プレイス・8241
(33) 優先権主張国	米国 (US)		

最終頁に続く

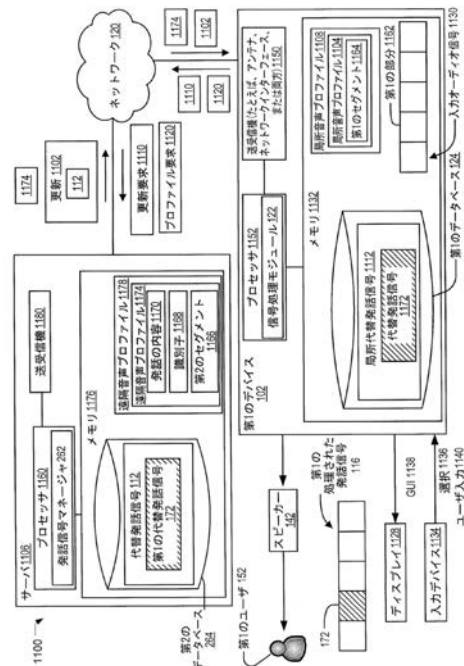
(54) 【発明の名称】 音声プロファイルの管理および発話信号の生成

(57) 【要約】

【課題】 発話信号のオーディオ品質を改善する。

【解決手段】 デバイスは、受信機、メモリ、およびプロセッサを含む。受信機は、遠隔音声プロファイルを受信するように構成される。メモリは、受信機に電氣的に結合される。メモリは、ある人物と関連付けられる局所音声プロファイルを記憶するように構成される。プロセッサは、メモリおよび受信機に電氣的に結合される。プロセッサは、遠隔音声プロファイルと関連付けられる発話の内容または遠隔音声プロファイルと関連付けられる識別子に基づいて、遠隔音声プロファイルがその人物と関連付けられると決定するように構成される。プロセッサはまた、この決定に基づいて、プロフィール管理のために局所音声プロファイルを選択するように構成される。

【選択図】 図 1 1



【特許請求の範囲】**【請求項 1】**

遠隔音声プロファイルを受信するように構成される受信機と、
前記受信機に電氣的に結合されるメモリであって、人物と関連付けられる局所音声プロファイルを記憶するように構成される、メモリと、

前記メモリおよび前記受信機に電氣的に結合されるプロセッサとを備え、前記プロセッサが、

前記遠隔音声プロファイルと関連付けられる発話の内容または前記遠隔音声プロファイルと関連付けられる識別子に基づいて、前記遠隔音声プロファイルが前記人物と関連付けられると決定し、

前記決定に基づいて、プロファイル管理のために前記局所音声プロファイルを選択する

ように構成される、デバイス。

【請求項 2】

前記プロセッサがさらに、前記局所音声プロファイルの第1のセグメントを前記遠隔音声プロファイルの第2のセグメントで置き換えるように構成される、請求項1に記載のデバイス。

【請求項 3】

前記局所音声プロファイルが第1の複数の代替発話信号と関連付けられ、前記遠隔音声プロファイルが第2の複数の代替発話信号と関連付けられ、前記局所音声プロファイルの前記第1のセグメントが、前記第1の複数の代替発話信号から第1の代替発話信号を除去し、前記第2の複数の代替発話信号の第2の代替発話信号を前記第1の複数の代替発話信号に追加することによって、前記遠隔音声プロファイルの前記第2のセグメントで置き換えられる、請求項2に記載のデバイス。

【請求項 4】

前記受信機が、アンテナ、ネットワークインターフェース、または前記アンテナと前記ネットワークインターフェースの組合せの、1つまたは複数を含む、請求項1に記載のデバイス。

【請求項 5】

前記受信機が、複数の人物と関連付けられる複数の遠隔音声プロファイルを受信するように構成される、請求項1に記載のデバイス。

【請求項 6】

前記遠隔音声プロファイルの表現、前記局所音声プロファイルの表現、または両方を含むグラフィカルユーザインターフェースを提示するように構成されるディスプレイをさらに備える、請求項1に記載のデバイス。

【請求項 7】

前記プロセッサがさらに、

テキスト発話変換器から発話信号を受信することと、

複数の使用モードのうちの使用モードを検出することと、

複数の属性領域のうち属性領域を選択することと、

前記発話信号、前記使用モード、および前記属性領域に基づいて、複数の代替発話信号のうち代替発話信号を選択することと、前記局所音声プロファイルが前記複数の代替発話信号と関連付けられる、選択することと、

前記代替発話信号に基づいて、処理された発話信号を生成することとを行うように構成される、請求項1に記載のデバイス。

【請求項 8】

前記複数の代替発話信号が分類と関連付けられ、前記分類の各々がある優先度を有し、前記代替発話信号が、前記発話信号の一部分を優先度の順番で前記複数の代替発話信号と比較することによって選択される、請求項7に記載のデバイス。

【請求項 9】

10

20

30

40

50

前記人物と関連付けられる訓練発話信号を受信するように構成されるマイクロフォンをさらに備え、前記複数の代替発話信号の1つまたは複数の発話信号が前記訓練発話信号に基づいて生成される、請求項7に記載のデバイス。

【請求項10】

前記プロセッサがさらに、

訓練発話信号と関連付けられる前記複数の使用モードの第1の使用モードを検出し、前記第1の使用モードを、前記訓練発話信号に基づいて生成される前記複数の代替発話信号の1つまたは複数の代替発話信号と関連付けるように構成される、請求項7に記載のデバイス。

【請求項11】

前記プロセッサがさらに、

訓練発話信号が受信されるときの使用モードの設定に基づいて、前記複数の使用モードの第1の使用モードを検出し、

前記第1の使用モードを、前記訓練発話信号に基づいて生成される前記複数の代替発話信号の1つまたは複数の発話信号と関連付けるように構成される、請求項7に記載のデバイス。

【請求項12】

前記プロセッサがさらに、

訓練発話信号が歌唱に対応すると決定したことに基づいて、歌唱モードを前記複数の代替発話信号の1つまたは複数の発話信号と関連付けることであって、前記1つまたは複数の発話信号が前記訓練発話信号に基づいて生成される、関連付けることと、

前記訓練発話信号が受信されたときに読書アプリケーションが有効モードにあったと決定したことに基づいて、読書モードを前記1つまたは複数の発話信号と関連付けることと

、
前記訓練発話信号が音声呼の間に捉えられたと決定したことに基づいて、会話モードを前記1つまたは複数の発話信号と関連付けることとを行うように構成される、請求項7に記載のデバイス。

【請求項13】

通信のための方法であって、

局所音声プロファイルを記憶するデバイスにおいて遠隔音声プロファイルを受信するステップであって、前記局所音声プロファイルが人物と関連付けられる、ステップと、

前記遠隔音声プロファイルと前記局所音声プロファイルとの比較に基づいて、または、前記遠隔音声プロファイルと関連付けられる識別子に基づいて、前記遠隔音声プロファイルが前記人物と関連付けられると決定するステップと、

前記デバイスにおいて、前記決定に基づいて、プロファイル管理のために前記局所音声プロファイルを選択するステップとを備える、方法。

【請求項14】

前記デバイスにおいて、前記遠隔音声プロファイルと関連付けられる複数の代替発話信号に対応するデータを含む更新を受信するステップと、

前記複数の代替発話信号を前記局所音声プロファイルと関連付けるステップと、

前記デバイスにおいて、前記局所音声プロファイルと関連付けられる選択を受信するステップと、

前記デバイスにおいて、入力オーディオ信号を受信するステップと、

前記デバイスにおいて、前記入力オーディオ信号の一部分を前記複数の代替発話信号のうちの代替発話信号で置き換えることによって、処理された発話信号を生成するステップとをさらに備える、請求項13に記載の方法。

【請求項15】

前記代替発話信号が、前記入力オーディオ信号の前記一部分よりも高いオーディオ品質を有する、請求項14に記載の方法。

【請求項16】

10

20

30

40

50

スピーカーを介して前記処理された発話信号を出力するステップをさらに備える、請求項14に記載の方法。

【請求項17】

前記選択が、前記デバイスから前記局所音声プロファイルと関連付けられる第2のデバイスへの音声呼の開始に対応する、請求項14に記載の方法。

【請求項18】

前記選択が、第2のデバイスからの音声呼の前記デバイスによる受け入れに対応し、前記第2のデバイスが前記局所音声プロファイルと関連付けられる、請求項14に記載の方法。

【請求項19】

前記デバイスにおいて、前記入力オーディオ信号の前記一部分を前記複数の代替発話信号と比較するステップと、

前記デバイスにおいて、前記入力オーディオ信号の前記一部分が前記代替発話信号と一致すると決定するステップとをさらに備え、

前記決定に基づいて、前記入力オーディオ信号の前記一部分が前記代替発話信号で置き換えられる、請求項14に記載の方法。

【請求項20】

前記更新が前記デバイスから送られる更新要求に応答して受信される、請求項14に記載の方法。

【請求項21】

前記デバイスにおいて、更新要求がいつ送られるべきかを指定するためのオプションを含むグラフィカルユーザインターフェースを生成するステップと、

前記グラフィカルユーザインターフェースをディスプレイデバイスに提供するステップと、

前記オプションに対応するユーザ入力を受け取ったことに基づいて前記更新要求を送るステップとをさらに備え、いつ前記更新要求が送られるべきかを前記ユーザ入力が見し、前記更新要求に応答して前記更新が受信される、請求項14に記載の方法。

【請求項22】

前記デバイスにおいて、第2のオーディオ信号を受信するステップと、

前記第2のオーディオ信号に基づいて、第2の複数の代替発話信号を生成するステップと

、
第2の更新を第2のデバイスに提供するステップとをさらに備え、前記第2の更新が、前記第2の複数の代替発話信号に対応する第2のデータを含む、請求項13に記載の方法。

【請求項23】

前記第2の更新を前記第2のデバイスにいつ送るかを選択するためのオプションを含むグラフィカルユーザインターフェースを生成するステップと、

前記グラフィカルユーザインターフェースをディスプレイデバイスに提供するステップとをさらに備え、

前記第2の更新が、前記オプションに対応するユーザ入力を受け取ったことに基づいて前記第2のデバイスに送られ、

前記ユーザ入力、前記第2の更新が前記第2のデバイスにいつ送られるべきかを示す、請求項22に記載の方法。

【請求項24】

前記第2の複数の代替発話信号が、前記第2のデバイスへ前記第2の更新を送ることを前記デバイスのユーザが認めていると決定したことに基づいて、前記第2のデバイスに送られる、請求項22に記載の方法。

【請求項25】

複数の代替発話信号を記憶するように構成されるメモリと、

前記メモリに電氣的に結合され、

テキスト発話変換器から発話信号を受信することと、

10

20

30

40

50

複数の使用モードのうちの使用モードを検出することと、
 複数の属性領域のうちの属性領域を選択することと、
 前記発話信号、前記属性領域、および前記使用モードに基づいて、前記複数の代替発話信号のうちの代替発話信号を選択することと、
 前記代替発話信号に基づいて、処理された発話信号を生成することと、
 前記処理された発話信号を少なくとも1つのスピーカーに提供することと
 を行うように構成されるプロセッサとを備える、デバイス。

【請求項26】

前記少なくとも1つのスピーカーが車両に含まれる、請求項25に記載のデバイス。

【請求項27】

前記少なくとも1つのスピーカーがモバイルデバイスに含まれる、請求項25に記載のデバイス。

【請求項28】

前記属性領域が、言語領域、性別領域、または年齢領域の少なくとも1つを含み、前記使用モードが、読書モード、会話モード、歌唱モード、または命令および制御モードを含む、請求項25に記載のデバイス。

【請求項29】

遠隔音声プロフィールを受信するための手段と、

人物と関連付けられる局所音声プロフィールを記憶するための手段と、

前記遠隔音声プロフィールと関連付けられる発話の内容または前記遠隔音声プロフィールの識別子に基づいて、前記遠隔音声プロフィールが前記人物と関連付けられると決定したことを基づいて、プロフィール管理のために前記局所音声プロフィールを選択するための手段とを備える、装置。

【請求項30】

受信するための前記手段、記憶するための前記手段、および選択するための前記手段が、車両、電子書籍デバイス、音響キャプチャデバイス、モバイル通信デバイス、スマートフォン、携帯電話、ラップトップコンピュータ、コンピュータ、エンコーダ、デコーダ、タブレット、携帯情報端末、ディスプレイデバイス、テレビジョン、ゲームコンソール、音楽プレーヤ、ラジオ、デジタルビデオプレーヤ、デジタルビデオディスクプレーヤ、チューナー、カメラ、またはナビゲーションデバイスの少なくとも1つへと組み込まれる、請求項29に記載の装置。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本出願は、内容の全体が参照によって明細書に明確に組み込まれる、2014年4月30日に出願された同一出願人が所有する米国仮特許出願第61/986,701号、および2015年4月29日に出願された米国非仮特許出願第14/700,009号の優先権を主張する。

【0002】

本開示は全般に、音声プロフィールの管理および発話信号の生成に関する。

【背景技術】

【0003】

技術の進歩は、より小型でより強力なコンピューティングデバイスをもたらした。たとえば、小型、軽量で、ユーザによって容易に携帯されるポータブルワイヤレス電話、携帯情報端末(PDA)、およびページングデバイスなどのワイヤレスコンピューティングデバイスを含む、種々のポータブルパーソナルコンピューティングデバイスが、現在存在している。より具体的には、携帯電話およびインターネットプロトコル(IP)電話などのポータブルワイヤレス電話は、ワイヤレスネットワークを通じて音声およびデータパケットを通信することができる。さらに、多くのそのようなワイヤレス電話には、内部に他のタイプのデバイスが組み込まれている。たとえば、ワイヤレス電話は、デジタルスチルカメラ、デ

10

20

30

40

50

デジタルビデオカメラ、デジタルレコーダ、およびオーディオファイルプレーヤも含み得る。また、そのようなワイヤレス電話は、ウェブブラウザアプリケーションのような、インターネットにアクセスするために使用され得るソフトウェアアプリケーションを含む、実行可能な命令を処理することができる。したがって、これらのワイヤレス電話は、かなりのコンピューティング能力を含み得る。

【0004】

デジタル技法による音声の送信が、特に長距離のデジタル無線電話の適用例において広まっている。再構築される発話の知覚される品質を保ちながらチャネルを通じて送信されることが可能な、最小の量の情報を決定することについて、関心があり得る。発話がサンプリングおよびデジタル化によって送信される場合、64キロビット毎秒(kbps)のオーダーのデータレートが、アナログ電話の発話品質を達成するために使用され得る。発話分析の使用と、それに続くコーディング、送信、および受信機における再合成とを通じて、データレートの大幅な低減が達成され得る。

10

【0005】

発話を圧縮するためのデバイスが、遠隔通信の多くの分野において使われることがある。例示的な分野はワイヤレス通信である。ワイヤレス通信の分野には、たとえば、コードレス電話、ページング、ワイヤレスローカルループ、セルラー電話システムおよびパーソナル通信サービス(PCS)電話システムのようなワイヤレス電話、モバイルインターネットプロトコル(IP)電話、および衛星通信システムを含む、多くの適用例がある。特定の適用例は、モバイル加入者のためのワイヤレス電話である。

20

【0006】

たとえば、周波数分割多重接続(FDMA)、時分割多重接続(TDMA)、符号分割多重接続(CDMA)、および時分割同期CDMA(TD-SCDMA)を含むワイヤレス通信システムのために、様々なover-the-airインターフェースが開発されている。それに関連して、たとえば、Advanced Mobile Phone Service(AMPS)、Global System for Mobile Communications(GSM(登録商標))、およびInterim Standard 95(IS-95)を含む、様々な国内規格および国際規格が確立されている。例示的なワイヤレス電話通信システムは、符号分割多重接続(CDMA)システムである。IS-95規格およびその派生物、IS-95A、ANSI J-STD-008、およびIS-95B(本明細書ではまとめてIS-95と呼ばれる)が、米国電気通信工業会(TIA)および他のよく知られている標準化団体によって、セルラーまたはPCS電話通信システムのためのCDMA over-the-airインターフェースの使用を規定するために公布されている。

30

【0007】

IS-95規格はその後、cdma2000およびWCDMA(登録商標)のような「3G」システムへと進化し、これはさらなる容量と高速なパケットデータサービスとを提供する。cdma2000の2つの変形が文書IS-2000(cdma2000 1xRTT)およびIS-856(cdma2000 1xEV-DO)によって発表されており、これらはTIAによって発行されている。cdma2000 1xRTT通信システムが153kbpsのピークデータレートを提供する一方、cdma2000 1xEV-DO通信システムは38.4kbpsから2.4Mbpsにわたるデータレートのセットを定義する。WCDMA(登録商標)規格は、第3世代パートナーシッププロジェクト「3GPP」、文書番号3G TS 25.211、3G TS 25.212、3G TS 25.213、および3G TS 25.214において具現化されている。International Mobile Telecommunications Advanced(IMT-Advanced)規格は、「4G」規格を提示している。IMT-Advanced規格は、4Gサービスのピークデータレートを、(たとえば、列車および自動車からの)高モビリティ通信に対しては100メガビット毎秒(Mbit/s)に、および(たとえば、歩行者および静止しているユーザからの)低モビリティ通信に対しては1ギガビット毎秒(Gbit/s)に設定している。

40

【0008】

人の発話生成のモデルに関するパラメータを抽出することによって発話を圧縮するための技法を利用するデバイスは、発話コーダと呼ばれる。発話コーダは、エンコーダおよびデコーダを備え得る。エンコーダは、入来する発話信号を時間のブロックまたは分析フレームへと分割する。時間的な各セグメント(またはフレーム)の時間長は、信号のスペクト

50

ルエンベローブが比較的变化しないままであることが予想され得るのに十分短くなるように選択され得る。たとえば、1つのフレーム長は20ミリ秒であり、これは8キロヘルツ(kHz)のサンプリングレートでは160個のサンプルに対応するが、特定の適用例に対して適切であると見なされる任意のフレーム長またはサンプリングレートが使用され得る。

【0009】

エンコーダは、入来する発話フレームを分析していくつかの関連するパラメータを抽出し、次いで、パラメータをバイナリの表現へと、たとえばビットのセットまたはバイナリデータパケットへと量子化する。データパケットは、通信チャネル(すなわち、有線および/またはワイヤレスネットワーク接続)を通じて受信機およびデコーダに送信される。デコーダは、データパケットを処理し、処理されたデータパケットを逆量子化してパラメータを作成し、逆量子化されたパラメータを使用して発話フレームを再合成する。

10

【0010】

発話コーダの機能は、発話に内在する天然の冗長性を除去することによって、デジタル化された発話信号を低ビットレートの信号へと圧縮することである。デジタル圧縮は、パラメータのセットにより入力発話フレームを表し、量子化を利用してビットのセットによりパラメータを表すことによって達成され得る。入力発話フレームがビットの数 N_i を有し、発話コーダによって作成されるデータパケットがビットの数 N_o を有する場合、発話コーダによって達成される圧縮係数は $C_r = N_i / N_o$ である。課題は、目標の圧縮係数を達成しながら復号される発話の高い信号品質を保つことである。発話コーダの性能は、(1)発話モデル、または上で説明された分析処理および合成処理の組合せがどの程度良好に機能するか、および(2)パラメータ量子化処理がフレームごとに N_o ビットの目標ビットレートにおいてどの程度良好に実行されるかに依存する。したがって、発話モデルの目標は、各フレームに対するパラメータの小さなセットにより、発話信号の本質または目標の音声品質を捉えることである。

20

【0011】

発話コーダは一般に、パラメータのセット(ベクトルを含む)を利用して発話信号を表現する。パラメータの良好なセットは、理想的には、知覚的に正確な発話信号の再構築のために低いシステム帯域幅を提供する。ピッチ、信号出力、スペクトルエンベローブ(またはフォルマント)、振幅、および位相スペクトルが、発話コーディングパラメータの例である。

30

【0012】

発話コーダは時間領域コーダとして実装されてよく、時間領域コーダは、高い時間分解能の処理を利用してある時間において発話の小さなセグメント(たとえば、5ミリ秒(ms)のサブフレーム)を符号化することによって、時間領域の発話波形を捉えることを試みる。各サブフレームに対して、コードブック空間からの高精度の代表が、検索アルゴリズムによって発見される。代替的に、発話コーダは周波数領域コーダとして実装されてよく、周波数領域コーダは、パラメータ(分析)のセットにより入力発話フレームの短期発話スペクトルを捉え、対応する合成処理を利用してスペクトルパラメータから発話波形を再作成することを試みる。パラメータ量子化器は、既知の量子化技法に従ってコードベクトルの記憶された表現によりパラメータを表すことによって、パラメータを維持する。

40

【0013】

1つの時間領域の発話コーダは、符号励振線形予測(CELP)コーダである。CELPコーダでは、発話信号における短期の相関または冗長性は、短期フォルマントフィルタの係数を発見する線形予測(LP)の分析によって除去される。短期予測フィルタを入来する発話フレームに適用することでLP残差信号が生成され、LP残差信号はさらに、長期予測フィルタのパラメータおよびその後の統計的コードブックによってモデル化され量子化される。したがって、CELPコーディングは、時間領域の発話波形を符号化するタスクを、LP短期フィルタ係数を符号化することおよびLP残差を符号化することという別々のタスクに分割する。時間領域のコーディングは、固定されたレートで(すなわち、各フレームに対して、ビットの同じ数 N_o を使用して)、または可変のレートで(この場合異なるビットレートが異なるタ

50

イブのフレーム内容のために使用される)実行され得る。可変レートコードは、目標の品質を得るのに適切なレベルへとコーデックパラメータを符号化するのに必要とされるビットの量を使用することを試みる。

【0014】

CELPコードのような時間領域コードは、時間領域の発話波形の精度を維持するために、フレームごとに N_0 個という多数のビットに依存し得る。そのようなコードは、フレームごとのビットの数 N_0 が比較的大きければ(たとえば、8kbps以上)、素晴らしい音声品質をもたらし得る。低ビットレート(たとえば、4kbps以下)では、時間領域コードは、利用可能なビットの数が限られていることにより、高い品質と安定した性能とを保てないことがある。低ビットレートでは、限られたコードブック空間が、より高レートの商用の適用例において展開される時間領域コードの波形照合能力を制限する。したがって、時間とともに改善されてはいるが、低ビットレートで動作する多くのCELPコーディングシステムには、ノイズとして特徴付けられる知覚的に重大な歪みという課題がある。

10

【0015】

低ビットレートにおけるCELPコードに対する代替は、「ノイズ励振線形予測」(NELP)コードであり、これはCELPコードと同様の原理で動作する。NELPコードは、コードブックではなく、フィルタリングされた疑似ランダムノイズ信号を使用して発話をモデル化する。NELPはコーディングされた発話のためにより簡単なモデルを使用するので、NELPはCELPよりも低いビットレートを達成する。NELPは、無声の発話または静寂を圧縮または表現するために使用され得る。

20

【0016】

2.4kbpsのオーダーのレートで動作するコーディングシステムは、一般に本質的にパラメトリックである。すなわち、そのようなコーディングシステムは、規則的な間隔で発話信号のピッチ周期およびスペクトルエンベロープ(またはフォルマント)を表現するパラメータを送信することによって動作する。これらのいわゆるパラメトリックコードの実例は、LPボコードシステムである。

【0017】

LPボコードは、ピッチ周期ごとに信号パルスを伴う有声の発話信号をモデル化する。この基本的な技法は、とりわけ、スペクトルエンベロープについての送信情報を含むように補強され得る。LPボコードは、一般に妥当な性能を提供するが、バズとして特徴付けられる知覚的な重大な歪みをもたらし得る。

30

【0018】

近年では、波形コードとパラメトリックコードの両方のハイブリッドであるコードが出現している。これらのいわゆるハイブリッドコードの実例は、プロトタイプ波形補間(PWI)発話コーディングシステムである。PWIコーディングシステムはまた、プロトタイプピッチ周期(PPP)発話コードとしても知られていることがある。PWIコーディングシステムは、有声の発話をコーディングするための効率的な方法を提供する。PWIの基本的な概念は、固定された間隔で代表的なピッチサイクル(プロトタイプ波形)を抽出し、その記述を送信し、複数のプロトタイプ波形の間を補間することによって発話信号を再構築することである。PWI方法は、LP残差信号または発話信号のいずれかに対して動作し得る。

40

【0019】

発話信号(たとえば、コーディングされた発話信号、再構築された発話信号、または両方)のオーディオ品質を改善することについて、研究上の関心および商業的な関心があり得る。たとえば、通信デバイスは、理想的な音声品質よりも低品質の発話信号を受信することがある。例示すると、通信デバイスは、音声呼の間に別の通信デバイスから発話信号を受信することがある。音声呼の品質は、環境ノイズ(たとえば、風、道のノイズ)、通信デバイスのインターフェースの制限、通信デバイスによる信号処理、パケットロス、帯域幅制限、ビットレート制限などのような、様々な理由で劣化し得る。

【発明の概要】

【課題を解決するための手段】

50

【0020】

ある特定の態様では、デバイスは、受信機、メモリ、およびプロセッサを含む。受信機は、遠隔音声プロファイルを受信するように構成される。メモリは、受信機に電氣的に結合される。メモリは、ある人物と関連付けられる局所音声プロファイルを記憶するように構成される。プロセッサは、メモリおよび受信機に電氣的に結合される。プロセッサは、遠隔音声プロファイルと関連付けられる発話の内容または遠隔音声プロファイルと関連付けられる識別子に基づいて、遠隔音声プロファイルがその人物と関連付けられると決定するように構成される。プロセッサはまた、この決定に基づいて、プロファイル管理のために局所音声プロファイルを選択するように構成される。

【0021】

別の態様では、通信のための方法は、局所音声プロファイルを記憶するデバイスにおいて遠隔音声プロファイルを受信するステップを含み、その局所音声プロファイルはある人物と関連付けられる。方法はまた、遠隔音声プロファイルと局所音声プロファイルとの比較に基づいて、または、遠隔音声プロファイルと関連付けられる識別子に基づいて、遠隔音声プロファイルがその人物と関連付けられると決定するステップを含む。方法はさらに、この決定に基づいて、デバイスにおいて、プロファイル管理のために局所音声プロファイルを選択するステップを含む。

【0022】

別の態様では、デバイスは、メモリおよびプロセッサを含む。メモリは、複数の代替発話信号を記憶するように構成される。プロセッサは、メモリに電氣的に結合される。プロセッサは、テキスト発話変換器から発話信号を受信するように構成される。プロセッサはまた、複数の使用モードのうちの使用モードを検出するように構成される。プロセッサはさらに、複数の属性領域のうちの属性領域を選択するように構成される。プロセッサはまた、発話信号、属性領域、および使用モードに基づいて、複数の代替発話信号のうちの代替発話信号を選択するように構成される。プロセッサはさらに、代替発話信号に基づいて、処理された発話信号を生成するように構成される。プロセッサはまた、処理された発話信号を少なくとも1つのスピーカに提供するように構成される。

【0023】

本開示の他の態様、利点、および特徴は、以下のセクション、すなわち、図面の簡単な説明、発明を実施するための形態、および特許請求の範囲を含む本出願の検討後に明らかになるであろう。

【図面の簡単な説明】

【0024】

【図1】発話信号を置き換えるように動作可能なシステムの特定の例示的な態様のブロック図である。

【図2】発話信号を置き換えるように動作可能なシステムの別の例示的な態様の図である。

【図3】発話信号を置き換えるように動作可能なシステムによって表示され得るユーザインターフェースの特定の例示的な態様の図である。

【図4】発話信号を置き換えるように動作可能なシステムの別の例示的な態様の図である。

【図5】発話信号を置き換えるように動作可能なシステムの別の例示的な態様の図である。

【図6】発話信号を置き換えるように動作可能なシステムの別の例示的な態様の図である。

【図7】発話信号を置き換えるように動作可能なシステムの別の例示的な態様の図である。

【図8】発話信号を置き換えるように動作可能なシステムの別の例示的な態様の図である。

【図9】発話信号を置き換えるように動作可能なシステムの別の例示的な態様の図である。

10

20

30

40

50

。

【図10】発話信号を置き換えるように動作可能なシステムによって使用され得るデータベースの例示的な態様の図である。

【図11】発話信号を置き換えるように動作可能なシステムの別の例示的な態様の図である。

【図12】発話信号を置き換えるように動作可能なシステムの別の例示的な態様の図である。

【図13】発話信号を置き換えるように動作可能なシステムの別の例示的な態様の図である。

【図14】発話信号を置き換える方法の特定の態様を示すフローチャートである。

10

【図15】図11の動作802に対応し得る、複数の代替発話信号を取得する方法の特定の態様を示すフローチャートである。

【図16】発話信号を置き換える方法の別の態様を示すフローチャートである。

【図17】ユーザインターフェースを生成する方法の別の態様を示すフローチャートである。

【図18】発話信号を置き換える方法の別の態様を示すフローチャートである。

【図19】図1～図18のシステムおよび方法による、発話信号を置き換えるように動作可能なデバイスの特定の例示的な態様のブロック図である。

【発明を実施するための形態】

【0025】

20

本明細書で説明される原理は、たとえば、ヘッドセット、ハンドセット、または発話信号の置換を実行するように構成される他のオーディオデバイスに適用され得る。文脈により明示的に限定されない限り、「信号」という用語は、ワイヤ、バス、または他の送信媒体上で表現されるようなメモリ位置(またはメモリ位置のセット)の状態を含む、その通常の意味のいずれをも指すために本明細書において使用される。文脈により明示的に限定されない限り、「生成する」という用語は、計算することまたは別様に作成することのような、その通常の意味のいずれをも指すために本明細書において使用される。文脈により明示的に限定されない限り、「算出する」という用語は、計算すること、評価すること、平滑化すること、および/または複数の値から選択することのような、その通常の意味のいずれをも指すために本明細書において使用される。文脈によって明示的に限定されない限り、「取得する」という用語は、算出すること、導出すること、(たとえば、別のコンポーネント、ブロック、またはデバイスから)受信すること、および/または(たとえば、メモリレジスタまたは記憶素子のアレイから)取り出すことのような、その通常の意味のいずれをも指すために使用される。

30

【0026】

文脈により明示的に限定されない限り、「作成する」という用語は、算出すること、生成すること、および/または提供することのような、その通常の意味のいずれをも指すために使用される。文脈により明示的に限定されない限り、「提供する」という用語は、算出すること、生成すること、および/または作成することのような、その通常の意味のいずれをも指すために使用される。文脈によって明示的に限定されない限り、「結合される」という用語は、直接的または間接的な、電氣的または物理的な接続を指すために使用される。接続が間接的である場合、「結合されている」構造物の間に他のブロックまたはコンポーネントがあり得ることが、当業者によりよく理解される。

40

【0027】

「構成」という用語は、特定の文脈によって示されるような、方法、装置/デバイス、および/またはシステムに言及する際に使用され得る。「備える」という用語がこの説明および請求項において使用される場合、それは他の要素または動作を除外しない。(「AはBに基づく」のような)「に基づく」という用語は、(i)「に少なくとも基づく」(たとえば、「Aは少なくともBに基づく」)という場合と、特定の文脈において適切であれば、(ii)「に等しい」(たとえば、「AはBに等しい」)という場合とを含む、その通常の意味のい

50

れをも指すために使用される。「AがBに基づく」が「AがBに少なくとも基づく」を含む場合(i)において、これは、AがBに結合される構成を含み得る。同様に、「に応答して」という用語は、「に少なくとも応答して」を含む、その通常の意味のいずれをも指すために使用される。「少なくとも1つの」という用語は、「1つまたは複数」を含む、その通常の意味のいずれをも指すために使用される。「少なくとも2つの」という用語は、「2つ以上」を含む、その通常の意味のいずれをも指すために使用される。

【0028】

「装置」および「デバイス」という用語は、特定の文脈によって別段示されない限り、汎用的および交換可能に使用される。別段示されない限り、特定の特徴を有する装置の動作の任意の開示が、類似の特徴を有する方法を開示すること(およびその逆)も明示的に意図されており、特定の構成に従った装置の動作の任意の開示が、類似の構成に従った方法を開示すること(およびその逆)も明示的に意図されている。「方法」、「処理」、「手順」および「技法」という用語は、特定の文脈によって別段示されない限り、汎用的および交換可能に使用される。「要素」および「モジュール」という用語は、より大きな構成の一部を指すために使用され得る。文書の一部の参照によるあらゆる組込みはまた、その部分において参照される用語または変数の定義を組み込むものとして理解されるものとし、ここでそのような定義は、文書中のあらゆる箇所に、さらには組み込まれる部分において参照されるあらゆる図面に現れる。

10

【0029】

本明細書で使用される場合、「通信デバイス」という用語は、ワイヤレス通信ネットワークを通じた音声および/またはデータ通信に使用され得る電子デバイスを指す。通信デバイスの例は、携帯電話、携帯情報端末(PDA)、ハンドヘルドデバイス、ヘッドセット、ワイヤレスモデム、ラップトップコンピュータ、パーソナルコンピュータなどを含む。

20

【0030】

図1を参照すると、発話信号を置き換えるように動作可能なシステムの特定の例示的な態様が開示され、全体が100と指定される。システム100は、ネットワーク120を介して1つまたは複数の他のデバイス(たとえば、モバイルデバイス104、デスクトップコンピュータ106など)と通信している第1のデバイス102を含み得る。モバイルデバイス104は、マイクロフォン146に結合されてよく、またはそれと通信してよい。デスクトップコンピュータ106は、マイクロフォン144に結合されてよく、またはそれと通信してよい。第1のデバイス102は、1つまたは複数のスピーカ142に結合され得る。第1のデバイス102は、信号処理モジュール122と、第1のデータベース124とを含み得る。第1のデータベース124は、代替発話信号112を記憶するように構成され得る。

30

【0031】

第1のデバイス102は、図1に示されるものよりも少数または多数のコンポーネントを含み得る。たとえば、第1のデバイス102は、1つまたは複数のプロセッサ、1つまたは複数のメモリユニット、または両方を含み得る。第1のデバイス102は、ネットワーク化されたコンピューティングシステムまたは分散されたコンピューティングシステムを含み得る。特定の例示的な態様では、第1のデバイス102は、モバイル通信デバイス、スマートフォン、携帯電話、ラップトップコンピュータ、コンピュータ、タブレット、携帯情報端末、ディスプレイデバイス、テレビジョン、ゲームコンソール、音楽プレーヤ、ラジオ、デジタルビデオプレーヤ、デジタルビデオディスク(DVD)プレーヤ、チューナー、カメラ、ナビゲーションデバイス、またはこれらの組合せを含み得る。そのようなデバイスは、ユーザインターフェース(たとえば、タッチスクリーン、音声認識能力、または他のユーザインターフェース能力)を含み得る。

40

【0032】

動作の間、第1のデバイス102は、ユーザ発話信号(たとえば、第1のユーザ発話信号130、第2のユーザ発話信号132、または両方)を受信し得る。たとえば、第1のユーザ152は、第2のユーザ154との音声呼に参与していることがある。音声呼のために第1のユーザ152は第1のデバイス102を使用することがあり、第2のユーザ154はモバイルデバイス104を使用

50

することがある。音声呼の間、第2のユーザ154は、モバイルデバイス104に結合されるマイクロフォン146に向かって話し得る。第1のユーザ発話信号130は、第2のユーザ154によって話される複数の言葉、1つの言葉、または1つの言葉の一部に対応し得る。モバイルデバイス104は、第2のユーザ154から、マイクロフォン146を介して第1のユーザ発話信号130を受け取り得る。ある特定の態様では、マイクロフォン146はオーディオ信号を捕捉することができ、アナログデジタルコンバータ(ADC)が、捕捉されたオーディオ信号を、アナログ波形から、デジタルオーディオサンプルからなるデジタル波形へと変換することができる。デジタルオーディオサンプルは、デジタル信号プロセッサによって処理され得る。利得調整器は、オーディオ信号(たとえば、アナログ波形またはデジタル波形)の振幅レベルを増大または減少させることによって、(たとえば、アナログ波形またはデジタル波形)の利得を調整することができる。利得調整器は、アナログ領域またはデジタル領域のいずれかで動作し得る。たとえば、利得調整器は、デジタル領域において動作することができ、アナログデジタルコンバータによって作成されるデジタルオーディオサンプルを調整することができる。利得調整の後で、エコーキャンセラが、マイクロフォン146に入るスピーカーの出力によって生み出された可能性のある、あらゆるエコーを低減し得る。デジタルオーディオサンプルは、ボコーダ(音声エンコーダ-デコーダ)によって「圧縮」され得る。エコーキャンセラの出力は、たとえば、フィルタ、ノイズプロセッサ、レートコンバータなどの、ボコーダ前処理ブロックに結合され得る。ボコーダのエンコーダは、デジタルオーディオサンプルを圧縮し、送信パケット(デジタルオーディオサンプルの圧縮されたビットの表現)を形成することができる。送信パケットは、モバイルデバイス104のプロセッサと共有され得るメモリに記憶され得る。プロセッサは、デジタル信号プロセッサと通信している制御プロセッサであり得る。

10

20

【0033】

モバイルデバイス104は、ネットワーク120を介して、第1のユーザ発話信号130を第1のデバイス102に送信することができる。たとえば、モバイルデバイス104は送受信機を含み得る。送受信機は、ある形態の送信パケットを変調し(他の情報が送信パケットに付加され得る)、変調された情報を無線を通じてアンテナを介して送ることができる。

【0034】

第1のデバイス102の信号処理モジュール122は、第1のユーザ発話信号130を受信し得る。たとえば、第1のデバイス102のアンテナは、送信パケットを備えるある形態の到来パケットを受信し得る。送信パケットは、第1のデバイス102においてボコーダのデコーダによって「解凍」され得る。解凍された波形は、再構築されたオーディオサンプルと呼ばれ得る。再構築されたオーディオサンプルは、ボコーダ後処理ブロックによって後処理されてよく、エコーを除去するためにエコーキャンセラによって使用されてよい。明確にするために、ボコーダのデコーダおよびボコーダ後処理ブロックは、ボコーダのデコーダモジュールと呼ばれ得る。いくつかの構成では、エコーキャンセラの出力は、信号処理モジュール122によって処理され得る。代替的に、他の構成では、ボコーダのデコーダモジュールの出力は、信号処理モジュール122によって処理され得る。

30

【0035】

第1のデバイス102によって受信されるような第1のユーザ発話信号130のオーディオ品質は、第2のユーザ154が送った第1のユーザ発話信号130のオーディオ品質より低いことがある。第1のユーザ発話信号130のオーディオ品質は、様々な理由で第1のデバイス102への送信の間に劣化することがある。たとえば、第2のユーザ154は、音声呼の間に騒がしい位置(たとえば、交通の多い通り、コンサートなど)にいたことがある。マイクロフォン146によって受信される第1のユーザ発話信号130は、第2のユーザ154によって話される言葉に加えて音を含み得る。別の例として、マイクロフォン146は、音を捉えるための能力が限られていることがあり、モバイルデバイス104は、一部の音情報を失う方式で音を処理することがあり、第1のユーザ発話信号130に対応するパケットの送信の間にネットワーク120内でパケットロスが存在することがあり、またはこれらの任意の組合せであり得る。

40

【0036】

50

信号処理モジュール122は、発話信号(たとえば、第1のユーザ発話信号130、第2のユーザ発話信号132、またはこれらの両方)の一部分を代替発話信号112の1つまたは複数と比較することができる。たとえば、第1のユーザ発話信号130は、第2のユーザ154によって話される言葉(たとえば、「cat」)に対応することがあり、第1のユーザ発話信号130の第1の部分162は、その言葉の一部分(たとえば、「c」、「a」、「t」、「ca」、「at」、または「cat」)に対応することがある。信号処理モジュール122は、比較に基づいて、第1の部分162が代替発話信号112の第1の代替発話信号172と一致すると決定することができる。たとえば、第1の部分162および第1の代替発話信号172は、共通の音(たとえば、「c」)を表し得る。共通の音は、音素、ダイフオン、トライフオン、音節、言葉、またはこれらの組合せを含み得る。

10

【0037】

ある特定の態様では、信号処理モジュール122は、第1の部分162および第1の代替発話信号172がサンプルの有限セットのような共通の音を表さないときであっても、第1の部分162が第1の代替発話信号172と一致すると決定することができる。第1の部分162は、サンプルの第1のセット(たとえば、「do」という言葉)に対応することがあり、第1の代替発話信号172は、サンプルの第2のセット(たとえば、「to」という言葉)に対応することがある。たとえば、信号処理モジュール122は、代替発話信号112の他の代替発話信号と比較して、第1の部分162が第1の代替発話信号172に高い類似性を有すると決定することができる。

【0038】

第1の代替発話信号172は、第1の部分162よりも高いオーディオ品質(たとえば、より高い信号対雑音比の値)を有し得る。たとえば、代替発話信号112は、図2を参照してさらに説明されるように、以前に記録された言葉、または第2のユーザ154によって話される言葉の一部に対応し得る。

20

【0039】

特定の態様では、信号処理モジュール122は、第1の部分162に対応する波形を第1の代替発話信号172に対応する別の波形と比較することによって、第1の部分162が第1の代替発話信号172と一致すると決定することができる。ある特定の態様では、信号処理モジュール122は、第1の部分162に対応する特徴のサブセットを、第1の代替発話信号172に対応する特徴の別のサブセットと比較することができる。たとえば、信号処理モジュール122は、第1の部分162に対応する複数の発話パラメータを第1の代替発話信号172に対応する別の複数の発話パラメータと比較することによって、第1の部分162が第1の代替発話信号172と一致すると決定することができる。複数の発話パラメータは、ピッチパラメータ、エネルギーパラメータ、線形予測コーディング(LPC)パラメータ、メル周波数ケプストラル係数(MFCC)、線スペクトル対(LSP)、線スペクトル周波数(LSF)、ケプストラル、線スペクトル情報(LSI)、離散コサイン変換(DCT)パラメータ(たとえば、係数)、離散フーリエ変換(DFT)パラメータ、高速フーリエ変換(FFT)パラメータ、フォルマント周波数、またはこれらの任意の組合せを含み得る。信号処理モジュール122は、ベクトル量子化器、隠れマルコフモデル(HMM)、または混合ガウスモデル(GMM)の少なくとも1つを使用して、複数の発話パラメータの値を決定することができる。

30

【0040】

信号処理モジュール122は、第1のユーザ発話信号130の第1の部分162を第1の代替発話信号172で置き換えることによって、処理された発話信号(たとえば、第1の処理された発話信号116)を生成することができる。たとえば、信号処理モジュール122は、第1のユーザ発話信号130を複製することができ、複製されたユーザ発話信号から第1の部分162を除去することができる。信号処理モジュール122は、第1の代替発話信号172と複製されたユーザ発話信号とを、第1の部分162が除去された複製されたユーザ発話信号の位置において連結する(たとえば、継ぎ合わせる)ことによって、第1の処理された発話信号116を生成することができる。

40

【0041】

ある特定の態様では、第1の処理された発話信号116は、複数の代替発話信号を含み得る

50

。ある特定の態様では、信号処理モジュール122は、第1の代替発話信号172と別の代替発話信号とを継ぎ合わせて、中間の代替発話信号を生成することができる。この態様では、信号処理モジュール122は、中間の代替発話信号と複製されたユーザ発話信号とを、第1の部分162が除去された位置において継ぎ合わせることができる。代替的な態様では、第1の処理された発話信号116は繰り返し生成され得る。第1の繰り返しの間、信号処理モジュール122は、他の代替発話信号と複製されたユーザ発話信号とを、複製されたユーザ発話信号から第2の部分162が除去される位置において継ぎ合わせることによって、中間の処理された発話信号を生成することができる。中間の処理された発話信号は、後続の繰り返しの間の第1のユーザ発話信号130に対応し得る。たとえば、信号処理モジュール122は、中間の処理された発話信号から第1の部分162を除去することができる。信号処理モジュール122は、第1の代替発話信号172と中間の処理された発話信号とを、第1の部分162が除去された位置において継ぎ合わせることによって、第1の処理された発話信号116を生成することができる。

10

20

30

40

50

【0042】

そのような継ぎ合わせ動作および複製動作は、オーディオ信号のデジタル表現に対してプロセッサによって実行され得る。ある特定の態様では、第1の処理された発話信号116は、利得調整器によって増幅または抑制され得る。第1のデバイス102は、第1のユーザ152へ、スピーカー142を介して第1の処理された発話信号116を出力し得る。たとえば、利得調整器の出力は、デジタルアナログコンバータによってデジタル信号からアナログ信号に変換され、スピーカー142を介して再生され得る。

【0043】

ある特定の態様では、モバイルデバイス104は第1のユーザ発話信号130を第1のデバイス102に送る前に、第1のユーザ発話信号130と関連付けられる1つまたは複数の発話パラメータを修正することができる。ある特定の態様では、モバイルデバイス104は信号処理モジュール122を含み得る。この態様では、1つまたは複数の発話パラメータは、モバイルデバイス104の信号処理モジュール122によって修正され得る。ある特定の態様では、モバイルデバイス104は代替発話信号112へのアクセス権を有し得る。たとえば、モバイルデバイス104は、図2を参照してさらに説明されるように、サーバから代替発話信号112を以前に受信していることがある。

【0044】

モバイルデバイス104は、信号処理モジュール122が第1のユーザ発話信号130の第1の部分162を代替発話信号112と比較するとき一致を見出すのを助けるために、1つまたは複数の発話パラメータを修正することができる。たとえば、モバイルデバイス104は、第1のユーザ発話信号130(たとえば、「cat」)の第1の部分162(たとえば、「c」)に対して、第2の代替発話信号174(たとえば、「cola」からの「c」に対応する)よりも第1の代替発話信号172(たとえば、「catatonic」からの「c」に対応する)がより一致していると決定することができる。モバイルデバイス104は、第1の代替発話信号172が第2の代替発話信号174よりも一致していることを信号処理モジュール122が決定するのを助けるために、第1の部分162の1つまたは複数の発話パラメータを修正することができる。たとえば、モバイルデバイス104は、修正されたパラメータが第2の代替発話信号174の対応するパラメータよりも第1の代替発話信号172の対応するパラメータに近いように、ピッチパラメータ、エネルギーパラメータ、線形予測コーディング(LPC)パラメータ、またはこれらの組合せを修正することができる。

【0045】

ある特定の態様では、モバイルデバイス104は、1つまたは複数の送信パラメータを第1のデバイス102に送ることができる。ある特定の態様では、モバイルデバイス104の信号処理モジュール122は、送信パラメータを第1のデバイス102に送ることができる。送信パラメータは、第1の部分162に一致するものとしてモバイルデバイス104が決定した、特定の代替発話信号(たとえば、第1の代替発話信号172)を特定することができる。別の例として、送信パラメータは、ベクトル量子化器テーブルにおいて特定のベクトル量子化器を第1

のデバイス102が見つけるのを助けるための、特定のベクトル量子化器テーブルエントリーインデックス番号を含み得る。

【0046】

修正された発話パラメータ、送信パラメータ、またはこれらの両方が、第1の部分162と一致する代替発話信号を信号処理モジュール122が選択するのを助けることができる。処理された発話信号(たとえば、第1の処理された発話信号116、第2の処理された発話信号118)を生成する効率も向上し得る。たとえば、信号処理モジュール122は、送信パラメータが特定の代替発話信号(たとえば、第1の代替発話信号172)を特定するとき、第1の部分162を代替発話信号112の各々と比較しなくてよく、これは、第1の処理された発話信号116を生成する際の効率の向上をもたらす。

10

【0047】

ある特定の態様では、第1のデバイス102信号処理モジュール122は、言葉予測アルゴリズムを使用して、第1の部分162を含み得る言葉を予測することができる。たとえば、言葉予測アルゴリズムは、第1のユーザ発話信号130に先行した言葉に基づいて、その言葉を予測することができる。例示すると、言葉予測アルゴリズムは、その言葉および先行する言葉に関連する、主語-動詞の呼応、言葉の関係、文法規則、またはこれらの組合せに基づいて、その言葉を決定することができる。別の例として、言葉予測アルゴリズムは、先行する言葉におけるその言葉の頻度に基づいてその言葉を決定することができ、言葉は第1の部分162に対応する音を含む。例示すると、特定の言葉が特定の会話において頻繁に繰り返されることがある。信号処理モジュール122は、先行する言葉が閾値の回数(たとえば、2)よりも多い「cat」という言葉を含むことに基づいて、「c」に対応する第1の部分162が「cat」という言葉の一部であると予測することができる。信号処理モジュール122は、予測される言葉に基づいて第1の代替発話信号172を選択することができる。たとえば、第1の代替発話信号172、代替発話信号112の別の代替発話信号、および第1の部分162はすべて、共通の音(たとえば、「c」)に対応し得る。信号処理モジュール122は、第1の代替発話信号172が「catatonic」から生成されたことと、他の代替発話信号が「cola」から生成されたことを決定することができる。信号処理モジュール122は、第1の代替発話信号172が他の代替発話信号よりも予測される言葉に近いと決定したことに基づいて、他の代替発話信号ではなく第1の代替発話信号172を選択することができる。

20

【0048】

ある特定の態様では、信号処理モジュール122は、処理された発話信号(たとえば、第1の処理された発話信号116または第2の処理された発話信号118)が発話信号の閾値を満たさないと決定することがある。たとえば、信号処理モジュール122は、第1の処理された発話信号116の発話パラメータの変動が閾値の発話パラメータの変動を満たさないと決定することがある。例示すると、発話パラメータの変動は、第1の代替発話信号172と第1の処理された発話信号116の別の部分との間の遷移に対応し得る。ある特定の態様では、信号処理モジュール122は、発話パラメータの変動を減らすために、平滑化アルゴリズムを使用して第1の処理された発話信号116を修正することができる。別の特定の態様では、信号処理モジュール122は、発話パラメータの変動が閾値の発話パラメータを満たさないことに応答して、第1の処理された発話信号116を廃棄することができる。この態様では、信号処理モジュール122は、第1のユーザ152へ、スピーカ-142を介して第1のユーザ発話信号130を出力し得る。

30

40

【0049】

ある特定の態様では、第1のデバイス102は、ユーザ入力を受け取ったことに基づいて、処理された発話信号(たとえば、第1の処理された発話信号116、第2の処理された発話信号118)を出力することと、ユーザ発話信号(たとえば、第1のユーザ発話信号130、第2のユーザ発話信号132)を出力することとを切り替えることができる。たとえば、第1のデバイス102は、発話信号の置換が有効にされるべきであることを示す第1の入力を第1のユーザ152から受け取り得る。第1の入力に応答して、第1のデバイス102は、スピーカ-142を介して第1の処理された発話信号116を出力することができる。別の例として、第1のデバイス102

50

は、発話信号の置換が無効にされるべきであることを示す第2の入力を第1のユーザ152から受け取り得る。第2の入力に回答して、第1のデバイス102は、スピーカ142を介して第1のユーザ発話信号130を出力することができる。ある特定の態様では、信号処理モジュール122は、ユーザ(たとえば、第1のユーザ152)が代替発話信号112を管理することを可能にするための、ユーザインターフェースを含み得る。この態様では、第1のデバイス102は、信号処理モジュール122のユーザインターフェースを介して、第1の入力、第2の入力、または両方を受け取り得る。

【0050】

ある特定の態様では、信号処理モジュール122は、音声呼、たとえば第1のユーザ152と第2のユーザ154との間の呼の間に、第1のユーザ発話信号130を受け取ることがあり、音声呼の間に第1の処理された発話信号116を生成することがある。別の特定の態様では、信号処理モジュール122は、第1のユーザ発話信号130を(たとえば、第2のユーザ154からのメッセージとして)受け取ることがあり、続いて第1の処理された発話信号116を生成することがあり、第1の処理された発話信号116を後で第1のユーザ152に対して再生するために記憶することがある。

10

【0051】

ある特定の態様では、信号処理モジュール122は、どのデバイス(たとえば、デスクトップコンピュータ106、モバイルデバイス104)がユーザ発話信号を送るかとは無関係に、特定のユーザ(たとえば、第2のユーザ154)に対応するユーザ発話信号130を(たとえば、第1のユーザ発話信号130、第2のユーザ発話信号132)の一部分を置き換えるために、代替発話信号112を使用することができる。たとえば、信号処理モジュール122は、マイクロフォン144、デスクトップコンピュータ106、およびネットワーク120を介して、第2のユーザ154から第2のユーザ発話信号132を受け取り得る。信号処理モジュール122は、第2のユーザ発話信号132の第2の部分164を代替発話信号112の第2の代替発話信号174で置き換えることによって、第2の処理された発話信号118を生成することができる。

20

【0052】

ある特定の態様では、信号処理モジュール122は、ユーザ発話信号(たとえば、第1のユーザ発話信号130、第2のユーザ発話信号132)に基づいて、修正された代替発話信号(たとえば、修正された代替発話信号176)を生成することができる。たとえば、第1のユーザ発話信号130の第1の部分162および第1の代替発話信号172は、共通の音に対応してよく、第2のユーザ154により使用される異なるトーンまたは抑揚に各々対応してよい。たとえば、第1のユーザ発話信号130は第2のユーザ154による驚いた話しぶりに対応することがあり、一方で第1の代替発話信号172は穏やかな話しぶりに対応することがある。信号処理モジュール122は、第1のユーザ発話信号130に対応する発話パラメータに基づいて第1の代替発話信号172を修正することによって、修正された代替発話信号176を生成することができる。たとえば、信号処理モジュール122は、第1の代替発話信号172のピッチパラメータ、エネルギーパラメータ、または線形予測コーディングパラメータの少なくとも1つを修正して、修正された代替発話信号176を生成することができる。信号処理モジュール122は、第1のデータベース124の中の代替発話信号112に、修正された代替発話信号176を追加することができる。ある特定の態様では、第1のデバイス102は、第2のユーザ154と関連付けられる代替発話信号112の複製を保持するサーバに、修正された代替発話信号176を送ることができる。

30

40

【0053】

したがって、システム100は発話信号の置換を可能にし得る。特定のユーザと関連付けられるユーザ発話信号の一部分は、同じユーザと関連付けられるより高いオーディオ品質の代替発話信号によって置き換えられ得る。

【0054】

図2を参照すると、発話信号を置き換えるように動作可能なシステムの特定の例示的な態様が開示され、全体が200と指定される。システム200は、図1のネットワーク120を介して、第1のデバイス102およびモバイルデバイス104に結合された、またはそれらと通信し

50

ているサーバ206を含み得る。モバイルデバイス104は、図1の信号処理モジュール122を含み得る。サーバ206は、発話信号マネージャ262および第2のデータベース264を含み得る。

【0055】

図2は、第1のデバイス102、モバイルデバイス104、または両方による、代替発話信号112の取得を示す。動作の間、第2のユーザ154は、訓練発話信号(たとえば、訓練発話信号272)をサーバ206に送ることができる。たとえば、第2のユーザ154は、新しい従業員であることがあり、従業員のオリエンテーションの一部としてテキストの台本を読むように求められ得る。第2のユーザ154は、言葉の台本をモバイルデバイス104のマイクロフォン244に向かって読むことによって、訓練発話信号272を送ることができる。たとえば、第2のユーザ154は、静かな環境で台本を読む目的で、音響調整室において台本を読むことができる。モバイルデバイス104の信号処理モジュール122は、ネットワーク120を介して訓練発話信号272をサーバ206に送ることができる。

10

【0056】

発話信号マネージャ262は、訓練発話信号272から代替発話信号112を生成することができる。たとえば、訓練発話信号272は言葉(たとえば、「catatonic」)に対応し得る。発話信号マネージャ262は、訓練発話信号272の特定の部分を複製して代替発話信号112の各々を生成することができる。例示すると、発話信号マネージャ262は、訓練発話信号272を複製して第3の代替発話信号220(たとえば、「catatonic」)を生成することができる。発話信号マネージャ262は、訓練発話信号272の一部分を複製して第4の代替発話信号224(たとえば、「ta」)を生成することができる。発話信号マネージャ262は、訓練発話信号272の別の部分を複製して第5の代替発話信号222(たとえば、「t」)を生成することができる。第4の代替発話信号224および第5の代替発話信号222は、訓練発話信号272の重複する部分に対応し得る。ある特定の態様では、訓練発話信号272の特定の部分は、音素、ダイフォーン、トライフォーン、音節、言葉、またはこれらの組合せに対応する特定の部分に基づいて代替発話信号を生成するために複製され得る。ある特定の態様では、訓練発話信号272の特定の部分は、特定の部分の大きさに基づいて代替発話信号を生成するために複製され得る。たとえば、発話信号マネージャ262は、特定の信号サンプルの大きさ(たとえば、100ミリ秒)に対応する訓練発話信号272の一部分を複製することによって、代替発話信号を生成することができる。

20

【0057】

ある特定の態様では、発話信号マネージャ262は、特定の代替発話信号に対応する音のテキスト表現を決定することができる。たとえば、第5の代替発話信号222は特定の音(たとえば、「t」の音)に対応し得る。発話信号マネージャ262は、テキスト表現に対応する別の発話信号(たとえば、第2のユーザ154の以前に生成された代替発話信号、別のユーザに対応する別の代替発話信号、または合成発話信号)に対する特定の音の比較に基づいて、特定の音のテキスト表現(たとえば、文字「t」)を決定することができる。

30

【0058】

発話信号マネージャ262は、第2のデータベース264に代替発話信号112を記憶することができる。たとえば、発話信号マネージャ262は、変更されない発話として、圧縮されたフォーマットで、または別のフォーマットで、代替発話信号112を第2のデータベース264に記憶することができる。発話信号マネージャ262は、第2のデータベース264に代替発話信号112の選択された特徴を記憶することができる。発話信号マネージャ262は、第2のデータベース264に代替発話信号112の各々のテキスト表現を記憶することができる。

40

【0059】

サーバ206は、代替発話信号(たとえば、代替発話信号112)をデバイス(たとえば、モバイルデバイス104、第1のデバイス102、デスクトップコンピュータ106)に送ることができる。サーバ206は、代替発話信号112をデバイス(たとえば、モバイルデバイス104、第1のデバイス102、デスクトップコンピュータ106)へ定期的に送ることができる。ある特定の態様では、サーバ206は、デバイス(たとえば、モバイルデバイス104、第1のデバイス102、デスクトップコンピュータ106)からの要求を受信したことに応答して、代替発話信号11

50

2をデバイスに送ることができる。ある代替的な態様では、代替発話信号112は、デバイス(たとえば、モバイルデバイス104、第1のデバイス102、またはデスクトップコンピュータ106)から別のデバイス(たとえば、モバイルデバイス104、第1のデバイス102、またはデスクトップコンピュータ106)へ、定期的に、または要求を受信したことに応答してのいずれかで送られ得る。

【0060】

ある特定の態様では、サーバ206は、第2のユーザ154と関連付けられる1つまたは複数のデバイス(たとえば、モバイルデバイス104、デスクトップコンピュータ106)に、第2のユーザ154と関連付けられる代替発話信号112を送ることができる。たとえば、第2のユーザ154は、静かな環境で(たとえば、自宅において)マイクロフォン244に向かって話すことによって、訓練発話信号272をサーバ206に送ることができ、モバイルデバイス104においてサーバ206から代替発話信号112を受信することができる。第2のユーザ154と関連付けられる代替発話信号112は、第2のユーザ154と関連付けられる1つまたは複数のデバイス(たとえば、モバイルデバイス104またはデスクトップコンピュータ106)上の様々なアプリケーションによって使用され得る。たとえば、第2のユーザ154は後で、騒がしい環境において(たとえば、コンサートで)モバイルデバイス104上の音声により有効にされるアプリケーションを使用することがある。図1を参照して説明されたように、モバイルデバイス104の信号処理モジュール122は、第2のユーザ154から受信されるユーザ発話信号の一部分を代替発話信号112の1つで置き換えて、処理された発話信号を生成することができる。処理された発話信号は、モバイルデバイス104上の音声により有効にされるアプリケーションによって使用され得る。したがって、モバイルデバイス104の音声により有効にされるアプリケーションは、騒がしい環境において(たとえば、騒がしいコンサートで)使用されることが可能である。

【0061】

別の例として、モバイルデバイス104上の電子書籍アプリケーションが、オーディオを出力するために代替発話信号112を使用することができる。電子書籍アプリケーションは、電子メール(eメール)、電子ブック(eブック)、記事、音声フィードバック、またはこれらの任意の組合せに対応するオーディオを出力することができる。たとえば、第2のユーザ154は、電子書籍アプリケーションを有効にしてeメールを読むことがある。eメールの一部分が「cat」に対応することがある。代替発話信号112は、代替発話信号112の各々に対応するテキスト表現を示し得る。たとえば、代替発話信号112は、「catatonic」、「ca」、「および「t」という音を含んでよく、各々の音のテキスト表現を示し得る。電子書籍アプリケーションは、代替発話信号112のサブセットがeメールのその部分と一致することを、サブセットのテキスト表現(たとえば、「ca」および「t」)に対するeメールのその部分(たとえば、「cat」)の比較に基づいて決定することができる。電子書籍アプリケーションは、代替発話信号112のサブセットを出力することができる。したがって、電子書籍アプリケーションは、第2のユーザ154のように聞こえる音声でeメールを読み上げることができる。

【0062】

ある特定の態様では、システム200は、代替発話信号112への認められていないアクセスを不可能にするためのロック機構を含み得る。デバイス(たとえば、第1のデバイス102、モバイルデバイス104、またはデスクトップコンピュータ106)は、(たとえば、サーバから)認証を受ける前に代替発話信号112を生成する(または使用する)ことが不可能であり得る。この認証は、第1のユーザ152、第2のユーザ154、または両方が特定のサービスに対して支払いを行ったかどうかに基づき得る。この認証は、使用ごと(たとえば、音声呼ごと、アプリケーションの起動ごと)に有効であり、期間ごと(たとえば、課金サイクル、週、月など)に有効であり、またはこれらの組合せであり得る。ある特定の態様では、この認証は、代替発話信号112の更新ごとに有効であり得る。ある特定の態様では、この認証は、アプリケーションごとに有効であり得る。たとえば、モバイルデバイス104は、音声呼の間に代替発話信号112を使用するために第1の認証を受け、電子書籍アプリケーションと

10

20

30

40

50

もに代替発話信号112を使用するために別の認証を受けることができる。ある特定の態様では、この認証は、代替発話信号112を生成(または使用)することをデバイスが免許されているかどうかに基づき得る。たとえば、モバイルデバイス104は、モバイルデバイス104上の電子書籍アプリケーションを使用するためのライセンスを第2のユーザ154が取得したことに応答して、電子書籍アプリケーションにおいて代替発話信号112を使用するための認証を受け取ることができる。

【0063】

ある特定の態様では、発話信号マネージャ262は、代替発話信号112にアクセスすることが認められている、デバイス、ユーザ、または両方のリストを保持し得る。発話信号マネージャ262は、デバイスが代替発話信号112にアクセスすることが認められていないと決定したことに応答して、ロック要求をデバイス(たとえば、第1のデバイス102、モバイルデバイス104、デスクトップコンピュータ106、またはこれらの組合せ)に送ることができる。デバイスは、ロック要求に応答して、ローカルに記憶された代替発話信号112を削除する(またはその信号へのアクセスを不可能にする)ことができる。たとえば、発話信号マネージャ262は、代替発話信号112にアクセスすることが以前は認められていたデバイス(たとえば、第1のデバイス102、モバイルデバイス104、またはデスクトップコンピュータ106)に、ロック要求を送ることができる。

10

【0064】

ある特定の態様では、信号処理モジュール122は、代替発話信号112の各アクセスの前に対応するデバイス(たとえば、第1のデバイス102)の認証ステータスを決定することができる。たとえば、信号処理モジュール122は、認証ステータス要求をサーバ206に送ることができる。サーバ206から認証ステータスを受信することができる。信号処理モジュール122は、第1のデバイス102が認証されていないと決定したことに応答して、代替発話信号112にアクセスするのを控えることができる。ある特定の態様では、信号処理モジュール122は代替発話信号112を削除する(または信号へのアクセスを不可能にする)ことができる。たとえば、信号処理モジュール122は、代替発話信号112を削除し、または、特定のアプリケーションによる代替発話信号112へのアクセスを不可能にすることができる。

20

【0065】

ある特定の態様では、信号処理モジュール122は、第1のユーザ入力信号が信号処理モジュール122のユーザインターフェースを介して受け取られたことに応答して、代替発話信号112を削除することができる。ある特定の態様では、信号処理モジュール122は、第2のユーザ入力信号が信号処理モジュール122のユーザインターフェースを介して受け取られたことに応答して、削除要求をサーバ206に送ることができる。削除要求に応答して、発話信号マネージャ262は代替発話信号112を削除する(または信号へのアクセスを不可能にする)ことができる。ある特定の態様では、発話信号マネージャ262は、削除要求が代替発話信号112に対応するユーザ(たとえば、第2のユーザ154)と関連付けられるデバイス(たとえば、モバイルデバイス104またはデスクトップコンピュータ106)から受け取られたと決定したことに応答して、代替発話信号112を削除する(または信号へのアクセスを不可能にする)ことができる。

30

【0066】

ある特定の態様では、モバイルデバイス104は発話信号マネージャ262を含み得る。モバイルデバイス104は訓練発話信号272を受信することができ、モバイルデバイス104上の発話信号マネージャ262は代替発話信号112を生成することができる。発話信号マネージャ262は、モバイルデバイス104に代替発話信号112を記憶することができる。この態様では、モバイルデバイス104は、代替発話信号112をサーバ206に送ることができる。

40

【0067】

ある特定の態様では、サーバ206は、第2のユーザ154とは別のユーザ(たとえば、第1のユーザ152)と関連付けられる1つまたは複数のデバイス(たとえば、第1のデバイス102)に、第2のユーザ154と関連付けられる代替発話信号112を送ることができる。たとえば、サーバ206は、閾値の呼の頻度を満たすある特定の期間にわたる呼の頻度を示す呼プロファ

50

イル(第1の呼プロファイル270、第2の呼プロファイル274)に基づいて、代替発話信号112を第1のデバイス102に送ることができる。第1の呼プロファイル270は、第1のユーザ152、第1のデバイス102、または両方と関連付けられ得る。第2の呼プロファイル274は、第2のユーザ154、モバイルデバイス104、または両方と関連付けられ得る。呼の頻度は、第1のユーザ152、第1のデバイス102、または両方と、第2のユーザ154、モバイルデバイス104、または両方との間の呼の頻度を示し得る。例示すると、サーバ206は、第1のユーザ152と第2のユーザ154との間の呼の頻度が閾値の呼の頻度(たとえば、先週に3回)を満たすことを示す第1の呼プロファイル270に基づいて、代替発話信号112を第1のデバイス102に送ることができる。ある特定の態様では、サーバ206は、音声呼の前に代替発話信号112を第1のデバイス102に送ることができる。たとえば、サーバ206は、呼プロファイル(たとえば、第1の呼プロファイル270または第2の呼プロファイル274)に基づいて、音声呼の前に代替発話信号112を第1のデバイス102に送ることができる。

10

【0068】

別の例として、サーバ206は、特定のユーザ(たとえば、第1のユーザ152、第2のユーザ154)、特定のデバイス(たとえば、第1のデバイス102、モバイルデバイス104)、またはこれらの組合せを示す連絡先リスト(たとえば、第1の連絡先リスト276、第2の連絡先リスト278)に基づいて、代替発話信号112を第1のデバイス102に送ることができる。たとえば、サーバ206は、ある特定のユーザ(たとえば、第1のユーザ152)を示す第2の連絡先リスト278に基づいて、その特定のユーザと関連付けられる1つまたは複数のデバイス(たとえば、第1のデバイス102)に、第2のユーザ154と関連付けられる代替発話信号112を送ることができる。ある特定の態様では、サーバ206は、音声呼の前に代替発話信号112を第1のデバイス102に送ることができる。たとえば、サーバ206は、連絡先リスト(たとえば、第1の連絡先リスト276、第2の連絡先リスト278)に基づいて、音声呼の前に代替発話信号112を第1のデバイス102に送ることができる。サーバ206は、訓練発話信号272を受信したことに応答して、代替発話信号112を第1のデバイス102に送ることができる。

20

【0069】

ある特定の態様では、サーバ206は、第1のデバイス102から代替発話信号112に対する要求を受信したことに応答して、代替発話信号112を第1のデバイス102へ送るためのモバイルデバイス104からの要求を受信したことに応答して、またはこれらの両方に応答して、代替発話信号112を送ることができる。たとえば、第1のデバイス102は、第1のデバイス102の信号処理モジュール122のユーザインターフェースを介して受け取られたユーザ入力に応答して、代替発話信号112に対する要求を送ることができる。別の例として、モバイルデバイス104は、モバイルデバイス104の信号処理モジュール122のユーザインターフェースを介して受け取られたユーザ入力に応答して、代替発話信号112を送るための要求を送ることができる。ある特定の態様では、サーバ206は、第1の連絡先リスト276、第2の連絡先リスト278、または両方に関する連絡先リスト更新通知を受信したことに基づいて、代替発話信号112を第1のデバイス102に送ることができる。サーバ206は、第1のデバイス102から第1の連絡先リスト276に関する連絡先リスト更新通知を受信することができ、かつサーバ206は、モバイルデバイス104から第2の連絡先リスト278に関する連絡先リスト更新通知を受信することができる。たとえば、サーバ206は、対応するデバイスの信号処理モジュール122から連絡先リスト更新通知を受信することができる。

30

40

【0070】

ある特定の態様では、サーバ206は、第1の呼プロファイル270、第2の呼プロファイル274、または両方に関する呼プロファイル更新通知を受信したことに基づいて、代替発話信号112を第1のデバイス102に送ることができる。サーバ206は、第1のデバイス102から第1の呼プロファイル270に関する呼プロファイル更新通知を受信することができ、かつサーバ206は、モバイルデバイス104から第2の呼プロファイル274に関する呼プロファイル更新通知を受信することができる。たとえば、サーバ206は、対応するデバイスの信号処理モジュール122から呼プロファイル更新通知を受信することができる。

【0071】

50

さらなる例として、サーバ206は、第2のユーザ154と関連付けられるデバイス(たとえば、モバイルデバイス104、デスクトップコンピュータ106)と第1のデバイス102との間の音声呼に関する呼開始通知に応答して、第2のユーザ154と関連付けられる代替発話信号112を第1のデバイス102に送ることができる。たとえば、サーバ206は、音声呼がモバイルデバイス104から行われるとき、モバイルデバイス104から呼開始通知を受信することができる。ある特定の態様では、モバイルデバイス104の信号処理モジュール122は、呼開始通知をサーバ206に送ることができる。サーバ206は、音声呼の初めに、音声呼の間に、またはその両方において、代替発話信号112の少なくともサブセットを送ることができる。たとえば、サーバ206は、呼開始通知を受信したことに応答して、代替発話信号112を送ることを開始し得る。代替発話信号112の第1のサブセットは、音声呼の初めにサーバ206によって送られ得る。代替発話信号112の第2のサブセットは、音声呼の間にサーバ206によって送られ得る。

10

【0072】

ある特定の態様では、発話信号マネージャ262は、システム管理者が診断を実行することを可能にするユーザインターフェースを含み得る。たとえば、ユーザインターフェースは、代替発話信号112と関連付けられるアップロードとダウンロードの頻度を含む使用プロファイルを表示することができる。使用プロファイルはまた、デバイス(たとえば、第1のデバイス102、モバイルデバイス104、および/またはデスクトップコンピュータ106)ごとのアップロードおよびダウンロードの頻度を含み得る。

20

【0073】

したがって、システム200は、図1のシステム100によって発話信号を置き換えるために使用され得る、代替発話信号の取得を可能にし得る。

【0074】

図3を参照すると、ユーザインターフェースの例示的な態様が開示され、全体が300と指定される。ある特定の態様では、ユーザインターフェース300は、図1のシステム100、図2のシステム200、または両方によって表示され得る。

【0075】

動作の間、デバイス(たとえば、第1のデバイス102)の信号処理モジュール122は、ユーザ(たとえば、第1のユーザ152)の連絡先リストへのアクセス権を有し得る。信号処理モジュール122は、ユーザインターフェース300を第1のデバイス102のディスプレイに提供することができる。たとえば、信号処理モジュール122は、第1のユーザ152から要求を受け取ったことに応答して、ユーザインターフェース300を第1のデバイス102のディスプレイに提供することができる。

30

【0076】

ユーザインターフェース300の水平軸(たとえば、x軸)は、連絡先リスト(たとえば、「Betsy Curtis」、「Brett Dean」、「Henry Lopez」、「Sabrina Sanders」、および「Randal Hughes」)から1人または複数のユーザの識別子(たとえば、名前)を表示することができる。ある特定の態様では、第1のユーザ152は、連絡先リストから1人または複数のユーザを選択することができ、ユーザインターフェース300は、水平軸に沿って選択されたユーザの名前を含み得る。別の態様では、信号処理モジュール122は、連絡先リストからユーザのサブセットを自動的に選択することができる。たとえば、ユーザのサブセットは、第1のユーザ152が最も頻繁に通信する、第1のユーザ152が直近に通信した、第1のユーザ152がある特定の時間間隔の間に最も長い時間(たとえば、電話呼の時間の合計)通信した、またはこれらの組合せの、特定の数(たとえば、5人)のユーザであり得る。

40

【0077】

ユーザインターフェース300は、垂直軸(たとえば、y軸)に沿った時間長を示し得る。この時間長は、1人または複数のユーザに対応する発話信号と関連付けられ得る。たとえば、ユーザインターフェース300は、第2のユーザ154(たとえば、「Randal Hughes」)に対応する第1のバー302を含み得る。第1のバー302は、第2のユーザ154と関連付けられる1つまたは複数の発話信号の再生時間長を示し得る。たとえば、第1のバー302は、訓練発話信号

50

272の再生時間長を示し得る。ある特定の態様では、第1のバー302は、ある特定の時間間隔(たとえば、特定の日、特定の週、または特定の月)にわたって捉えられた第2のユーザ154と関連付けられる発話信号の再生時間長を示し得る。

【0078】

ある特定の態様では、ユーザインターフェース300は、別個の時間間隔にわたって捉えられた特定のユーザの発話信号の時間長を示し得る。たとえば、第1のバー302は、第1の時間間隔(たとえば、先週)にわたって捉えられた第2のユーザ154の発話信号の第1の時間長を示してよく、別のユーザ(たとえば、「Sabrina Sanders」)と関連付けられる第2のバー314は、第2の時間間隔(たとえば、先月)にわたって捉えられた他のユーザの発話信号の第2の時間長を示してよい。たとえば、第1のユーザ152は、1週間前に第2のユーザ154と関連付けられる代替発話信号をダウンロードしていることがあり、第1のバー302は、そのダウンロードの後に捕捉された第2のユーザ154の発話信号の第1の時間長を示し得る。別の例として、第1のユーザ152は、1ヶ月前に他のユーザと関連付けられる代替発話信号をダウンロードしていることがあり、第2のバー314は、そのダウンロードの後に捕捉された他のユーザ(たとえば、「Sabrina Sanders」)の発話信号と関連付けられる第2の時間長を示し得る。

10

【0079】

第1のユーザ152は、カーソル306を使用して、第2のユーザ154の名前(たとえば、「Randal Hughes」)または第1のバー302を選択することができる。ユーザインターフェース300は、この選択を受け取ったことに応答して、追加オプション308、削除オプション310、または両方を表示することができる。第1のユーザ152は削除オプション310を選択することができる。たとえば、第1のユーザ152は、第2のユーザ154と関連付けられる追加の代替発話信号が生成されるべきではないと決定することができる。例示すると、第1のユーザ152は第2のユーザ154と関連付けられる代替発話信号が十分収集されていると決定することがあり、第1のユーザ152は近い将来第2のユーザ154と通信することを予想しないことがあり、第1のユーザ152はリソース(たとえば、メモリまたは処理サイクル)を節約するのを望むことがあり、またはこれらの組合せであり得る。削除オプション310の選択を受け取ったことに応答して、信号処理モジュール122は、第2のユーザ154と関連付けられる1つまたは複数の発話信号を除去することができ、第2のユーザ154の発話信号と関連付けられる第1の時間長の指示をユーザインターフェース300から除去することができ、またはこれらの両方であり得る。

20

30

【0080】

ある特定の態様では、第2のユーザ154と関連付けられる発話信号は、第1のデバイス102がアクセス可能なメモリに記憶されてよく、信号処理モジュール122は、削除オプション310の選択を受け取ったことに応答して、第2のユーザ154と関連付けられる1つまたは複数の発話信号を除去することができる。ある特定の態様では、第2のユーザ154と関連付けられる1つまたは複数の発話信号は、サーバ(たとえば、図2のサーバ206)がアクセス可能なメモリに記憶され得る。信号処理モジュール122は、削除オプション310の選択を受け取ったことに応答して、削除要求をサーバ206に送ることができる。削除要求を受け取ったことに応答して、サーバ206は、第2のユーザ154と関連付けられる発話信号を除去することができ、第2のユーザ154と関連付けられる発話信号を第1のデバイス102に対してアクセス不可能にでき、第2のユーザ154と関連付けられる発話信号を第1のユーザ152に対してアクセス可能にでき、またはこれらの組合せであり得る。たとえば、第1のユーザ152は、第2のユーザ154と関連付けられる発話信号を第1のデバイス102に対してアクセス不可能にするために、削除オプション310を選択することができる。この例では、第1のユーザ152は、第2のユーザ154と関連付けられる1つまたは複数の発話信号を、他のデバイスがアクセス可能なままにすることを望み得る。

40

【0081】

代替的に、第1のユーザ152は追加オプション308を選択することができる。追加オプション308の選択を受け取ったことに応答して、信号処理モジュール122は、第2のユーザ154と

50

関連付けられる発話信号から生成される代替発話信号(たとえば、代替発話信号112)を提供するように、図2の発話信号マネージャ262に要求し得る。図2を参照して説明されたように、発話信号マネージャ262は、第2のユーザ154と関連付けられる発話信号を使用して、代替発話信号112を生成することができる。発話信号マネージャ262は、信号処理モジュール122から要求を受け取ったことに応答して、代替発話信号112を信号処理モジュール122に提供することができる。

【 0 0 8 2 】

ある特定の態様では、信号処理モジュール122は、削除オプション310の選択を受け取った後に、追加オプション308の選択を受け取った後に、または、発話信号マネージャ262から代替発話信号112を受け取った後に、第1のパー302を(たとえば、0に)リセットすることができる。第1のパー302は、第2のユーザ154と関連付けられる発話信号の以前の削除の後に、または第2のユーザ154と関連付けられる他の代替発話信号の以前のダウンロード(または生成)の後に捕捉された、第2のユーザ154の発話信号と関連付けられる第1の時間長を示し得る。

10

【 0 0 8 3 】

ある特定の態様では、信号処理モジュール122は、特定のユーザ(たとえば、「Brett Dean」)の発話信号に対応する(たとえば、第3のパー316によって示される)時間長が自動更新の閾値312(たとえば、1時間)を満たすと決定したことに応答して、代替発話信号に対する要求を発話信号マネージャ262に自動的に送ることができる。

【 0 0 8 4 】

ある特定の態様では、信号処理モジュール122は、第2のユーザ154と関連付けられる発話信号に対応する代替発話信号112を定期的に(たとえば、1週間に1回)要求することができる。別の態様では、発話信号マネージャ262は、第1のデバイス102に代替発話信号112を定期的に送ることができる。ある特定の態様では、ユーザインターフェース300は、第1のユーザ152が定期的な更新に加えて代替発話信号112をダウンロードすることを可能にし得る。

20

【 0 0 8 5 】

ある特定の態様では、第2のユーザ154と関連付けられる発話信号は、第2のユーザ154と関連付けられるデバイス(たとえば、モバイルデバイス104)において信号処理モジュール122によって、第1のユーザ152と関連付けられるデバイス(たとえば、第1のデバイス102)によって、サーバ(たとえば、図2のサーバ206)によって、またはこれらの組合せによって捕捉され得る。たとえば、第1のデバイス102(またはモバイルデバイス104)における信号処理モジュール122は、第2のユーザ154と第1のユーザ152との間の電話呼の間に発話信号を記録することができる。別の例として、発話信号は、サーバ206に記憶された第2のユーザ154からのオーディオメッセージに対応し得る。ある特定の態様では、モバイルデバイス104、第1のデバイス102、または両方が、第2のユーザ154と関連付けられる発話信号をサーバ206に提供することができる。ある特定の態様では、ユーザインターフェース300は、テキスト表現、グラフィカル表現(たとえば、棒グラフ、円グラフ、または両方)、または両方を使用して、時間長を示し得る。

30

【 0 0 8 6 】

ユーザインターフェース300はこうして、代替発話信号の生成のためにユーザと関連付けられる発話信号の利用可能性を第1のユーザ152が監視することを可能にし得る。ユーザインターフェース300は、第1のユーザ152が、代替発話信号を生成することとリソース(たとえば、メモリ、処理サイクル、または両方)を節約することとの間から選ぶことを可能にし得る。

40

【 0 0 8 7 】

図4を参照すると、発話信号の「符号化前の」置換を実行するように動作可能なシステムの例示的な態様が開示され、全体が400と指定される。システム400は、ネットワーク120を介して第1のデバイス102と通信しているモバイルデバイス104を含み得る。モバイルデバイス104は、エンコーダ426に結合された、またはそれと通信している、信号処理モジュ

50

ール422を含み得る。モバイルデバイス104は、データベース分析器410を含み得る。モバイルデバイス104は、第1のデータベース124、パラメータ化されたデータベース424、または両方を含み得る。

【0088】

図4は、置き換えられた発話信号をエンコーダ(たとえば、エンコーダ426)に提供する前の、発話信号の置換を示す。動作の間、モバイルデバイス104は、第2のユーザ154から、マイクロフォン146を介して第1のユーザ発話信号130を受信し得る。第1のデータベース124(たとえば、「生の」データベース)は、代替発話信号112を含み得る。代替発話信号112は、第1の代替発話信号172、第2の代替発話信号174、およびN個の代替発話信号478を含み得る。代替発話信号112の各々は、図2を参照して説明されたように生成され得る。

10

【0089】

パラメータ化されたデータベース424は、代替発話信号112と関連付けられる発話パラメータのセットを含み得る。たとえば、パラメータ化されたデータベース424は、第1の代替発話信号172の第1の発話パラメータ472、第2の代替発話信号174の第2の発話パラメータ474、およびN個の代替発話信号478のN個の発話パラメータ476を含み得る。発話パラメータ472、474、および476は、パラメータ化されたデータベース424が第1のデータベース124よりも小さくなり得るように、代替発話信号112よりも少ない記憶空間を占有し得る。

【0090】

ある特定の態様では、パラメータ化されたデータベース424は、第1のデータベース124中の対応する代替発話信号の位置に対する、発話パラメータ472、474、および476の各々と関連付けられる特定の参照(たとえば、インデックス、データベースアドレス、データベースポインタ、メモリアドレス、uniform resource locator(URL)、uniform resource identifier(URI)、またはこれらの組合せ)を含み得る。たとえば、パラメータ化されたデータベース424は、第1のデータベース124中の第1の代替発話信号172の第1の位置に対する、第1の発話パラメータ472と関連付けられる第1の参照を含み得る。信号処理モジュール422は、特定の発話パラメータと関連付けられる対応する代替発話信号にアクセスするために、特定の参照を使用し得る。たとえば、信号処理モジュール422は、第1の発話パラメータ472と関連付けられる第1の代替発話信号172にアクセスするために、第1の参照を使用し得る。

20

【0091】

ある特定の態様では、第1のデータベース124は、パラメータ化されたデータベース424中の対応する発話パラメータの位置に対する、代替発話信号112の各々と関連付けられる特定の参照(たとえば、インデックス、データベースアドレス、データベースポインタ、メモリアドレス、uniform resource locator(URL)、uniform resource identifier(URI)、またはこれらの組合せ)を含み得る。たとえば、第1のデータベース124は、パラメータ化されたデータベース424中の第1の発話パラメータ472の第1の位置に対する、第1の代替発話信号172と関連付けられる第1の参照を含み得る。信号処理モジュール422は、特定の代替発話信号と関連付けられる対応する発話パラメータにアクセスするために、特定の参照を使用し得る。たとえば、信号処理モジュール422は、第1の代替発話信号172の第1の発話パラメータ472にアクセスするために、第1の参照を使用し得る。

30

40

【0092】

ある特定の態様では、データベース分析器410は、モバイルデバイス104が第1のユーザ発話信号130を受け取ったことに応答して、パラメータ化されたデータベース424が最新であるかどうかを決定することができる。たとえば、データベース分析器410は、パラメータ化されたデータベース424が代替発話信号112と関連付けられる発話パラメータ(たとえば、発話パラメータ472、474、および476)を含むかどうかを決定することができる。例示すると、データベース分析器410は、パラメータ化されたデータベース424が代替発話信号112の1つまたは複数の代替発話信号に対応する発話パラメータを含むかどうかを決定することができる。パラメータ化されたデータベース424が代替発話信号112と関連付けられる発話パラメータを含むと決定したことに応答して、データベース分析器410は、パラメー

50

タ化されたデータベース424の以前の更新の後に第1のデータベース124に対して修正が行われたかどうかを決定することができる。例示すると、データベース分析器410は、第1のデータベース124の修正タイムスタンプがパラメータ化されたデータベース424の更新タイムスタンプより後であるかどうかを決定することができる。パラメータ化されたデータベース424の以前の更新の後に第1のデータベース124に対する修正が行われなかったと決定したことに応答して、データベース分析器410は、パラメータ化されたデータベース424が最新であると決定することができる。

【0093】

パラメータ化されたデータベース424が代替発話信号112と関連付けられる発話パラメータを含まない、または、パラメータ化されたデータベース424の以前の更新の後に第1のデータベース124に対する修正が行われたと決定したことに応答して、データベース分析器410は、代替発話信号112に対応する発話パラメータを生成することができる。たとえば、データベース分析器410は、第1の代替発話信号172から第1の発話パラメータ472を、第2の代替発話信号174から第2の発話パラメータ474を、N個の代替発話信号478からN個の発話パラメータ476を生成し得る。第1の発話パラメータ472、第2の発話パラメータ474、およびN個の発話パラメータ476は、それぞれ、第1の代替発話信号172、第2の代替発話信号174、およびN個の代替発話信号478の、ピッチパラメータ、エネルギーパラメータ、線形予測コーディング(LPC)パラメータ、またはこれらの任意の組合せを含み得る。

【0094】

信号処理モジュール422は、分析器402、検索器404、制約分析器406、および合成器408を含み得る。分析器402は、第1のユーザ発話信号130に基づいて検索基準418を生成することができる。検索基準418は、第1のユーザ発話信号130と関連付けられる複数の発話パラメータを含み得る。たとえば、複数の発話パラメータは、第1のユーザ発話信号130の、ピッチパラメータ、エネルギーパラメータ、線形予測コーディング(LPC)パラメータ、またはこれらの任意の組合せを含み得る。分析器402は、検索基準418を検索器404に提供することができる。

【0095】

検索器404は、検索基準418に基づいてパラメータ化されたデータベース424を検索することによって、検索結果414を生成することができる。たとえば、検索器404は、複数の発話パラメータを発話パラメータ472、474、および476の各々と比較することができる。検索器404は、第1の発話パラメータ472および第2の発話パラメータ474が複数の発話パラメータと一致すると決定することができる。たとえば、検索器404は、この比較に基づいて第1の発話パラメータ472および第2の発話パラメータ474を選択(または特定)することができる。例示すると、第1のユーザ発話信号130は「bat」に対応することがある。第1の代替発話信号172は、図2を参照して説明されたように、「cat」から生成された「a」に対応し得る。第2の代替発話信号174は、図2を参照して説明されたように、「tag」から生成された「a」に対応し得る。

【0096】

ある特定の態様では、検索器404は、発話パラメータ472、474、および476の各々と関連付けられる類似性の尺度に基づいて、検索結果414を生成することができる。特定の類似性の尺度は、対応する発話パラメータが、第1のユーザ発話信号130と関連付けられる複数の発話パラメータとどの程度密接に関連するかを示し得る。たとえば、検索器404は、第1のユーザ発話信号130の複数の発話パラメータと第1の発話パラメータ472との差を算出することによって、第1の発話パラメータ472の第1の類似性の尺度を決定することができる。

【0097】

検索器404は、類似性の閾値を満たす類似性の尺度を有する発話パラメータを選択することができる。たとえば、検索器404は、第1の類似性の尺度および第2の発話パラメータ474の第2の類似性の尺度が類似性の閾値を満たす(たとえば、それ未満である)と決定したことに応答して、第1の発話パラメータ472および第2の発話パラメータ474を選択すること

ができる。

【0098】

別の例として、検索器404は、類似性の尺度に基づいて、第1のユーザ発話信号130の複数の発話パラメータと最も類似している特定の数(たとえば、2個)の発話パラメータを選択することができる。検索器404は、選択された発話パラメータ(たとえば、第1の発話パラメータ472および第2の発話パラメータ474)を検索結果414に含めることができる。ある特定の態様では、検索器404は、選択された複数の発話パラメータに対応する代替発話信号を検索結果414に含め得る。たとえば、検索器404は、第1の代替発話信号172および第2の代替発話信号174を検索結果414に含め得る。検索器404は、検索結果414を制約分析器406に提供することができる。

10

【0099】

制約分析器406は、制約に基づいて検索結果414から特定の検索結果(たとえば、選択された結果416)を選択することができる。この制約は、誤差の制約、コストの制約、または両方を含み得る。たとえば、制約分析器406は、誤差の制約またはコストの制約に基づいて、検索結果414から第1の発話パラメータ472(または第1の代替発話信号172)を選択することができる。例示すると、制約分析器406は、検索結果414の各々に基づいて処理された発話信号を生成することができる。処理された発話信号の各々は、対応する検索結果によって置き換えられる部分を有する第1のユーザ発話信号130を表し得る。制約分析器406は、処理された発話信号の最小の誤差を有する特定の処理された発話信号を特定することができる。たとえば、制約分析器406は、第1のユーザ発話信号130との処理された発話信号の比較に基づいて、処理された発話信号の各々と関連付けられる誤差の尺度を決定することができる。ある特定の態様では、制約分析器406は、最小のコストを有する特定の処理された発話信号を特定することができる。たとえば、制約分析器406は、対応する検索結果、1つまたは複数の以前の代替発話信号、1つまたは複数の後続の代替発話信号、またはこれらの組合せと関連付けられる代替発話信号を取り出すコストに基づいて、処理された発話信号の各々と関連付けられるコストの尺度を決定することができる。例示すると、複数の発話信号を取り出すコストは、複数の発話信号に対応するメモリ位置の違いに基づいて決定され得る。たとえば、最小のコストは、メモリの中で以前に取り出された代替発話信号のより近くに位置する後続の代替発話信号を取り出すことと関連付けられ得る。制約分析器406は、特定の処理された発話信号に対応する検索結果を選択された結果416として選択することができる。

20

30

【0100】

合成器408は、選択された結果416に基づいて第1の処理された発話信号116を生成することができる。たとえば、選択された結果416が第1の代替発話信号172と関連付けられる場合、合成器408は、第1のユーザ発話信号130の少なくとも一部分を第1の代替発話信号172で置き換えることによって、第1の処理された発話信号116を生成することができる。ある特定の態様では、合成器408は、第1の発話パラメータ472に基づいて置換発話信号を生成(たとえば、合成)することができ、第1のユーザ発話信号130の少なくとも一部分を置換発話信号で置き換えることによって、第1の処理された発話信号116を生成することができる。

40

【0101】

合成器408は、第1の処理された発話信号116をエンコーダ426に提供することができる。エンコーダ426は、第1の処理された発話信号116を符号化することによって出力信号430を生成することができる。ある特定の態様では、出力信号430は第1の代替発話信号172を示し得る。たとえば、出力信号430は、第1のデータベース124中の第1の代替発話信号172への参照、パラメータ化されたデータベース424中の第1の発話パラメータ472への参照、または両方を含み得る。モバイルデバイス104は、第1のデバイス102へ、ネットワーク120を介して出力信号430を送り得る。

【0102】

システム400はこうして、ユーザ発話信号を符号化する前の、代替発話信号によるユー

50

ザ発話信号の一部分の「符号化前の」置換を可能にし得る。代替発話信号は、ユーザ発話信号の置き換えられる部分よりも高いオーディオ品質を有し得る。たとえば、第2のユーザ154は、騒がしい環境において(たとえば、コンサートまたは交通の多い通りで)モバイルデバイス104を使って話していることがある。図2を参照して説明されたように、第2のユーザ154が静かな環境において提供した訓練信号から、代替発話信号が生成され得る。信号処理モジュール422は、元のユーザ発話信号よりも高い品質のオーディオをエンコーダに提供することができる。エンコーダは、より高い品質のオーディオを符号化することができ、符号化された信号を別のデバイスに送信することができる。

【0103】

図5を参照すると、発話信号の「符号化中の」置換を実行するように動作可能なシステムの例示的な態様が開示され、全体が500と指定される。システム500は、エンコーダ526、データベース分析器410、第1のデータベース124、パラメータ化されたデータベース424、またはこれらの組合せを含み得る。

10

【0104】

図5は、エンコーダ(たとえば、エンコーダ526)における発話信号の置換を示す。図4の態様はエンコーダ426に結合された信号処理モジュール422を示すが、エンコーダ526は信号処理モジュール422の1つまたは複数のコンポーネントを含む。動作の間、モバイルデバイス104は、第1のユーザ発話信号130を受け取り得る。図4を参照して説明されたように、エンコーダ526内で、分析器402が検索基準418を生成することができ、検索器404が検索結果414を生成することができ、制約分析器406が選択された結果416を生成することができ、合成器408が第1の処理された発話信号116を生成することができる。エンコーダ526は、第1の処理された発話信号116を符号化することによって出力信号430を生成することができる。モバイルデバイス104は、第1のデバイス102へ、ネットワーク120を介して出力信号430を送り得る。

20

【0105】

ある特定の態様では、制約分析器406は、図4を参照して説明されたように、選択された結果416に対応する第1の処理された信号を生成することができる。エンコーダ526は、第1の処理された信号を符号化することによって出力信号430を生成することができる。この態様では、信号処理モジュール422は、第1の処理された発話信号116を生成するのを控えることができる。ある特定の態様では、図4を参照して説明されたように、出力信号430は第1の代替発話信号172を示し得る。

30

【0106】

システム500はこうして、代替発話信号と関連付けられるユーザ発話信号の一部分の「符号化中の」置換を実行することによって、エンコーダにおける処理された発話信号の生成を可能にし得る。処理された発話信号は、ユーザ発話信号よりも高いオーディオ品質を有し得る。たとえば、第2のユーザ154は、騒がしい環境において(たとえば、コンサートまたは交通の多い通りで)モバイルデバイス104を使って話していることがある。図2を参照して説明されたように、第2のユーザ154が静かな環境において提供した訓練信号から、代替発話信号が生成され得る。エンコーダは、元のユーザ発話信号よりも高い品質のオーディオを符号化することができ、符号化された信号を別のデバイスに送信することができる。ある特定の態様では、ユーザ発話信号のその部分の符号化中の置換は、ユーザ発話信号から符号化された発話信号を生成することと関連付けられるステップを減らすことができ、ユーザ発話信号のその部分の符号化前の置換よりも高速および効率的な信号処理をもたらす。

40

【0107】

図4～図5の態様はエンコーダシステムにおける発話信号の置換を示すが、図6および図7の態様は、デコーダシステムにおける発話信号の置換を示す。ある特定の態様では、発話信号の置換は、エンコーダシステムのローカルデコーダにおいて行われ得る。たとえば、第1の処理された発話信号は、第1の処理された発話信号とユーザ発話信号(たとえば、第1のユーザ発話信号130)との比較に基づいて信号パラメータを決定するために、ローカルデ

50

コードによって生成され得る。例示すると、ローカルデコードは、別のデバイスにおけるデコードの挙動をエミュレートし得る。エンコードシステムは、他のデバイスに送信すべき信号パラメータを符号化することができる。

【0108】

図6を参照すると、デコードにおいて発話信号の「パラメトリックな」置換を実行するように動作可能なシステムの例示的な態様が開示され、全体が600と指定される。システム600は、第1のデバイス102を含む。第1のデバイス102は、第1のデータベース124、パラメータ化されたデータベース424、データベース分析器410、デコード626、またはこれらの組合せを含み得る。デコード626は、信号処理モジュール622を含んでよく、それに結合されてよく、またはそれと通信してよい。

【0109】

図6は、受信デバイスにおける発話信号の置換を示す。動作の間、第1のデバイス102は、モバイルデバイス104からネットワーク120を介して第1のデータ630を受信することができる。第1のデータ630は、オーディオ信号の1つまたは複数のフレームに対応するパケットデータを含み得る。オーディオ信号は、ユーザ(たとえば、図1の第2のユーザ154)に対応するユーザ発話信号であり得る。第1のデータ630は、オーディオ信号と関連付けられる複数のパラメータを含み得る。複数のパラメータは、ピッチパラメータ、エネルギーパラメータ、線形予測コーディング(LPC)パラメータ、またはこれらの任意の組合せを含み得る。

【0110】

分析器602は、第1のデータ630に基づいて検索基準612を生成することができる。たとえば、分析器602は、第1のデータ630からオーディオ信号と関連付けられる複数の発話パラメータを抽出することができる。探索基準612は、抽出された複数の発話パラメータを含み得る。分析器602は、検索基準612を検索器604に提供することができる。

【0111】

検索器604は、図4の発話パラメータ472、474、および476の1つまたは複数の発話パラメータを特定することができる。ある特定の態様では、図4の検索器404を参照して説明されたように、検索器604は、複数の発話パラメータを発話パラメータ472、474、および476の各々と比較することによって、第1の発話パラメータ472および第2の発話パラメータ474を特定することができる。検索器604は、発話パラメータ614のセットに特定された発話パラメータ(たとえば、第1の発話パラメータ472および第2の発話パラメータ474)を含めることができる。検索器604は、発話パラメータ614のセットを制約分析器606に提供することができる。図4の制約分析器406を参照して説明されたように、制約分析器606は、制約に基づいて発話パラメータ614のセットの第1の発話パラメータ472を選択することができる。制約分析器606は、第1の発話パラメータ472を合成器608に提供することができる。

【0112】

合成器608は、第1のデータ630中の複数の発話パラメータを第1の発話パラメータ472で置き換えることによって、第2のデータ618を生成することができる。デコード626は、第2のデータ618を復号することによって、第1の処理された発話信号116を生成することができる。たとえば、第1の処理された発話信号116は、第1の発話パラメータ472に基づいて生成された処理された信号であり得る。第1のデバイス102は、第1の処理された発話信号116をスピーカ142に提供することができる。

【0113】

ある特定の態様では、分析器602は、第1のデータ630が第1の代替発話信号172を示すと決定することができる。たとえば、分析器602は、第1のデータ630が、第1のデータベース124中の第1の代替発話信号172への参照、パラメータ化されたデータベース424中の第1の発話パラメータ472への参照、または両方を含むと決定することができる。分析器602は、第1のデータ630が第1の代替発話信号172を示すと決定したことに応答して、第1の発話パラメータ472を合成器608に提供することができる。この態様では、分析器602は、検索基準612を生成するのを控えることができる。

10

20

30

40

50

【0114】

システム600はこうして、代替発話信号に対応する発話パラメータによる受信されたデータの第1の複数の発話パラメータの「パラメトリックな」置換によって、処理された発話信号を生成することを可能にし得る。代替発話信号に対応する発話パラメータを使用して処理された信号を生成することは、第1の複数の発話パラメータを使用して処理された信号を生成することよりも、高いオーディオ品質をもたらし得る。たとえば、受信されたデータは、高品質のオーディオ信号を表さないことがある。例示すると、受信されたデータは、騒がしい環境にいる(たとえば、コンサートにおける)ユーザから受け取られたユーザ発話信号に対応することがあり、発話の置換を実行することなく別のデバイスによって生成されることがある。他のデバイスが発話の置換を実行したとしても、他のデバイスによって送られるデータの一部は、たとえばパケットロス、帯域幅制限、および/またはビットレート制限が原因で、受信されないことがある。図2を参照して説明されたように、発話パラメータは、第2のユーザ154が静かな環境において提供した訓練信号から生成された代替発話信号に対応し得る。デコーダは、その発話パラメータを使用して、第1の複数の発話パラメータを使用して生成され得るものより高品質の処理された信号を生成することができる。

10

【0115】

図7を参照すると、デコーダにおいて発話信号の「波形の」置換を実行するように動作可能なシステムの例示的な態様が開示され、全体が700と指定される。システム700は、第1のデータベース124、パラメータ化されたデータベース424、データベース分析器410、デコーダ726、またはこれらの組合せを含み得る。デコーダ726は、信号処理モジュール722を含んでよく、それに結合されてよく、またはそれと通信してよい。

20

【0116】

図7は、受信デバイスにおける発話信号の置換を示す。図6の態様は、受信された発話パラメータを代替発話信号に対応する発話パラメータにより置き換えることで、処理された発話信号を生成することを示すが、図7の態様は、受信された発話パラメータに基づいて代替発話信号を特定し、代替発話信号を使用して処理された発話信号を生成することを示す。

【0117】

動作の間、第1のデバイス102は、第1のデータ630を受信することができる。図6を参照して説明されたように、分析器602は、第1のデータ630に基づいて検索基準612を生成することができる。図4の検索器404を参照して説明されたように、検索器604は、検索基準612に基づいて代替発話信号112のサブセット(たとえば、複数の代替発話信号714)を特定することができる。たとえば、複数の代替発話信号714は、第1の代替発話信号172および第2の代替発話信号174を含み得る。図4の制約分析器406を参照して説明されたように、制約分析器706は、制約に基づいて複数の代替発話信号714の第1の代替発話信号(たとえば、第1の代替発話信号172)を選択することができる。ある特定の態様では、制約分析器706は、図4の制約分析器406を参照して説明されたように、制約に基づいて第1の発話パラメータ472を選択することができ、選択された第1の発話パラメータ472に対応する第1の代替発話信号172を選択することができる。

30

40

【0118】

制約分析器706は、第1の代替発話信号172を合成器708に提供することができる。合成器708は、第1の代替発話信号172に基づいて第1の処理された発話信号116を生成することができる。たとえば、合成器708は、入力信号の一部を第1の代替発話信号172で置き換えることによって、第1の処理された発話信号116を生成することができる。第1のデバイス102は、スピーカー142を介して第1の処理された発話信号116を出力することができる。

【0119】

システム700はこうして、ユーザ発話信号に対応するデータを受信し、ユーザ発話信号の一部を代替発話信号で置き換えることにより処理された信号を生成することによって、「波形の」置換を可能にし得る。処理された信号は、ユーザ発話信号よりも高いオーデ

50

ィオ品質を有し得る。たとえば、受信されたデータは、高品質のオーディオ信号を表さないことがある。例示すると、受信されたデータは、騒がしい環境にいる(たとえば、コンサートにおける)ユーザから受け取られたユーザ発話信号に対応することがあり、発話の置換を実行することなく別のデバイスによって生成されることがある。他のデバイスが発話の置換を実行したとしても、他のデバイスによって送られるデータの一部は、たとえばパケットロス、帯域幅制限、および/またはビットレート制限が原因で、受信されないことがある。図2を参照して説明されたように、第2のユーザ154が静かな環境において提供した訓練信号から、代替発話信号が生成され得る。デコーダは、代替発話信号を使用して、受信されたデータを使用して生成され得るものより高品質の処理された信号を生成することができる。

10

【0120】

図8を参照すると、発話信号の「デバイス中の」置換を実行するように動作可能なシステムの例示的な態様が開示され、全体が800と指定される。システム800は、第1のデータベース124、データベース分析器410、信号処理モジュール422、パラメータ化されたデータベース424、アプリケーション822、またはこれらの組合せを含み得る。

【0121】

動作の間、信号処理モジュール422は、アプリケーション822から入力信号830を受信することができる。入力信号830は、発話信号に対応し得る。アプリケーション822は、ドキュメントビューワアプリケーション、電子ブックビューワアプリケーション、テキスト発話変換アプリケーション、電子メールアプリケーション、コミュニケーションアプリケーション、インターネットアプリケーション、音声記録アプリケーション、またはこれらの組合せを含み得る。図4を参照して説明されたように、分析器402は入力信号830に基づいて検索基準418を生成することができ、検索器404は検索基準418に基づいて検索結果414を生成することができ、制約分析器406は選択された結果416を特定することができ、合成器408は入力信号830の一部分を選択された結果416に基づいて代替発話信号で置き換えることによって第1の処理された発話信号116を生成することができる。第1のデバイス102は、スピーカ142を介して第1の処理された発話信号116を出力することができる。ある特定の態様では、信号処理モジュール422は、第1の処理された発話信号116を生成するための第1のユーザ152からのユーザ要求を受け取ったことに応答して、検索基準418を生成することができる。

20

30

【0122】

ある特定の態様では、第1のデータベース124は代替発話信号の複数のセットを含み得る。代替発話信号の各セットは、特定の有名人、特定のキャラクター(たとえば、漫画のキャラクター、テレビのキャラクター、または映画のキャラクター)、特定のユーザ(たとえば、第1のユーザ152または第2のユーザ154)、またはこれらの組合せに対応し得る。第1のユーザ152は、代替発話信号の特定のセット(たとえば、代替発話信号112)を選択することができる。たとえば、第1のユーザ152は、対応する有名人、対応するキャラクター、対応するユーザ、またはこれらの組合せを選択することができる。本明細書で説明されるように、信号処理モジュール422は、代替発話信号(たとえば、代替発話信号112)の選択されたセットを使用して、第1の処理された発話信号116を生成することができる。

40

【0123】

ある特定の態様では、第1のデバイス102は、要求を別のデバイス(たとえば、図2のサーバ206またはモバイルデバイス104)に送ることができる。この要求は、特定の有名人、特定のキャラクター、特定のユーザ、またはこれらの組合せを特定することができる。他のデバイスは、要求を受信したことに応答して、代替発話信号112を第1のデバイス102に送ることができる。

【0124】

システム800はこうして、アプリケーションによって生成された入力発話信号の一部分を代替発話信号で置き換えることによって、「デバイス中の」置換を可能にし得る。代替発話信号は、入力発話信号よりも高いオーディオ品質を有し得る。ある特定の態様では、

50

代替発話信号は、第1のユーザ152の訓練発話から生成され得る。たとえば、アプリケーションによって生成される入力発話信号は、ロボットのように聞こえることがあり、またはデフォルトの音声に対応することがある。第1のユーザ152は、別の声(たとえば、図1の第1のユーザ152、第2のユーザ154、特定の有名人、特定のキャラクターなどの声)を好むことがある。代替発話信号は、好まれる声に対応し得る。

【0125】

図4~図8の態様はローカルデータベース分析器410を示すが、ある特定の態様では、データベース分析器410またはそのコンポーネントは、サーバ(たとえば、図2のサーバ206)に含まれ得る。この態様では、モバイルデバイス104、第1のデバイス102、またはこれらの両方は、サーバ206からパラメータ化されたデータベース424の少なくとも一部分を受信し得る。たとえば、サーバ206は、代替発話信号112を分析して、第1の発話パラメータ472、第2の発話パラメータ474、N個の発話パラメータ476、またはこれらの組合せを生成することができる。

10

【0126】

サーバ206は、パラメータ化されたデータベース424の少なくともその部分をモバイルデバイス104に、第1のデバイス102に、または両方に提供することができる。たとえば、モバイルデバイス104、第1のデバイス102、または両方が、定期的に、または、図1の第1のユーザ発話信号130、図6の第1のデータ630、入力信号830、ユーザ要求、もしくはこれらの組合せを受信したことに応答してのいずれかで、サーバ206からパラメータ化されたデータベース424の少なくともその部分を要求することができる。ある代替的な態様では、サーバ206は、パラメータ化されたデータベース424の少なくともその部分をモバイルデバイス104に、第1のデバイス102に、または両方に定期的に送ることができる。

20

【0127】

ある特定の態様では、サーバ(たとえば、図2のサーバ206)は信号処理モジュール422を含み得る。サーバ206は、モバイルデバイス104から第1のユーザ発話信号130を受信することができる。サーバ206は、図3および図4を参照して説明されたように、出力信号430を生成することができる。サーバ206は、第1のデバイス102へ出力信号430を送ることができる。ある特定の態様では、出力信号430は図6の第1のデータ630に対応し得る。

【0128】

図9を参照すると、発話信号を置き換えるように動作可能なシステムのある特定の態様の図が示され、全体が900と指定される。ある特定の態様では、システム900は、第1のデバイス102に対応してよく、またはそれに含まれてよい。

30

【0129】

システム900は、合成器408に結合される検索モジュール922を含む。検索モジュール922は、検索器904に結合されたデータベース924を含んでよく、またはそれへのアクセス権を有してよい。ある特定の態様では、データベース924は、図1の第1のデータベース124、図4のパラメータ化されたデータベース424、または両方を含み得る。たとえば、データベース924は、代替発話信号112および/または代替発話信号112に対応する発話パラメータ(たとえば、発話パラメータ472、474、および/または476)のセットを記憶することができる。ある特定の態様では、データベース924は、代替発話信号112および発話パラメータのセットを受信して記憶することができる。たとえば、データベース924は、代替発話信号112および発話パラメータのセットを別のデバイス(たとえば、図2のサーバ206)から受信することができる。別の例として、データベース924は、サーバ206から代替発話信号112を受信することができ、検索器904は、ベクトル量子化器、隠れマルコフモデル(HMM)、または混合ガウスモデル(GMM)の少なくとも1つを使用して、発話パラメータのセットを生成することができる。検索器904は、発話パラメータの生成されたセットをデータベース924に記憶することができる。

40

【0130】

ある特定の態様では、検索器904は図4の分析器402および検索器404を含み得る。合成器408は置換器902を含み得る。置換器902は、検索器904およびデータベース924に結合され

50

得る。

【0131】

動作の間、検索器904は、入力信号930を受信することができる。入力信号930は、発話に対応し得る。ある特定の態様では、入力信号930は、図1および図4～図5の第1のユーザ発話信号130、図6～図7の第1のデータ630、または図8の入力信号830を含み得る。たとえば、検索器904は、ユーザ(たとえば、図1の第1のユーザ152)、別のデバイス、アプリケーション、またはこれらの組合せから入力信号930を受信することができる。アプリケーションは、ドキュメントビューワアプリケーション、電子ブックビューワアプリケーション、テキスト発話変換アプリケーション、電子メールアプリケーション、コミュニケーションアプリケーション、インターネットアプリケーション、音声記録アプリケーション、またはこれらの組合せを含み得る。

10

【0132】

検索器904は、入力信号930を代替発話信号112と比較することができる。ある特定の態様では、検索器904は、入力信号930の第1の複数の発話パラメータを発話パラメータのセットと比較することができる。検索器904は、ベクトル量子化器、隠れマルコフモデル(HMM)、または混合ガウスモデル(GMM)の少なくとも1つを使用して、第1の複数の発話パラメータを決定することができる。

【0133】

検索器904は、この比較に基づいて、特定の代替発話信号(たとえば、第1の代替発話信号172)が入力信号930と一致すると決定することができる。たとえば、図4を参照して説明されたように、検索器904は、第1の代替発話信号172、第1の発話パラメータ472、またはこれらの両方を、第1の複数の発話パラメータと、代替発話信号112に対応する発話パラメータのセット、制約、またはこれらの両方との比較に基づいて選択することができる。

20

【0134】

検索器904は、選択された結果416を置換器902に提供することができる。選択された結果416は、第1の代替発話信号172、第1の発話パラメータ472、または両方を示し得る。置換器902は、選択された結果416に基づいて、第1の代替発話信号172をデータベース924から取り出すことができる。図4を参照して説明されたように、置換器902は、入力信号930の一部を第1の代替発話信号172で置き換えることによって、第1の処理された発話信号116を生成することができる。

30

【0135】

ある特定の態様では、データベース924は、代替発話信号112と関連付けられるラベルを含み得る。たとえば、第1のラベルは、第1の代替発話信号172に対応する音(たとえば、音素、ダイフォンの、トライフォンの、音節、言葉、またはこれらの組合せ)を示し得る。第1のラベルは、音と関連付けられるテキスト識別子を含み得る。ある特定の態様では、置換器902は、選択された結果416に基づいてデータベース924から第1のラベルを取り出すことができる。たとえば、選択された結果416は第1の代替発話信号172を示してよく、置換器902はデータベース924から対応する第1のラベルを取り出すことができる。合成器408は、第1のラベルを含むテキスト出力を生成することができる。

40

【0136】

既存の発話認識システムは、言語モデルを使用し、言葉または音のようなより高次の構造に対して動作し得る。対照的に、検索モジュール922は、言語モデルを使用せず、入力信号930と代替発話信号112の比較を実行して選択された結果416を決定するために、パラメータレベルで、または信号レベルで動作することができる。

【0137】

システム900は、第1の処理された発話信号116を生成するために、第1の代替発話信号172による入力信号930の一部の置換を可能にし得る。第1の処理された発話信号116は、入力信号930よりも高いオーディオ品質を有し得る。

【0138】

図10を参照すると、データベースの特定の態様の図が示されており、全体が1024と指定

50

される。ある特定の態様では、データベース1024は、図1の第1のデータベース124、図4のパラメータ化されたデータベース424、図9のデータベース924、またはこれらの組合せに対応し得る。

【0139】

データベース1024は、メル周波数ケプストラル係数(MFCC)1002、線スペクトル対(LSP)1004、線スペクトル周波数(LSF)1006、ケプストラル1010、線スペクトル情報(LSI)1012、離散コサイン変換(DCT)パラメータ1014、離散フーリエ変換(DFT)パラメータ1016、高速フーリエ変換(FFT)パラメータ1018、フォルマント周波数1020、パルス符号変調(PCM)サンプル1022、またはこれらの組合せを含み得る。

【0140】

ある特定の態様では、デバイス(たとえば、図2のサーバ206)のアナログデジタルコンバータ(ADC)は、訓練発話信号(たとえば、訓練発話信号272)に基づいてPCMサンプル1022を生成することができる。発話信号マネージャ262は、ADCからPCMサンプル1022を受信することができ、PCMサンプル1022をデータベース1024に記憶することができる。図2を参照して説明されたように、発話信号マネージャ262は、PCMサンプル1022に基づいて代替発話信号112を生成することができる。

【0141】

発話信号マネージャ262は、PCMサンプル1022の各々のスペクトルの表現を算出することができる。たとえば、発話信号マネージャ262は、PCMサンプル1022の各々(または代替発話信号112の各々)に対応する、MFCC 1002、LSP 1004、LSF 1006、ケプストラル1010、LSI 1012、DCTパラメータ1014、DFTパラメータ1016、FFTパラメータ1018、フォルマント周波数1020、またはこれらの組合せを生成することができる。

【0142】

たとえば、MFCC 1002は、特定のPCMサンプルに対応する音の短期出力スペクトルの表現であり得る。発話信号マネージャ262は、周波数の非線形なメル尺度上での対数出力スペクトルの線形コサイン変換に基づいて、MFCC 1002を決定することができる。対数出力スペクトルは、特定のPCMサンプルに対応し得る。

【0143】

別の例として、LSP 1004またはLSF 1006は、特定のPCMサンプルに対応する線形予測係数(LPC)の表現であり得る。LPCは、特定のPCMサンプルのスペクトルエンベロープを表し得る。発話信号マネージャ262は、線形予測モデルに基づいて、特定のPCMサンプルのLPCを決定することができる。発話信号マネージャ262は、LPCに基づいて、LSP 1004、LSF 1006、または両方を決定することができる。

【0144】

さらなる例として、ケプストラル1010は、特定のPCMサンプルの出力スペクトルを表し得る。発話信号マネージャ262は、逆フーリエ変換(IFT)を特定のPCMサンプルの推定されるスペクトルの対数に適用することによって、ケプストラル1010を決定することができる。

【0145】

追加の例として、LSI 1012は、特定のPCMサンプルのスペクトルを表し得る。発話信号マネージャ262は、フィルタを特定のPCMサンプルに適用してLSI 1012を生成し得る。

【0146】

発話信号マネージャ262は、特定の離散コサイン変換(DCT)を特定のPCMサンプルに適用してDCTパラメータ1014を生成することができ、特定の離散フーリエ変換(DFT)を特定のPCMサンプルに適用してDFTパラメータ1016を生成することができ、特定の高速フーリエ変換(FFT)を特定のPCMサンプルに適用してFFTパラメータ1018を生成することができ、またはこれらの組合せであり得る。

【0147】

フォルマント周波数1020は、特定のPCMサンプルのスペクトルのスペクトルピークを表し得る。発話信号マネージャ262は、特定のPCMサンプルの位相情報に基づいて、バンドパ

10

20

30

40

50

スフィルタを特定のPCMサンプルに適用することによって、特定のPCMサンプルのLPC分析を実行することによって、またはこれらの組合せで、フォルマント周波数1020を決定することができる。

【0148】

ある特定の態様では、MFCC 1002、LSP 1004、LSF 1006、ケプストラル1010、LSI 1012、DCTパラメータ1014、DFTパラメータ1016、FFTパラメータ1018、フォルマント周波数1020、またはこれらの組合せは、第1の発話パラメータ472に対応してよく、特定のPCMサンプルは、第1の代替発話信号172に対応してよい。

【0149】

データベース1024は、一致する代替発話信号の検索の間に入力発話信号を複数の代替発話信号と比較するために信号処理モジュールによって使用され得る、代替発話信号のパラメータの例を示す。信号処理モジュールは、入力発話信号の一部を一致する代替発話信号で置き換えることによって、処理された信号を生成することができる。処理された信号は、入力発話信号よりも良いオーディオ品質を有し得る。

【0150】

図11を参照すると、システムの特定の態様が開示され、全体が1100と指定される。システム1100は、図1～図2および図4～図9のシステム100～200および400～900を参照して説明された1つまたは複数の動作を実行することができる。

【0151】

システム1100は、ネットワーク120を介して第1のデバイス102に結合された、またはそれと通信しているサーバ1106を含み得る。サーバ1106は、メモリ1176に電氣的に結合されたプロセッサ1160を含み得る。プロセッサ1160は、ネットワーク120へ送受信機1180を介して電氣的に結合され得る。送受信機(たとえば、送受信機1180)は、受信機、送信機、または両方を含み得る。受信機は、アンテナ、ネットワークインターフェース、またはアンテナおよびネットワークインターフェースの組合せの、1つまたは複数を含み得る。送信機は、アンテナ、ネットワークインターフェース、またはアンテナおよびネットワークインターフェースの組合せの、1つまたは複数を含み得る。プロセッサ1160は、発話信号マネージャ262を含んでよく、またはそれに電氣的に結合されてよい。メモリ1176は、第2のデータベース264を含み得る。第2のデータベース264は、特定のユーザ(たとえば、図1の第2のユーザ154)と関連付けられる代替発話信号112を記憶するように構成され得る。たとえば、図2を参照して説明されるように、発話信号マネージャ262は、訓練信号に基づいて代替発話信号112を生成することができる。代替発話信号112は、第1の代替発話信号172を含み得る。

【0152】

メモリ1176は、1つまたは複数の遠隔音声プロファイル1178を記憶するように構成され得る。遠隔音声プロファイル1178は、複数の人物と関連付けられ得る。たとえば、遠隔音声プロファイル1178は、遠隔音声プロファイル1174を含み得る。遠隔音声プロファイル1174は、ある人物(たとえば、図1の第2のユーザ154)と関連付けられ得る。例示すると、遠隔音声プロファイル1174は、第2のユーザ154と関連付けられる識別子1168(たとえば、ユーザ識別子)を含み得る。遠隔音声プロファイル1178の別の遠隔音声プロファイルは、別の人物(たとえば、第1のユーザ152)と関連付けられ得る。

【0153】

遠隔音声プロファイル1174は、代替発話信号112と関連付けられ得る。たとえば、遠隔音声プロファイル1174は、代替発話信号112と関連付けられる発話の内容1170を含み得る。例示すると、発話の内容1170は、第2のユーザ154の発話信号または発話モデルに対応し得る。発話の内容1170は、代替発話信号112の1つまたは複数から抽出される特徴に基づき得る。遠隔音声プロファイル1174は、代替発話信号112が第1の再生時間長を有するオーディオデータに対応することを示し得る。たとえば、代替発話信号112を生成するために使用される訓練信号は、図2を参照して説明されたように、第1の再生時間長を有し得る。

【0154】

10

20

30

40

50

第1のデバイス102は、ネットワーク120へ送受信機1150を介して電氣的に結合されるプロセッサ1152を含み得る。プロセッサ1152は、メモリ1132に電氣的に結合され得る。メモリ1132は、第1のデータベース124を含み得る。第1のデータベース124は、局所代替発話信号1112を記憶するように構成され得る。プロセッサ1152は、信号処理モジュール122を含んでよく、またはそれに電氣的に結合されてよい。メモリ1132は、1つまたは複数の局所音声プロファイル1108を記憶するように構成され得る。たとえば、局所音声プロファイル1108は、特定のユーザ(たとえば、図1の第2のユーザ154)と関連付けられる局所音声プロファイル1104を含み得る。第1のデバイス102は、少なくとも1つのスピーカ(たとえば、スピーカ142)、ディスプレイ1128、入力デバイス1134(たとえば、タッチスクリーン、キーボード、マウス、またはマイクロフォン)、またはこれらの組合せに電氣的に結合されてよく、またはそれらを含んでよい。

10

【0155】

動作の間、信号処理モジュール122は、サーバ1106から遠隔音声プロファイル1178の1つまたは複数の遠隔音声プロファイル(たとえば、遠隔音声プロファイル1174)を受信することができる。信号処理モジュール122は、識別子1168、発話の内容1170、または両方に基づいて、遠隔音声プロファイル1174がある人物(たとえば、図1の第2のユーザ154)と関連付けられると決定することができる。たとえば、信号処理モジュール122は、識別子1168(たとえば、ユーザ識別子)が第2のユーザ154に対応すると決定したことに応答して、遠隔音声プロファイル1174が第2のユーザ154と関連付けられると決定することができる。別の例として、信号処理モジュール122は、発話の内容1170が第2のユーザ154に対応すると決定したことに応答して、遠隔音声プロファイル1174が第2のユーザ154と関連付けられると決定することができる。信号処理モジュール122は、局所代替発話信号1112に基づいて、第2の発話の内容(たとえば、発話モデル)を生成することができる。信号処理モジュール122は、発話の内容1170と第2の発話の内容との差が閾値を満たす(すなわち、それより小さい)と決定したことに応答して、遠隔音声プロファイル1174が第2のユーザ154と関連付けられると決定することができる。たとえば、発話の内容1170は、代替発話信号1112の1つまたは複数と関連付けられる特徴に対応し得る。第2の発話の内容は、第2のユーザ154の発話モデルに対応し得る。信号処理モジュール122は、特徴が発話モデルに対応するかどうかを示す信頼値を決定することができる。信号処理モジュール122は、信頼値が信頼値の閾値を満たす(すなわち、それより高い)と決定したことに応答して、遠隔音声プロファイル1174が第2のユーザ154と関連付けられると決定することができる。

20

30

【0156】

信号処理モジュール122は、局所音声プロファイル1104および遠隔音声プロファイル1174が同じ人物(たとえば、第2のユーザ154)と関連付けられると決定したことに応答して、プロファイル管理のために局所音声プロファイル1104を選択することができる。本明細書で説明されたように、信号処理モジュール122は、局所音声プロファイル1104を選択したことに応答して、グラフィカルユーザインターフェース(GUI)1138を生成して第1のユーザ152によるプロファイル管理を可能にし、更新要求1110をサーバ1106に送って局所音声プロファイル1104を更新し、またはその両方を行うことができる。

【0157】

ある特定の実装形態では、信号処理モジュール122は、プロファイル要求1120をサーバ1106に送ることができる。プロファイル要求1120は、局所音声プロファイル1104、遠隔音声プロファイル1174、または両方を示し得る。たとえば、プロファイル要求1120は、第2のユーザ154と関連付けられる識別子1168を含み得る。発話信号マネージャ262は、プロファイル要求1120を受信したことに応答して、遠隔音声プロファイル1174を第1のデバイス102に送ることができる。

40

【0158】

ある特定の態様では、信号処理モジュール122は、プロファイル要求1120がいつサーバ1106に送られるべきであることを示すユーザ入力1140を、入力デバイス1134を介して受信することができる。信号処理モジュール122は、ユーザ入力1140に基づいてプロファイル要

50

求1120を送ることができる。たとえば、信号処理モジュール122は、ユーザ入力1140を受信し、ユーザ入力1140の受信後に、ある特定の時間において、および/またはある特定の条件が満たされたことに応答して、プロファイル要求1120がサーバ1106に送られるべきであることをユーザ入力1140が示すと決定したことに応答して、プロファイル要求1120をサーバ1106に送ることができる。別の例として、信号処理モジュール122は、プロファイル要求1120の複数のインスタンスをサーバ1106へ定期的に送ることができる。ある特定の態様では、信号処理モジュール122は、第1のデバイス102の特定のアプリケーション(たとえば、プロファイル管理アプリケーション)が有効モードにあると決定したことに応答して、プロファイル要求1120をサーバ1106に送ることができる。

【0159】

信号処理モジュール122は、局所音声プロファイル1104を選択した後にGUI 1138を生成することができる。ある特定の実装形態では、GUI 1138は図3のGUI 300に対応し得る。GUI 1138は、局所音声プロファイル1104の表現、遠隔音声プロファイル1174の表現、または両方を含み得る。たとえば、GUI 1138は、遠隔音声プロファイル1174に対応する代替発話信号112と関連付けられる第1の再生時間長を示し得る。例示すると、図3の第1のバー302、第2のバー314、または第3のバー316は、それぞれ、遠隔音声プロファイル1174が「Randal Hughes」、「Sabrina Sanders」、または「Brett Dean」に対応するとき、第1の再生時間長を示し得る。第1のバー302、第2のバー314、および第3のバー316の各々は、遠隔音声プロファイル1178の別個の遠隔音声プロファイルと関連付けられ得る。たとえば、第1のバー302は遠隔音声プロファイル1174と関連付けられてよく、第2のバー314は遠隔音声プロファイル1178の第2の遠隔音声プロファイルと関連付けられてよく、第3のバー316は遠隔音声プロファイル1178の第3の遠隔音声プロファイルと関連付けられてよい。

【0160】

ある特定の態様では、GUI 1138は、局所音声プロファイル1104に対応する局所代替発話信号1112と関連付けられる第2の再生時間長を示し得る。例示すると、図3の第1のバー302、第2のバー314、または第3のバー316は、それぞれ、局所音声プロファイル1104が「Randal Hughes」、「Sabrina Sanders」、または「Brett Dean」に対応するとき、第2の再生時間長を示し得る。第1のバー302、第2のバー314、および第3のバー316の各々は、局所音声プロファイル1108の別個の局所音声プロファイルと関連付けられ得る。たとえば、第1のバー302は局所音声プロファイル1104と関連付けられてよく、第2のバー314は局所音声プロファイル1108の第2の局所音声プロファイルと関連付けられてよく、第3のバー316は局所音声プロファイル1108の第3の局所音声プロファイルと関連付けられてよい。ある特定の態様では、GUI 1138は、遠隔音声プロファイル1174と関連付けられる第1の再生時間長と、局所音声プロファイル1104と関連付けられる第2の再生時間長とを示してよく、遠隔音声プロファイル1174および局所音声プロファイル1104は同じ人物(たとえば、図1の第2のユーザ154)と関連付けられる。第1のユーザ152は、GUI 1138に基づいて局所音声プロファイル1104を更新するかどうかを決定することができる。たとえば、局所音声プロファイル1104が短い再生時間長に対応すると第1のユーザ152が決定する場合、第1のユーザ152は、遠隔音声プロファイル1174に基づいて、局所音声プロファイル1104を更新すると決めることができる。別の例として、第1のユーザ152は、遠隔音声プロファイル1174が短い再生時間長に対応することをGUI 1138が示すと決定したことに応答して、遠隔音声プロファイル1174を使用して局所音声プロファイル1104を更新するのを控えることによって、更新の頻度を下げることができる。更新の頻度を下げることによって、リソース(たとえば、電力、帯域幅、または両方)を節約することができる。

【0161】

GUI 1138は、音声プロファイル(たとえば、局所音声プロファイル1104または遠隔音声プロファイル1174)と関連付けられる更新要求1110がいつ送られるべきかを指定するためのオプションを示し得る。信号処理モジュール122は、GUI 1138をディスプレイ1128に提供することができる。たとえば、オプションは、自動更新の閾値312または追加オプション308に対応し得る。信号処理モジュール122は、第1のユーザ152から入力デバイス1134を

10

20

30

40

50

介して、オプションに対応するユーザ入力1140を受け取ることができる。ユーザ入力1140は、更新要求1110がいつ送られるべきかを示し得る。たとえば、第1のユーザ152は、自動更新の閾値312を上げ、または下げ得る。信号処理モジュール122は、第1の再生時間長が自動更新の閾値312を満たす(たとえば、それ以上である)と決定したことに応答して、更新要求1110を送ることができる。例示すると、信号処理モジュール122は、代替発話信号112に対応する第1の再生時間長が自動更新の閾値312を満たすことを遠隔音声プロファイル1174が示すと決定したことに応答して、更新要求1110を送ることができる。信号処理モジュール122はこうして、第1の再生時間長が自動更新の閾値312を満たさない(たとえば、それ未満である)ときに更新要求1110を送るのを控えることによって、更新の頻度を下げることができる。別の例として、第1のユーザ152は、音声プロファイル(たとえば、遠隔音声プロファイル1174、局所音声プロファイル1104、または両方)と関連付けられる特定のパー(たとえば、第1のパー302)を選択することができ、追加オプション308を選択することができる。信号処理モジュール122は、追加オプション308および第1のパー302の選択を示すユーザ入力1140を受け取ったことに応答して、音声プロファイル(たとえば、遠隔音声プロファイル1174、局所音声プロファイル1104、または両方)を示す更新要求1110をサーバ1106に送ることができる。

10

【 0 1 6 2 】

ある特定の実装形態では、GUI 1138は、音声プロファイル(たとえば、遠隔音声プロファイル1174、局所音声プロファイル1104、または両方)と関連付けられる更新要求(たとえば、更新要求1110)がサーバ1106に定期的に(たとえば、毎日、毎週、または毎月)送られるべきであることを指定するためのオプションを含み得る。信号処理モジュール122は、たとえば、更新要求が定期的に送られるべきであることを示すユーザ入力1140を受け取ったことに応答して、更新要求1110の複数のインスタンスをサーバ1106へ定期的に送ることができる。ある特定の態様では、ユーザ入力1140は更新時間の閾値を示し得る。信号処理モジュール122は、音声プロファイル(たとえば、遠隔音声プロファイル1174、局所音声プロファイル1104、または両方)と関連付けられる第1の更新が第1の時間においてサーバ1106から受信されたと決定することができる。信号処理モジュール122は、第2の時間において、第1の時間と第2の時間との差を決定することができる。信号処理モジュール122は、この差がユーザ入力1140によって示される更新時間の閾値を満たすと決定したことに応答して、更新要求1110をサーバ1106に送ることができる。

20

30

【 0 1 6 3 】

ある特定の実装形態では、GUI 1138は、音声プロファイル(たとえば、遠隔音声プロファイル1174、局所音声プロファイル1104、または両方)と関連付けられる更新要求(たとえば、更新要求1110)を送るのをいつ控えるべきかを指定するためのオプションを含み得る。このオプションは、更新停止の閾値、リソース使用の閾値、または両方に対応し得る。たとえば、信号処理モジュール122は、局所代替発話信号1112と関連付けられる第2の再生時間長が更新停止の閾値を満たす(たとえば、それ以上である)と決定したことに応答して、更新要求1110をサーバ1106に送るのを控えることができる。例示すると、第1のユーザ152は、局所代替発話信号1112に対応する第2の再生時間長が更新停止の閾値を満たす(たとえば、それ以上である)ときに、信号処理モジュール122が局所音声プロファイル1104を自動的に更新するのを止めるように、更新停止の閾値を規定することができる。別の例として、信号処理モジュール122は、リソース使用(たとえば、残っているバッテリー電力、利用可能なメモリ、課金サイクルのネットワーク使用量、またはこれらの組合せ)がリソース使用の閾値を満たすと決定したことに応答して、更新要求1110をサーバ1106に送るのを控えることができる。たとえば、信号処理モジュール122は、残りのバッテリー電力が電力節約の閾値を満たす(たとえば、それ未満である)とき、利用可能なメモリがメモリ節約の閾値を満たす(たとえば、それ未満である)とき、課金サイクルのネットワーク使用量がネットワーク使用量の閾値を満たす(たとえば、それ以上である)とき、またはこれらの組合せのとき、更新要求1110をサーバ1106に送るのを控えることができる。

40

【 0 1 6 4 】

50

ある特定の態様では、GUI 1138は削除オプション310を含み得る。第1のユーザ152は、第1のバー302および削除オプション310を選択することができる。信号処理モジュール122は、削除オプション310および第1のバー302の選択を示すユーザ入力1140を受け取ったことに応答して、対応する音声プロファイル(たとえば、遠隔音声プロファイル1174)を示す削除要求をサーバ1106に送ることができる。発話信号マネージャ262は、削除要求を受信したことに応答して代替発話信号112を削除することができる。ある特定の実装形態では、信号処理モジュール122は、削除オプション310および第1のバー302の選択を示すユーザ入力1140を受け取ったことに応答して、局所代替発話信号1112を削除することができる。

【0165】

本明細書で説明されるように、発話信号マネージャ262は、更新要求1110を受信したことに応答して、更新1102を第1のデバイス102に送ることができる。発話信号マネージャ262は、更新1102が第1のデバイス102にいつ送られるかを示すために、遠隔音声プロファイル1178および第1のデバイス102と関連付けられる最終更新送信時間を更新することができる。ある特定の实装形態では、発話信号マネージャ262は、第1の時間において、削除要求を受信したことに応答して、第1の時間を示すように第1のデバイス102および遠隔音声プロファイル1178と関連付けられる最終更新送信時間を更新することができる。たとえば、発話信号マネージャ262は、遠隔音声プロファイル1178と関連付けられる更新を別のデバイスに提供するために代替発話信号112を保持しながら、本明細書で説明されるように、代替発話信号112が第1のデバイス102に対する後続の更新から除外されるように、第1のデバイス102に更新を送ることなく最終更新送信時間を更新することができる。

【0166】

信号処理モジュール122は、更新1102をサーバ1106から受信することができる。更新1102は、局所音声プロファイル1104、遠隔音声プロファイル1174、または両方を示し得る。たとえば、更新1102は、図1の第2のユーザ154に対応する識別子1168を含み得る。更新1102は、代替発話信号112に対応する発話データを含み得る。信号処理モジュール122は、局所代替発話信号1112に代替発話信号112の各々を追加することができる。たとえば、信号処理モジュール122は、発話データを局所音声プロファイル1104に追加することができる。別の例として、信号処理モジュール122は、発話データに基づいて代替発話信号112を生成することができ、生成された代替発話信号112の各々を局所代替発話信号1112に追加することができる。

【0167】

ある特定の態様では、信号処理モジュール122は、局所音声プロファイル1104の第1のセグメント1164を、遠隔音声プロファイル1174の第2のセグメント1166で置き換えることができる。たとえば、第1のセグメント1164は、局所代替発話信号1112の代替発話信号1172に対応し得る。第2のセグメント1166は、代替発話信号112の第1の代替発話信号172に対応し得る。信号処理モジュール122は、局所代替発話信号1112から代替発話信号1172を除去することができ、第1の代替発話信号172を局所代替発話信号1112に追加することができる。

【0168】

第1の代替発話信号172および代替発話信号1172は、類似する音(たとえば、音素、ダイフオン、トライフオン、音節、言葉、またはこれらの組合せ)に対応し得る。ある特定の態様では、信号処理モジュール122は、第1の代替発話信号172が代替発話信号1172よりも高いオーディオ品質(たとえば、より高い信号対雑音比)を有すると決定したことに応答して、代替発話信号1172を第1の代替発話信号172で置き換えることができる。別の態様では、信号処理モジュール122は、代替発話信号1172が期限切れになった(たとえば、期限切れの閾値を超えるタイムスタンプを有する)と決定したことに応答して、代替発話信号1172を第1の代替発話信号172で置き換えることができる。

【0169】

ある特定の实装形態では、発話信号マネージャ262は、第1の時間において、遠隔音声プロファイル1174と関連付けられる第1の更新を第1のデバイス102に送ることができる。発

10

20

30

40

50

話信号マネージャ262は、第1の時間を示すように最終更新送信時間を更新することができる。発話信号マネージャ262はその後、遠隔音声プロファイル1174と関連付けられる代替発話信号112が第1のデバイス102に第1の更新を送った後に生成されたと決定することができる。たとえば、発話信号マネージャ262は、代替発話信号112の各々に対応するタイムスタンプが最終更新送信時間よりも後の時間を示すと決定したことに応答して、代替発話信号112を選択することができる。発話信号マネージャ262は、代替発話信号112が第1のデバイス102に第1の更新を送った後に生成されたと決定したことに応答して、更新1102を第1のデバイス102に送ることができる。この特定の実装形態では、更新1102は、遠隔音声プロファイル1174と関連付けられるすべての代替発話信号よりも少数のものを含み得る。たとえば、更新1102は、サーバ1106が以前に第1のデバイス102へ送っていない代替発話信号112だけを含み得る。

10

【0170】

ある特定の態様では、更新要求1110は、すべての代替発話信号が要求されるかどうか、または、以前に送られていない代替発話信号だけが要求されるかどうかを示し得る。たとえば、信号処理モジュール122は、第1の時間において第1の更新を第1のデバイス102に送ることができる。第1の時間を示すように最終更新送信時間を更新することができる。第2のデータベース264は、第1の時間よりも前に生成された遠隔音声プロファイル1174と関連付けられる第2の代替発話信号を含み得る。代替発話信号112は、第1の時間の後に生成され得る。

【0171】

信号処理モジュール122は、局所音声プロファイル1104と関連付けられるすべての代替発話信号を要求するために、更新要求1110を送ることができる。ある特定の態様では、信号処理モジュール122は、局所音声プロファイル1104と関連付けられる発話データが第1のデバイス102において削除されたと決定したことに応答して、更新要求1110を送ることができる。発話信号マネージャ262は、更新要求1110を受信することができる。発話信号マネージャ262は、更新要求1110が第2のユーザ154と関連付けられる識別子1168を含むと決定したことに応答して、更新要求1110が遠隔音声プロファイル1174に対応すると決定することができる。発話信号マネージャ262は、すべての対応する代替発話信号が要求されることを更新要求1110が示すと決定したことに応答して、遠隔音声プロファイル1174と関連付けられるすべての代替発話信号を含む更新1102を送ることができる。代替的に、発話信号マネージャ262は、以前に第1のデバイス102に送られていない代替発話信号だけが要求されることを更新要求1110が示すと決定したことに応答して、遠隔音声プロファイル1174と関連付けられるすべての代替発話信号よりも少数のものを含む更新1102を送ることができる。たとえば、更新1102は、最終更新送信時間の後の時間(たとえば、生成時間)を示すタイムスタンプを有する代替発話信号112を含み得る。

20

30

【0172】

ある特定の態様では、発話信号マネージャ262は、更新要求1110を受信したこととは独立に、更新1102を第1のデバイス102に送ることができる。たとえば、発話信号マネージャ262は、更新1102を定期的を送ることができる。別の例として、発話信号マネージャ262は、最終更新送信時間を超える時間(たとえば、生成時間)を示すタイムスタンプを有する代替発話信号(たとえば、代替発話信号112)と関連付けられる再生時間長を決定することができる。発話信号マネージャ262は、再生時間長が自動更新の閾値を満たすと決定したことに応答して、代替発話信号112を送ることができる。

40

【0173】

局所代替発話信号1112は、本明細書で説明されるように、処理された発話信号を生成するために使用され得る。第1のユーザ152は、局所音声プロファイル1104を示す選択1136を第1のデバイス102に提供することができる。たとえば、第1のデバイス102は、入力デバイス1134を介して選択1136を受け取ることができる。ある特定の態様では、選択1136は、局所音声プロファイル1104と関連付けられる第2のデバイスとの音声呼に対応し得る。たとえば、局所音声プロファイル1104は、図1の第2のユーザ154のユーザプロファイルと関連

50

付けられ得る。ユーザプロファイルは、第2のデバイスを示し得る。第1のユーザ152は、音声呼を開始することがあり、または、第2のデバイスからの音声呼を受け入れることがある。信号処理モジュール122は、選択1136を受け取った後に入力オーディオ信号1130を受信することがある。たとえば、信号処理モジュール122は、音声呼の間に入力オーディオ信号1130を受信することがある。信号処理モジュール122は、入力オーディオ信号1130の第1の部分1162を局所代替発話信号1112と比較することができる。局所代替発話信号1112は、第1の代替発話信号172を含み得る。信号処理モジュール122は、第1の部分1162が第1の代替発話信号172と一致する(たとえば、それに類似する)と決定することができる。図1を参照して説明されたように、信号処理モジュール122は、第1の部分1162が第1の代替発話信号172と一致すると決定したことに応答して、第1の部分1162を第1の代替発話信号172で置き換えることによって、第1の処理された発話信号116を生成することができる。第1の代替発話信号172は、第1の部分1162よりも高いオーディオ品質を有し得る。信号処理モジュール122は、第1のユーザ152へ、スピーカ142を介して第1の処理された発話信号116を出力し得る。

10

【0174】

ある特定の実装形態では、サーバ1106は、局所音声プロファイル1104と関連付けられるデバイス(たとえば、図1のモバイルデバイス104)に対応し得る。たとえば、サーバ1106によって実行されるものとして本明細書で説明される動作の1つまたは複数は、モバイルデバイス104によって実行され得る。この特定の実装形態では、第1のデバイス102によってサーバ1106へ送られるものとして本明細書で説明される1つまたは複数のメッセージ(たとえば、更新要求1110、プロファイル要求1120、または両方)が、モバイルデバイス104へ送られ得る。同様に、第1のデバイス102によってサーバ1106から受信されるものとして本明細書で説明される1つまたは複数のメッセージ(たとえば、更新1102、遠隔音声プロファイル1174、または両方)が、モバイルデバイス104から受信され得る。

20

【0175】

システム1100は、処理された発話信号をユーザに提供することによってユーザ体験を向上させることができ、ここで処理された発話信号は入力オーディオ信号の一部を代替発話信号で置き換えることによって生成され、代替発話信号は入力オーディオ信号のその一部分よりも高いオーディオ品質を有する。

システム1100はまた、音声プロファイルをユーザが管理することを可能にするために、GUIを提供することができる。

30

【0176】

図12を参照すると、システムの特定の態様が開示され、全体が1200と指定される。システム1200は、図1～図2、図4～図9、および図11のシステム100～200、400～900、および1100を参照して説明された1つまたは複数の動作を実行することができる。システム1200は、図11のシステム1100の1つまたは複数のコンポーネントを含み得る。たとえば、システム1200は、ネットワーク120を介して第1のデバイス102に結合された、またはそれと通信しているサーバ1106を含み得る。第1のデバイス102は、1つまたは複数のマイクロフォン1244に結合されてよく、またはそれを含んでよい。第1のデバイス102は、図1の第1のデータベース124を含み得る。第1のデータベース124は、1つまたは複数の局所代替発話信号1214を記憶するように構成され得る。局所代替発話信号1214は、ある人物(たとえば、第1のユーザ152)と関連付けられる局所音声プロファイル1204に対応し得る。

40

【0177】

動作の間、信号処理モジュール122は、第1のユーザ152からマイクロフォン1244を介して、訓練発話信号(たとえば、入力オーディオ信号1238)を受信することがある。たとえば、信号処理モジュール122は、音声呼の間に入力オーディオ信号1238を受信することがある。信号処理モジュール122は、図2を参照して説明されたように、入力オーディオ信号1238に基づいて代替発話信号1212を生成することができる。信号処理モジュール122は、局所代替発話信号1214に代替発話信号1212を追加することができる。信号処理モジュール122は、更新1202をサーバ206に提供することができる。更新1202は、代替発話信号1212に対

50

応する発話データを含み得る。更新1202は、第1のユーザ152と関連付けられる識別子、局所音声プロファイル1204、または両方を含み得る。サーバ206は更新1202を受信することができ、発話信号マネージャ262は、更新1202の識別子に基づいて、更新1202が遠隔音声プロファイル1178の遠隔音声プロファイル1274と関連付けられると決定することができる。

【0178】

ある特定の実装形態では、発話信号マネージャ262は、遠隔音声プロファイル1178のいずれもが更新1202に対応しないと決定することができる。それに応答して、発話信号マネージャ262は、遠隔音声プロファイル1274を遠隔音声プロファイル1178に追加することができる。遠隔音声プロファイル1274は、更新1202の識別子と関連付けられる(たとえば、それを含む)ことがある。発話信号マネージャ262は、第2のデータベース264に代替発話信号1212を記憶することができ、記憶された代替発話信号1212を遠隔音声プロファイル1274と関連付けることができる。たとえば、発話信号マネージャ262は、発話データに基づいて代替発話信号1212を生成することができ、代替発話信号1212、発話データ、または両方を第2のデータベース264に記憶することができる。

10

【0179】

発話信号マネージャ262は、代替発話信号1212と関連付けられるタイムスタンプ(たとえば、生成タイムスタンプ)を決定することができ、第2のデータベース264、メモリ1176、または両方にタイムスタンプを記憶することができる。たとえば、タイムスタンプは、入力オーディオ信号1238が第1のデバイス102によって受信される時間、代替発話信号1212が信号処理モジュール122によって生成される時間、更新1202が第1のデバイス102によって送られる時間、更新1202がサーバ1106によって受信される時間、代替発話信号1212が第2のデータベース264に記憶される時間、またはこれらの組合せを示し得る。発話信号マネージャ262は、代替発話信号1212と関連付けられる再生時間長を決定することができ、第2のデータベース264、メモリ1176、または両方に再生時間長を記憶することができる。たとえば、再生時間長は、入力オーディオ信号1238の再生時間長であり得る。ある特定の实装形態では、更新1202は再生時間長を示すことができ、発話信号マネージャ262は更新1202に基づいて再生時間長を決定することができる。

20

【0180】

ある特定の態様では、第1のデバイス102は、サーバ1106または別のデバイスから更新要求1210を受信したことに応答して、更新1202をサーバ1106に送ることができる。たとえば、サーバ1106は、定期的に、またはある事象(たとえば、音声呼の開始)を検出したことに応答して、更新要求1210を第1のデバイス102に送ることができる。信号処理モジュール122は、更新要求1210を受信したことに応答して更新1202を送ることができる。

30

【0181】

ある特定の態様では、信号処理モジュール122は、第1のデバイス102のユーザ(たとえば、第1のユーザ152)がサーバ1106への代替発話信号1212(または発話データ)の送信を認めていると決定したことに応答して、更新1202を送ることができる。たとえば、信号処理モジュール122は、第1のユーザ152から入力デバイス1134を介して、ユーザ入力1240を受け取ることができる。信号処理モジュール122は、代替発話信号1212がサーバ1106へ送られることを第1のユーザ152が認めていることをユーザ入力1240が示すと決定したことに応答して、更新1202をサーバ1106に送ることができる。

40

【0182】

ある特定の態様では、信号処理モジュール122は、特定のユーザプロファイル(たとえば、図1の第2のユーザ154のユーザプロファイル)と関連付けられる1つまたは複数のデバイスに代替発話信号1212(または発話データ)が送られることを第1のデバイス102のユーザ(たとえば、第1のユーザ152)が認めていると、決定することができる。更新1202は、代替発話信号1212を共有することが認められている、1つまたは複数のデバイス、特定のユーザプロファイル、またはこれらの組合せを示し得る。信号処理モジュール122は、1つまたは複数のデバイス、特定のユーザプロファイル、またはこれらの組合せを示す認証データ

50

を、第2のデータベース264、メモリ1176、または両方に記憶することができる。図11を参照して説明されたように、信号処理モジュール122は、認証データが特定のデバイスを示すと決定したことに応答して、代替発話信号1212と関連付けられるデータを含む更新を特定のデバイスに送ることができる。

【0183】

ある特定の態様では、信号処理モジュール122は、更新1202をサーバ1106にいつ送るかを選択するためのオプションを示すGUI 1232を生成することができる。信号処理モジュール122は、GUI 1232をディスプレイ1128に提供することができる。信号処理モジュール122は、第1のユーザ152から入力デバイス1134を介して、ユーザ入力1240を受け取ることができる。信号処理モジュール122は、ユーザ入力1240を受信したことに応答して更新1202を送ることができる。ある特定の態様では、ユーザ入力1240は、更新1202(たとえば、代替発話信号1212または発話データ)が定期的に(たとえば、毎時間、毎日、毎週、または毎月)送られるべきであることを示すことがあり、または、更新の閾値(たとえば、4時間)を示すことがある。それに応答して、信号処理モジュール122は、定期的に、または更新の閾値に基づいて、更新1202の複数のインスタンスをサーバ1106へ送ることができる。

10

【0184】

システム1200は、ユーザと関連付けられる代替発話信号をあるデバイスが別のデバイスに提供することを可能にし得る。他のデバイスは、ユーザに対応する入力オーディオ信号に基づいて、処理された発話信号を生成することができる。たとえば、他のデバイスは、入力オーディオ信号の一部分を代替発話信号で置き換えることによって、処理された発話信号を生成することができる。代替発話信号は、入力オーディオ信号のその一部分よりも高いオーディオ品質を有し得る。

20

【0185】

図13を参照すると、システムの特定の態様が開示され、全体が1300と指定される。システム1300は、図1～図2、図4～図9、および図11～図12のシステム100～200、400～900、および1100～1200を参照して説明された1つまたは複数の動作を実行することができる。

【0186】

システム1300は、図11のシステム1100、図12のシステム1200、または両方の1つまたは複数のコンポーネントを含み得る。たとえば、システム1300は、第1のデバイス102およびプロセッサ1152を含む。第1のデバイス102は、マイクロフォン1244、スピーカ142、または両方に結合され得る。プロセッサ1152は、信号処理モジュール122に結合されてよく、またはそれを含んでよい。信号処理モジュール122は、モード検出器1302、領域選択器1304、テキスト発話変換器1322、またはこれらの組合せに結合され得る。

30

【0187】

システム1300の1つまたは複数のコンポーネントは、車両、電子書籍(e書籍)デバイス、モバイルデバイス、音響キャプチャデバイス、テレビジョン、通信デバイス、タブレット、スマートフォン、ナビゲーションデバイス、ラップトップ、ウェアラブルデバイス、またはコンピューティングデバイスの少なくとも1つに含まれ得る。たとえば、スピーカ142の1つまたは複数は、車両またはモバイルデバイスに含まれ得る。モバイルデバイスは、電子書籍デバイス、タブレット、通信デバイス、スマートフォン、ナビゲーションデバイス、ラップトップ、ウェアラブルデバイス、コンピューティングデバイス、またはこれらの組合せを含み得る。メモリ1132は、車両、電子書籍デバイス、モバイルデバイス、音響キャプチャデバイス、テレビジョン、通信デバイス、タブレット、スマートフォン、ナビゲーションデバイス、ラップトップ、ウェアラブルデバイス、またはコンピューティングデバイスの少なくとも1つによってアクセス可能な記憶デバイスに含まれ得る。

40

【0188】

動作の間、信号処理モジュール122は、第1のユーザ152からマイクロフォン1244を介して、訓練発話信号1338を受信することができる。ある特定の態様では、訓練発話信号1338はテキスト依存であることがあり、第1のデバイス102の一般的なまたは「普通の」使用の間に受信されることがある。たとえば、第1のユーザ152は、信号処理モジュール122の「常

50

時オン」キャプチャモードを有効にすることができ、信号処理モジュール122は、第1のユーザ152がマイクロフォン1244の近くで話すにつれてバックグラウンドの訓練発話信号1338を受信することができる。信号処理モジュール122は、図2および図12を参照して説明されたように、訓練発話信号1338に基づいて代替発話信号1212を生成することができる。代替発話信号1212は、第1の代替発話信号1372を含み得る。図12を参照して説明されたように、信号処理モジュール122は、代替発話信号112、代替発話信号112と関連付けられる発話データ、または両方を、局所代替発話信号1214に追加することができる。

【0189】

モード検出器1302は、複数の使用モード(たとえば、読書モード、会話モード、歌唱モード、命令および制御モード、またはこれらの組合せ)のうちの使用モード1312を検出するように構成され得る。使用モード1312は、訓練発話信号1338と関連付けられ得る。たとえば、モード検出器1302は、訓練発話信号1338が受信されるときに信号処理モジュール122の使用モードの設定1362に基づいて、使用モード1312を検出することができる。例示すると、信号処理モジュール122は、第1のデバイス102の第1のアプリケーション(たとえば、書籍アプリケーションまたは別の読書アプリケーション)が有効にされていることを検出したことに応答して、使用モードの設定1362を読書モードに設定することができる。信号処理モジュール122は、進行中の音声呼を検出したこと、または第1のデバイス102の第2のアプリケーション(たとえば、オーディオチャットアプリケーションおよび/またはビデオ会議アプリケーション)が有効にされていることを検出したことに応答して、使用モードの設定1362を会話モードに設定することができる。

【0190】

信号処理モジュール122は、第1のデバイス102の第3のアプリケーション(たとえば、歌唱アプリケーション)が有効にされていると決定したことに応答して、使用モードの設定1362を歌唱モードに設定することができる。ある特定の態様では、信号処理モジュール122は、訓練発話信号1338が歌唱に対応すると決定したことに応答して、使用モードの設定1362を歌唱モードに設定することができる。たとえば、信号処理モジュール122は、訓練発話信号1338の特徴を抽出することができ、抽出された特徴および分類器(たとえば、サポートベクターマシン(SVM))に基づいて、訓練発話信号1338が歌唱に対応すると決定することができる。分類器は、有声の発話の期間、有声の発話の割合、発話の中断の割合、複数の有声の発話間隔の間の類似性の尺度(たとえば、音節に対応する)、またはこれらの組合せを分析することによって、歌唱を検出することができる。

【0191】

信号処理モジュール122は、第1のデバイス102の第4のアプリケーション(たとえば、パーソナルアシスタントアプリケーション)が有効にされていると決定したことに応答して、使用モードの設定1362を命令および制御モードに設定することができる。ある特定の態様では、信号処理モジュール122は、訓練発話信号1338が命令のキーワード(たとえば、「起動する」)を含むと決定したことに応答して、使用モードの設定1362を命令および制御モードに設定することができる。たとえば、信号処理モジュール122は、発話分析技法を使用して、訓練発話信号1338が命令のキーワードを含むと決定することができる。読書モード、会話モード、歌唱モード、または命令および制御モードの1つが、「デフォルトの」使用モードに対応し得る。使用モードの設定1362は、無効にされない限りデフォルトの使用モードに対応し得る。ある特定の態様では、信号処理モジュール122は、使用モードの設定1362をユーザ入力によって示される特定の使用モードに設定することができる。モード検出器1302は、使用モード1312を信号処理モジュール122に提供することができる。

【0192】

領域選択器1304は、複数の属性領域(たとえば、言語領域、性別領域、年齢領域、またはこれらの組合せ)のうちの属性領域1314を検出するように構成され得る。属性領域1314は、訓練発話信号1338と関連付けられ得る。たとえば、領域選択器1304は、第1のユーザ152と関連付けられる属性データ(たとえば、特定の言語、特定の性別、特定の年齢、またはこれらの組合せ)に基づいて属性領域1314を検出することができる。ある特定の態様で

は、メモリ1132は、第1のユーザ152と関連付けられるユーザプロフィールに属性データを含め得る。別の例として、領域選択器1304は、分類器を使用して訓練発話信号1338を分析することによって、属性領域1314を検出することができる。例示すると、分類器は、訓練発話信号1338を、特定の言語、特定の性別、特定の年齢、またはこれらの組合せへと分類することができる。属性領域1314は、特定の言語、特定の性別、特定の年齢、またはこれらの組合せを示すことができる。領域選択器1304は、属性領域1314を信号処理モジュール122に提供することができる。

【0193】

信号処理モジュール122は、使用モード1312、属性領域1314、または両方を、代替発話信号1212の各々と関連付けることができる。信号処理モジュール122は、代替発話信号1212と関連付けられる分類1346を決定することができる。分類1346は、代替発話信号1212が信号処理モジュール122によって生成されるときに時間および/または日付を示すタイムスタンプを含み得る。分類1346は、訓練発話信号1338が第1のデバイス102によって受信されたときに有効にされたアプリケーション(たとえば、読書アプリケーション)またはアプリケーションの構成要素(たとえば、特定の話)を示し得る。分類1346は、訓練発話信号1338に対応する言葉、語句、テキスト、またはこれらの組合せを示し得る。信号処理モジュール122は、発話分析技法を使用して、訓練発話信号1338を1つまたは複数の感情(たとえば、喜び、怒り、悲しみなど)に分類することができる。分類1346は、1つまたは複数の感情を示し得る。信号処理モジュール122は、分類1346を代替発話信号1212の各々と関連付けることができる。

10

20

【0194】

図2を参照して説明されたように、局所代替発話信号1214は、第2の訓練発話信号に基づいて信号処理モジュール122によって生成される1つまたは複数の追加の代替発話信号(たとえば、第2の代替発話信号1374)を含み得る。第2の代替発話信号1374は、第2の使用モード、第2の属性領域、第2の分類、またはこれらの組合せと関連付けられ得る。

【0195】

信号処理モジュール122は、テキスト発話変換器1322から発話信号1330を受け取ることができる。たとえば、第1のユーザ152はテキスト発話変換器1322を有効にすることができる。テキスト発話変換器1322は発話信号1330を信号処理モジュール122に提供することができる。本明細書で説明されるように、信号処理モジュール122は、ユーザモードの設定1362に基づいて特定の使用モードを検出することができる。信号処理モジュール122はまた、特定の属性領域を選択することができる。たとえば、信号処理モジュール122は、特定のユーザプロフィールを示すユーザ入力を第1のユーザ152から受け取ることができる。信号処理モジュール122は、特定のユーザプロフィールに対応する属性データに基づいて、特定の属性領域を選択することができる。別の例として、ユーザ入力は、特定の言語、特定の年齢、特定の性別、またはこれらの組合せを示し得る。信号処理モジュール122は、特定の言語、特定の年齢、特定の性別、またはこれらの組合せに対応する特定の属性領域を選択することができる。

30

【0196】

信号処理モジュール122は、検出された特定の使用モード、選択された特定の属性領域、または両方に基づいて、局所代替発話信号1214の特定の代替発話信号(たとえば、第1の代替発話信号1372または第2の代替発話信号1374)を選択することができる。たとえば、信号処理モジュール122は、発話信号1330の一部分を局所代替発話信号1214の各々と比較することができる。第1の代替発話信号1372、第2の代替発話信号1374、および発話信号1330のその部分の各々は、類似する音に対応し得る。信号処理モジュール122は、特定の使用モードが使用モード1312に対応する(たとえば、一致する)と決定したこと、特定の属性領域が属性領域1314に対応する(たとえば、少なくとも一部一致する)と決定したこと、または両方に基づいて、第1の代替発話信号1372を選択することができる。したがって、使用モードおよび属性領域は、発話信号1330のその部分が、同様の状況において生成された代替発話信号(たとえば、第1の代替発話信号1372)によって置き換えられることを確実にす

40

50

るために使用され得る。たとえば、発話信号1330が歌唱モードに対応するとき、発話信号1330のその部分は、より高いオーディオ品質の会話の発話ではなくより高いオーディオ品質の歌唱に対応する代替発話信号によって置き換えられ得る。

【0197】

ある特定の態様では、局所代替発話信号1214は分類と関連付けられ得る。分類の各々はある優先度を有し得る。信号処理モジュール122は、発話信号1330のその部分を優先度の順番で局所代替発話信号1214と比較することによって、第1の代替発話信号1372を選択することができる。たとえば、信号処理モジュール122は、分類1346によって示される第1のタイムスタンプが第2の代替発話信号1374と関連付けられる第2の分類によって示される第2のタイムスタンプよりも後であると決定したことに少なくとも一部基づいて、第1の代替発話信号1372を選択することができる。別の例として、信号処理モジュール122は、特定のアプリケーション(たとえば、電子書籍アプリケーション)が有効にされていることと、その特定のアプリケーションが分類1346によって示され第2の分類によって示されないことを決定したことに少なくとも一部基づいて、第1の代替発話信号1372を選択することができる。

10

【0198】

信号処理モジュール122は、発話信号1330のその部分を第1の代替発話信号1372で置き換えることによって、第1の処理された発話信号116を生成することができる。信号処理モジュール122は、第1の処理された発話信号116をメモリ1132に記憶することができる。代替的に、または加えて、信号処理モジュール122は、第1のユーザ152へ、スピーカー142を介して第1の処理された発話信号116を出力し得る。

20

【0199】

ある特定の態様では、訓練発話信号1338は特定のユーザ(たとえば、第1のユーザ152)から受け取られることがあり、第1の処理された発話信号116は同じユーザ(たとえば、第1のユーザ152)に出力されることがある。たとえば、第1の処理された発話信号116は、第1のユーザ152の発話と同じように聞こえ得る。

【0200】

別の態様では、訓練発話信号1338は特定のユーザ(たとえば、親)から受け取られることがあり、第1の処理された発話信号116は別のユーザ(たとえば、その親の子供)に出力されることがある。たとえば、親は、最初の日に電子書籍アプリケーションを有効にして特定の話を読むことがある。訓練発話信号1338は、その親の発話に対応し得る。信号処理モジュール122は、図2および図12を参照して説明されたように、訓練発話信号1338に基づいて代替発話信号1212を生成することができる。信号処理モジュール122は、使用モード1312(たとえば、読書モード)を、代替発話信号1212の各々と関連付けることができる。信号処理モジュール122は、属性領域1314(たとえば、大人、女性、および英語)を、代替発話信号1212の各々と関連付けることができる。信号処理モジュール122は、分類1346(たとえば、特定の話、電子書籍アプリケーション、最初の日に対応するタイムスタンプ、訓練発話信号に対応するテキスト、および訓練発話信号1338に対応する1つまたは複数の感情)を、代替発話信号1212の各々と関連付けることができる。

30

【0201】

他のユーザ(たとえば、子供)が、2日目にテキスト発話変換器1322を有効にすることができる。たとえば、子供は、特定の話を選択することがあり、親のユーザプロファイルを選択することがある。例示すると、その子供は、その特定の話の、親が以前に読んだ部分よりも後の部分を選択することがある。テキスト発話変換器1322は、その特定の話のテキストに基づいて、発話信号1330を生成し得る。信号処理モジュール122は、発話信号1330を受信することができる。信号処理モジュール122は、特定の使用モード(たとえば、読書モード)を検出することができ、親のユーザプロファイルに対応する特定の属性領域(たとえば、大人、女性、および英語)を選択することができる。信号処理モジュール122は、特定の分類(たとえば、特定の話)を決定することができる。信号処理モジュール122は、発話信号1330の一部分を優先度の順番で局所代替発話信号1214と比較することができる。信号

40

50

処理モジュール122は、使用モード1312が特定の使用モードと一致すると決定したこと、分類1346が特定の分類と一致すると決定したこと、または両方に基づいて、第1の代替発話信号1372を選択することができる。信号処理モジュール122は、発話信号1330のその部分を第1の代替発話信号1372で置き換えることによって、第1の処理された発話信号116を生成することができる。信号処理モジュール122は、スピーカー142を介して第1の処理された発話信号116を子供へ出力し得る。第1の処理された発話信号116は、その親の発話と同じように聞こえ得る。こうして、その子供は、親がいないとき(たとえば、街の外にいる)ときにその特定の話の後続の章を親の声で聞くことが可能になり得る。

【0202】

システム1300はこうして、ユーザから受け取られたオーディオ信号に基づいて、ユーザの発話に対応する代替発話信号の生成を可能にし得る。代替発話信号は、ユーザの発話と同じように聞こえる処理された発話信号を生成するために使用され得る。たとえば、処理された発話信号は、テキスト発話変換器によって生成された発話信号の一部分を代替発話信号で置き換えることによって生成され得る。

【0203】

図14を参照すると、発話信号を置き換える方法の特定の例示的な態様のフローチャートが示されており、全体が1400と指定される。方法1400は、図1の信号処理モジュール122、図4の信号処理モジュール422、図6の信号処理モジュール622、図7の信号処理モジュール722、またはこれらの組合せによって実行され得る。

【0204】

1402において、方法1400は、複数の代替発話信号を取得するステップを含む。たとえば、第1のデバイス102は、図2を参照してさらに説明されるように、代替発話信号112を取得することができる。

【0205】

1404において、方法1400はまた、デバイスにおいて、ユーザと関連付けられるユーザ発話信号を受信するステップを含み得る。ユーザ発話信号は、音声呼の間に受信され得る。たとえば、第1のデバイス102は、図1を参照してさらに説明されたように、モバイルデバイス104との音声呼の間に第1のユーザ発話信号130を受信することができる。

【0206】

1406において、方法1400はさらに、ユーザ発話信号の一部分を、ユーザと関連付けられる複数の代替発話信号と比較するステップを含み得る。複数の代替発話信号はデータベースに記憶され得る。たとえば、図1の信号処理モジュール122は、図1を参照してさらに説明されたように、第1のユーザ発話信号130の第1の部分162を第1のデータベース124に記憶されている代替発話信号112と比較することができる。

【0207】

1408において、方法1400はまた、その比較に基づいて、ユーザ発話信号のその部分が複数の代替発話信号の第1の代替発話信号と一致すると決定するステップを含み得る。たとえば、図1を参照して説明されたように、図1の信号処理モジュール122は、第1の部分162は第1の代替発話信号172と一致すると決定することができる。

【0208】

1410において、方法1400はさらに、その決定に応答して、ユーザ発話信号のその部分を第1の代替発話信号で置き換えることによって、処理された発話信号を生成するステップを含み得る。処理された発話信号は、音声呼の間に生成され得る。たとえば、信号処理モジュール122は、その決定に応答して、第1の部分162を第1の代替発話信号172で置き換えることによって、第1の処理された発話信号116を生成することができる。ある特定の態様では、図1を参照してさらに説明されたように、信号処理モジュール122は、第1の代替発話信号172と第1の処理された発話信号116の別の部分との間の遷移に対応する発話パラメータの変動を減らすために、平滑化アルゴリズムを使用することができる。

【0209】

1412において、方法1400はまた、スピーカーを介して処理された発話信号を出力するス

トップを含み得る。たとえば、図1の第1のデバイス102は、スピーカ-142を介して第1の処理された発話信号116を出力することができる。

【0210】

方法1400はさらに、1414において、ユーザ発話信号の複数の発話パラメータに基づいて第1の代替発話信号を修正するステップと、1416において、修正された第1の代替発話信号をデータベースに記憶するステップとを含み得る。たとえば、図1を参照してさらに説明されたように、信号処理モジュール122は、第1のユーザ発話信号130の複数の発話パラメータに基づいて第1の代替発話信号172を修正して、修正された代替発話信号176を生成することができる。信号処理モジュール122は、修正された代替発話信号176を第1のデータベース124に記憶することができる。

10

【0211】

したがって、方法1400は、あるユーザと関連付けられるユーザ発話信号の一部を同じユーザと関連付けられる代替発話信号で置き換えることによって、処理された発話信号の生成を可能にし得る。処理された発話信号は、ユーザ発話信号よりも高いオーディオ品質を有し得る。

【0212】

図14の方法1400は、フィールドプログラマブルゲートアレイ(FPGA)デバイス、特定用途向け集積回路(ASIC)、処理ユニット、たとえば中央処理装置(CPU)、デジタル信号プロセッサ(DSP)、コントローラ、別のハードウェアデバイス、ファームウェアデバイス、またはこれらの任意の組合せによって実施され得る。ある例として、図14の方法1400は、図19

20

【0213】

図15を参照すると、複数の代替発話信号を取得する方法の特定の例示的な態様のフローチャートが示されており、全体が1500と指定される。特定の態様では、方法1500は、図14の動作1402に対応し得る。方法1500は、図2の発話信号マネージャ262によって実行され得る。

【0214】

1502において、方法1500は、ユーザと関連付けられる訓練発話信号を受信するステップを含み得る。たとえば、図2を参照してさらに説明されたように、図2のサーバ206は、第2のユーザ154と関連付けられる訓練発話信号272を受信することができる。

30

【0215】

方法1500はまた、1504において、訓練発話信号から複数の代替発話信号を生成するステップと、1506において、複数の代替発話信号をデータベースに記憶するステップとを含み得る。たとえば、図2を参照して説明されるように、発話信号マネージャ262は、訓練発話信号272から代替発話信号112を生成することができる。発話信号マネージャ262は、第2のデータベース264に代替発話信号112を記憶することができる。

【0216】

したがって、方法1500は、ユーザ発話信号の一部を置き換えるために使用され得る代替発話信号の取得を可能にし得る。

【0217】

図15の方法1500は、フィールドプログラマブルゲートアレイ(FPGA)デバイス、特定用途向け集積回路(ASIC)、処理ユニット、たとえば中央処理装置(CPU)、デジタル信号プロセッサ(DSP)、コントローラ、別のハードウェアデバイス、ファームウェアデバイス、またはこれらの任意の組合せによって実施され得る。ある例として、図15の方法1500は、図19

40

【0218】

図16を参照すると、発話信号を置き換える方法の特定の例示的な態様のフローチャートが示されており、全体が1600と指定される。方法1600は、図1の信号処理モジュール122、図4の信号処理モジュール422、図6の信号処理モジュール622、図7の信号処理モジュール722、またはこれらの組合せによって実行され得る。

50

【0219】

1602において、方法1600は、デバイスにおいて、入力発話信号に基づいて検索基準を生成するステップを含み得る。たとえば、図4を参照して説明されたように、分析器402は、第1のユーザ発話信号130に基づいて検索基準418を生成することができる。別の例として、図6を参照して説明されたように、分析器602は、第1のデータ630に基づいて検索基準612を生成することができる。

【0220】

1604において、方法1600はまた、検索基準に基づいてデータベースを検索することによって、検索結果を生成するステップを含み得る。データベースは、複数の代替発話信号と関連付けられる発話パラメータのセットを記憶し得る。たとえば、図4を参照してさらに説明されたように、検索器404は、検索基準418に基づいてパラメータ化されたデータベース424を検索することによって、検索結果414を生成することができる。別の例として、図6を参照して説明されたように、検索器604は、検索基準612に基づいて複数の発話パラメータ614のセットを生成することができる。さらなる例として、検索器704は、検索基準612に基づいて複数の代替発話信号714を生成することができる。パラメータ化されたデータベース424は、発話パラメータ472、474、および476を含み得る。

10

【0221】

1606において、方法1600はさらに、制約に基づいて検索結果の特定の検索結果を選択するステップを含み得る。特定の検索結果は、複数の代替発話信号の第1の代替発話信号と関連付けられ得る。たとえば、図4を参照して説明されたように、制約分析器406は、制約に基づいて検索結果414の選択された結果416を選択することができる。選択された結果416は、第1の代替発話信号172、第1の発話パラメータ472、または両方を含み得る。別の例として、図6を参照して説明されたように、制約分析器606は、複数の発話パラメータ614のセットの第1の複数の発話パラメータ472を選択することができる。さらなる例として、図7を参照して説明されたように、制約分析器706は、複数の代替発話信号714の第1の代替発話信号172を選択することができる。

20

【0222】

1608において、方法1600はまた、ユーザ発話信号の一部を置換発話信号で置き換えることによって、処理された発話信号を生成するステップを含み得る。置換発話信号は、特定の検索結果に基づいて決定され得る。たとえば、図4を参照して説明されたように、合成器408は、第1のユーザ発話信号130の一部を置換発話信号で置き換えることによって、第1の処理された発話信号116を生成することができる。別の例として、図6を参照して説明されたように、合成器608は、第1のデータ630の入力信号の一部を置換発話信号で置き換えることによって、第1の処理された発話信号116を生成することができる。さらなる例として、図7を参照して説明されたように、合成器708は、第1のデータ630の入力信号の一部を置換発話信号で置き換えることによって、第1の処理された発話信号116を生成することができる。

30

【0223】

したがって、方法1600は、入力発話信号の一部を代替発話信号で置き換えることによって、処理された発話信号の生成を可能にし得る。処理された発話信号は、入力発話信号よりも高いオーディオ品質を有し得る。

40

【0224】

図16の方法1600は、フィールドプログラマブルゲートアレイ(FPGA)デバイス、特定用途向け集積回路(ASIC)、処理ユニット、たとえば中央処理装置(CPU)、デジタル信号プロセッサ(DSP)、コントローラ、別のハードウェアデバイス、ファームウェアデバイス、またはこれらの任意の組合せによって実施され得る。ある例として、図16の方法1600は、図19に関して説明されるように、命令を実行するプロセッサによって実行され得る。

【0225】

図17を参照すると、ユーザインターフェースを生成する方法の特定の例示的な態様のフローチャートが示されており、全体が1700と指定される。方法1700は、図1の信号処理モ

50

ジュール122、図4の信号処理モジュール422、図6の信号処理モジュール622、図7の信号処理モジュール722、またはこれらの組合せによって実行され得る。

【0226】

方法1700は、1702において、デバイスにおいてユーザインターフェースを生成するステップを含み得る。ユーザインターフェースは、ユーザの1つまたは複数の発話信号と関連付けられる時間長を示し得る。ユーザインターフェースは、追加オプションを含み得る。たとえば、図3を参照して説明されたように、図1の信号処理モジュール122は、ユーザインターフェース300を生成することができる。ユーザインターフェース300の第1のバー302は、図3を参照して説明されたように、ユーザ(たとえば、「Randal Hughes」)の1つまたは複数の発話信号と関連付けられる時間長を示し得る。ユーザインターフェース300は、図3を参照して説明されたように、追加オプション308を含み得る。

10

【0227】

方法1700はまた、1704において、ユーザインターフェースをディスプレイに提供するステップを含み得る。たとえば、図3を参照して説明されたように、図1の信号処理モジュール122は、ユーザインターフェース300を第1のデバイス102のディスプレイに提供することができる。

【0228】

1706において、方法1700はさらに、追加オプションの選択を受け取ったことに応答して、1つまたは複数の発話信号に対応する複数の代替発話信号を生成するステップを含み得る。たとえば、信号処理モジュール122は、図3の追加オプション308の選択を受け取ったことに応答して、1つまたは複数の発話信号に対応する複数の代替発話信号112を生成することができる。例示すると、図3を参照して説明されたように、信号処理モジュール122は、追加オプション308の選択を受け取ったことに応答して要求を発話信号マネージャ262に送ることができ、発話信号マネージャ262から複数の代替発話信号112を受信することができる。

20

【0229】

したがって、方法1700は、代替発話信号の管理を可能にし得る。方法1700は、ユーザと関連付けられる発話信号の監視を可能にし得る。方法1700は、選択されたユーザと関連付けられる代替発話信号の生成を可能にし得る。

【0230】

図17の方法1700は、フィールドプログラマブルゲートアレイ(FPGA)デバイス、特定用途向け集積回路(ASIC)、処理ユニット、たとえば中央処理装置(CPU)、デジタル信号プロセッサ(DSP)、コントローラ、別のハードウェアデバイス、ファームウェアデバイス、またはこれらの任意の組合せによって実施され得る。ある例として、図17の方法1700は、図19に関して説明されるように、命令を実行するプロセッサによって実行され得る。

30

【0231】

図18を参照すると、発話信号を置き換える方法の特定の例示的な態様のフローチャートが示されており、全体が1800と指定される。方法1800は、図1の信号処理モジュール122、図4の信号処理モジュール422、図6の信号処理モジュール622、図7の信号処理モジュール722、またはこれらの組合せによって実行され得る。

40

【0232】

1802において、方法1800は、局所音声プロファイルを記憶するデバイスにおいて遠隔音声プロファイルを受信するステップを含んでよく、その局所音声プロファイルはある人物と関連付けられる。たとえば、図11を参照して説明されたように、第1のデバイス102の信号処理モジュール122は、第2のユーザ154と関連付けられる遠隔音声プロファイル1174を受信することができる。第1のデバイス102は、第2のユーザ154と関連付けられる局所音声プロファイル1104を記憶することができる。

【0233】

1804において、方法1800はまた、遠隔音声プロファイルと局所音声プロファイルとの比較に基づいて、または、遠隔音声プロファイルと関連付けられる識別子に基づいて、遠隔

50

音声プロフィールがその人物と関連付けられると決定するステップを含み得る。たとえば、図11を参照して説明されたように、信号処理モジュール122は、遠隔音声プロフィール1174と関連付けられる発話の内容と、局所音声プロフィール1104と関連付けられる発話の内容との比較に基づいて、遠隔音声プロフィール1174が第2のユーザ154と関連付けられると決定することができる。別の例として、図11を参照して説明されたように、信号処理モジュール122は、遠隔音声プロフィール1174が第2のユーザ154と関連付けられる識別子1168を含むと決定したことに応答して、遠隔音声プロフィール1174が第2のユーザ154と関連付けられると決定することができる。

【0234】

1806において、方法1800はさらに、この決定に基づいて、デバイスにおいて、プロフィール管理のために局所音声プロフィールを選択するステップを含み得る。たとえば、図11を参照して説明されたように、第1のデバイス102の信号処理モジュール122は、遠隔音声プロフィール1174が第2のユーザ154に対応すると決定したことに基づいて、局所音声プロフィール1104を選択することができる。プロフィール管理の例は、限定はされないが、遠隔音声プロフィール1174に基づいて局所音声プロフィール1104を更新することと、遠隔音声プロフィール1174の表現、局所音声プロフィール1104の表現、または両方を含むGUIを生成することによって音声プロフィール情報をユーザに提供することを含み得る。

【0235】

方法1800はこうして、別のデバイスからある人物の遠隔音声プロフィールを受信したことに応答して、その人物の局所音声プロフィールがプロフィール管理のために選択されることを可能にし得る。たとえば、局所音声プロフィールは、遠隔音声プロフィールと比較され得る。別の例として、局所音声プロフィールは、遠隔音声プロフィールに基づいて更新され得る。

【0236】

図18の方法1800は、フィールドプログラマブルゲートアレイ(FPGA)デバイス、特定用途向け集積回路(ASIC)、処理ユニット、たとえば中央処理装置(CPU)、デジタル信号プロセッサ(DSP)、コントローラ、別のハードウェアデバイス、ファームウェアデバイス、またはこれらの任意の組合せによって実施され得る。ある例として、図18の方法1800は、図18に関して説明されるように、命令を実行するプロセッサによって実行され得る。

【0237】

図19を参照すると、デバイス(たとえば、ワイヤレス通信デバイス)の特定の例示的な態様のブロック図が示され、全体が1900と指定される。様々な態様において、デバイス1900は、図19に示されるものよりも多数または少数のコンポーネントを有し得る。ある例示的な態様では、デバイス1900は、図1の第1のデバイス102、モバイルデバイス104、デスクトップコンピュータ106、または図2のサーバ206に対応し得る。ある例示的な態様では、デバイス1900は、図1～図18を参照して本明細書で説明された1つまたは複数の動作を実行し得る。

【0238】

ある特定の態様では、デバイス1900は、プロセッサ1906(たとえば、中央処理装置(CPU))を含む。プロセッサ1906は、図11のプロセッサ1152、プロセッサ1160、または両方に対応し得る。デバイス1900は、1つまたは複数の追加のプロセッサ1910(たとえば、1つまたは複数のデジタル信号プロセッサ(DSP))を含み得る。プロセッサ1910は、発話および音楽のコーダ-デコーダ(コーデック)1908と、エコーキャンセラ1912とを含み得る。発話および音楽のコーデック1908は、ボコーダのエンコーダ1936、ボコーダのデコーダ1938、または両方を含み得る。

【0239】

デバイス1900は、メモリ1932とコーデック1934とを含み得る。メモリ1932は、図11のメモリ1132、メモリ1176、または両方に対応し得る。デバイス1900は、高周波(RF)モジュール1950を介してアンテナ1942に結合されるワイヤレスコントローラ1940を含み得る。デバイス1900は、ディスプレイコントローラ1926に結合されたディスプレイ1128を含み得る。

10
20
30
40
50

図1のスピーカ-142、1つまたは複数のマイクロフォン1946、またはこれらの組合せが、コーデック1934に結合され得る。コーデック1934は、デジタルアナログコンバータ1902とアナログデジタルコンバータ1904とを含み得る。ある例示的な態様では、マイクロフォン1946は、図1のマイクロフォン144、マイクロフォン146、図2のマイクロフォン246、図12のマイクロフォン1244、またはこれらの組合せに対応し得る。ある特定の態様では、コーデック1934は、マイクロフォン1946からアナログ信号を受け取り、アナログデジタルコンバータ1904を使用してそのアナログ信号をデジタル信号に変換し、そのデジタル信号を発話および音楽のコーデック1908に与えることができる。発話および音楽のコーデック1908は、デジタル信号を処理することができる。ある特定の態様では、発話および音楽のコーデック1908は、デジタル信号をコーデック1934に与えることができる。コーデック1934は、デジタルアナログコンバータ1902を使用してデジタル信号をアナログ信号に変換することができ、アナログ信号をスピーカ-142に与えることができる。

【0240】

メモリ1932は、図1の第1のデータベース124、図2の第2のデータベース264、図4のパラメータ化されたデータベース424、図11の局所音声プロファイル1108、遠隔音声プロファイル1178、入力オーディオ信号1130、図13の分類1346、使用モードの設定1362、またはこれらの組合せを含み得る。デバイス1900は、信号処理モジュール1948、図2の発話信号マネージャ262、図4のデータベース分析器410、図8のアプリケーション822、図9の検索モジュール922、図13のモード検出器1302、領域検出器1304、テキスト発話変換器1322、またはこれらの組合せを含み得る。信号処理モジュール1948は、図1の信号処理モジュール122、図4の信号処理モジュール422、図6の信号処理モジュール622、図7の信号処理モジュール722、またはこれらの組合せに対応し得る。ある特定の態様では、発話信号マネージャ262、データベース分析器410、図8のアプリケーション822、図9の検索モジュール922、図13のモード検出器1302、領域検出器1304、テキスト発話変換器1322、および/または信号処理モジュール1948の1つまたは複数のコンポーネントは、プロセッサ1906、プロセッサ1910、コーデック1934、またはこれらの組合せに含まれ得る。ある特定の態様では、発話信号マネージャ262、データベース分析器410、図8のアプリケーション822、図9の検索モジュール922、図13のモード検出器1302、領域検出器1304、テキスト発話変換器1322、および/または信号処理モジュール1948の1つまたは複数のコンポーネントは、ボコーダのエンコーダ1936、ボコーダのデコーダ1938、または両方に含まれ得る。

【0241】

信号処理モジュール1948、発話信号マネージャ262、データベース分析器410、アプリケーション822、検索モジュール922、モード検出器1302、領域選択器1304、テキスト発話変換器1322、またはこれらの組合せが、本明細書で説明される発話信号置換技法のハードウェアの態様を実装するために使用され得る。代替的に、または追加で、ソフトウェアの態様(またはソフトウェア/ハードウェアの組み合わせられた態様)が実装され得る。たとえば、メモリ1932は、プロセッサ1910またはデバイス1900の他の処理ユニット(たとえば、プロセッサ1906、コーデック1934、または両方)によって実行可能な命令1956を含み得る。命令1956は、発話信号マネージャ262、データベース分析器410、アプリケーション822、検索モジュール922、モード検出器1302、領域選択器1304、テキスト発話変換器1322、信号処理モジュール1948、またはこれらの組合せに対応し得る。

【0242】

特定の態様では、デバイス1900は、システムインパッケージまたはシステムオンチップデバイス1922に含まれ得る。ある特定の態様では、プロセッサ1906、プロセッサ1910、ディスプレイコントローラ1926、メモリ1932、コーデック1934、およびワイヤレスコントローラ1940は、システムインパッケージまたはシステムオンチップデバイス1922に含まれる。ある特定の態様では、入力デバイス1134および電源1944がシステムオンチップデバイス1922に結合される。その上、ある特定の態様では、図19に示されるように、ディスプレイ1128、入力デバイス1134、スピーカ-142、マイクロフォン1946、アンテナ1942、および電源1944は、システムオンチップデバイス1922の外部にある。ある特定の態様では、ディ

10

20

30

40

50

ディスプレイ1128、入力デバイス1134、スピーカー142、マイクロフォン1946、アンテナ1942、および電源1944の各々は、インターフェースまたはコントローラのようなシステムオンチップデバイス1922のコンポーネントに結合され得る。

【0243】

デバイス1900は、モバイル通信デバイス、スマートフォン、携帯電話、ラップトップコンピュータ、コンピュータ、タブレット、携帯情報端末、ディスプレイデバイス、テレビジョン、ゲームコンソール、音楽プレーヤ、ラジオ、デジタルビデオプレーヤ、デジタルビデオディスク(DVD)プレーヤ、チューナー、カメラ、ナビゲーションデバイス、またはこれらの任意の組合せを含み得る。

【0244】

ある例示的な態様では、プロセッサ1910は、図1~図18を参照して説明された方法または動作のすべてまたは一部分を実行するように動作可能であり得る。たとえば、マイクロフォン1946は、ユーザ発話信号に対応するオーディオ信号を捉えることができる。ADC 1904は、捉えられたオーディオ信号を、アナログ波形から、デジタルオーディオサンプルからなるデジタル波形へと変換することができる。プロセッサ1910は、デジタルオーディオサンプルを処理することができる。利得調整器は、デジタルオーディオサンプルを調整することができる。エコーキャンセラ1912は、マイクロフォン1946に入るスピーカー142の出力によって生み出された可能性のある、あらゆるエコーを低減することができる。

【0245】

プロセッサ1910は、ユーザ発話信号の一部分を複数の代替発話信号と比較することができる。たとえば、プロセッサ1910は、デジタルオーディオサンプルの一部分を複数の発話信号と比較することができる。プロセッサ1910は、複数の代替発話信号の第1の代替発話信号がユーザ発話信号のその部分と一致すると決定することができる。プロセッサ1910は、ユーザ発話信号のその部分を第1の代替発話信号で置き換えることによって、処理された発話信号を生成することができる。

【0246】

ボコーダのエンコーダ1936は、処理された発話信号に対応するデジタルオーディオサンプルを圧縮することができ、送信パケット(たとえば、デジタルオーディオサンプルの圧縮されたビットの表現)を形成することができる。送信パケットは、メモリ1932に記憶され得る。RFモジュール1950は、何らかの形態の送信パケットを変調することができ(たとえば、他の情報が送信パケットに付加され得る)、アンテナ1942を介して変調されたデータを送信することができる。別の例として、プロセッサ1910は、マイクロフォン1946を介して訓練信号を受け取ることができ、訓練信号から複数の代替発話信号を生成することができる。

【0247】

さらなる例として、アンテナ1942は、受信パケットを含む入来パケットを受信することができる。受信パケットは、ネットワークを介して別のデバイスによって送られ得る。たとえば、受信パケットは、ユーザ発話信号に対応し得る。ボコーダのデコーダ1938は、受信パケットを解凍することができる。解凍された波形は、再構築されたオーディオサンプルと呼ばれ得る。エコーキャンセラ1912は、再構築されたオーディオサンプルからエコーを除去することができる。

【0248】

プロセッサ1910は、ユーザ発話信号の一部分を複数の代替発話信号と比較することができる。たとえば、プロセッサ1910は、再構築されたオーディオサンプルの一部分を複数の発話信号と比較することができる。プロセッサ1910は、複数の代替発話信号の第1の代替発話信号がユーザ発話信号のその部分と一致すると決定することができる。プロセッサ1910は、ユーザ発話信号のその部分を第1の代替発話信号で置き換えることによって、処理された発話信号を生成することができる。利得調整器は、処理された発話信号を増幅または抑制することができる。DAC 1902は、処理された発話信号をデジタル波形からアナログ波形に変換することができ、変換された信号をスピーカー142に与えることができる。

10

20

30

40

50

【0249】

説明された態様とともに、装置は、遠隔音声プロファイルを受信するための手段を含み得る。たとえば、遠隔音声プロファイルを受信するための手段は、信号処理モジュール1948、プロセッサ1906、プロセッサ1910、RFモジュール1950、遠隔音声プロファイルを受信するように構成される1つまたは複数の他のデバイスもしくは回路、またはこれらの任意の組合せを含み得る。

【0250】

装置はまた、ある人物と関連付けられる局所音声プロファイルを記憶するための手段を含み得る。たとえば、記憶するための手段は、メモリ1932、信号処理モジュール1948、プロセッサ1906、プロセッサ1910、局所音声プロファイルを記憶するように構成される1つまたは複数の他のデバイスもしくは回路、またはこれらの任意の組合せを含み得る。

10

【0251】

装置はさらに、遠隔音声プロファイルと関連付けられる発話の内容または遠隔音声プロファイルの識別子に基づいて、遠隔音声プロファイルがその人物と関連付けられると決定したに基づいて、プロファイル管理のために局所音声プロファイルを選択するための手段を含み得る。たとえば、局所音声プロファイルを選択するための手段は、信号処理モジュール1948、プロセッサ1906、プロセッサ1910、局所音声プロファイルを選択するように構成される1つまたは複数の他のデバイスもしくは回路、またはこれらの任意の組合せを含み得る。

【0252】

受信するための手段、記憶するための手段、および選択するための手段は、車両、電子書籍デバイス、音響キャプチャデバイス、モバイル通信デバイス、スマートフォン、携帯電話、ラップトップコンピュータ、コンピュータ、エンコーダ、デコーダ、タブレット、携帯情報端末、ディスプレイデバイス、テレビジョン、ゲームコンソール、音楽プレーヤ、ラジオ、デジタルビデオプレーヤ、デジタルビデオディスクプレーヤ、チューナー、カメラ、またはナビゲーションデバイスの少なくとも1つへと組み込まれ得る。

20

【0253】

当業者はさらに、本明細書で開示される態様に関連して説明された様々な例示的な論理ブロック、構成、モジュール、回路、およびアルゴリズムステップが、電子ハードウェア、プロセッサによって実行されるコンピュータソフトウェア、または両方の組合せとして実装され得ることを理解するだろう。上では、様々な例示的なコンポーネント、ブロック、構成、モジュール、回路、およびステップが、全般にそれらの機能の観点から説明されている。そのような機能が、ハードウェアとして実装されるか、またはプロセッサ実行可能命令として実装されるかは、具体的な適用例と、システム全体に課される設計制約とによって決まる。当業者は、説明された機能を具体的な適用例ごとに様々な方法で実装し得るが、そのような実装の判断は、本開示の範囲からの逸脱をもたらすものと解釈されるべきではない。

30

【0254】

前述の例の1つまたは複数は、特定の人物と関連付けられる遠隔音声プロファイルおよび局所音声プロファイルを説明するが、そのような例は例示を目的とするものに過ぎず、限定するものとして見なされるべきではないことに留意されたい。特定の態様では、デバイスは「汎用的な」(またはデフォルトの)音声プロファイルを記憶し得る。たとえば、汎用音声プロファイルは、個々の人物と関連付けられる音声プロファイルの交換を通信インフラストラクチャまたは通信環境がサポートしないときに使用され得る。したがって、汎用音声プロファイルは、限定はされないが、特定の発呼者または被呼者と関連付けられる音声プロファイルが利用不可能であるとき、置換(たとえば、より高品質の)発話データが特定の使用モードまたは属性領域に対して利用不可能であるときなどを含む、いくつかの状況にあるデバイスにおいて選択され得る。いくつかの実装形態では、汎用音声プロファイルは、特定の人物と関連付けられる音声プロファイルを補足するために使用され得る。例示すると、特定の人物と関連付けられる音声プロファイルが特定の音、使用モード、属

40

50

性領域、またはこれらの組合せに対する置換発話データを含まないとき、置換発話データは代わりに、汎用音声プロファイルに基づいて導出され得る。特定の態様では、デバイスは、複数の汎用またはデフォルト音声プロファイル(たとえば、会話vs歌唱、英語vsフランス語などに対する別々の汎用またはデフォルト音声プロファイル)を記憶し得る。

【0255】

本明細書で開示される態様に関して説明される方法またはアルゴリズムのステップは、ハードウェアにおいて直接具現化されるか、プロセッサによって実行されるソフトウェアモジュールにおいて具現化されるか、またはその2つの組合せで具現化されることがある。ソフトウェアモジュールは、ランダムアクセスメモリ(RAM)、フラッシュメモリ、読取り専用メモリ(ROM)、プログラム可能読取り専用メモリ(PROM)、消去可能プログラム可能読取り専用メモリ(EPROM)、電氣的消去可能プログラム可能読取り専用メモリ(EEPROM)、レジスタ、ハードディスク、リムーバブルディスク、コンパクトディスク読取り専用メモリ(CD-ROM)、または当技術分野において既知の任意の他の形の非一時的記憶媒体内に存在し得る。例示的な記憶媒体は、プロセッサが記憶媒体から情報を読み取り、記憶媒体に情報を書き込むことができるように、プロセッサに結合される。代替的に、記憶媒体はプロセッサと一体化されてもよい。プロセッサおよび記憶媒体は、特定用途向け集積回路(ASIC)内に存在し得る。ASICは、コンピューティングデバイスまたはユーザ端末内に存在し得る。代替として、プロセッサおよび記憶媒体は、コンピューティングデバイスまたはユーザ端末内に個別のコンポーネントとして存在し得る。

10

【0256】

開示された態様の前述の説明は、開示された態様を当業者が作成または使用することを可能にするために提供される。これらの態様への様々な修正が当業者には容易に明らかになり、本明細書で定義された原理は本開示の範囲から逸脱することなく他の態様に適用され得る。したがって、本開示は、本明細書で示される態様に限定されることを意図するものではなく、以下の特許請求の範囲によって規定される原理および新規な特徴と一致する、取り得る最も広い範囲を与えられるべきである。

20

【符号の説明】

【0257】

- 100 システム
- 102 第1のデバイス
- 104 モバイルデバイス
- 106 デスクトップコンピュータ
- 112 代替発話信号
- 116 第1の処理された発話信号
- 118 第2の処理された発話信号
- 120 ネットワーク
- 122 信号処理モジュール
- 124 第1のデータベース
- 130 第1のユーザ発話信号
- 132 第2のユーザ発話信号
- 142 スピーカー
- 144 マイクロフォン
- 146 マイクロフォン
- 152 第1のユーザ
- 154 第2のユーザ
- 162 第1の部分
- 164 第2の部分
- 172 第1の代替発話信号
- 174 第2の代替発話信号
- 176 修正された代替発話信号

30

40

50

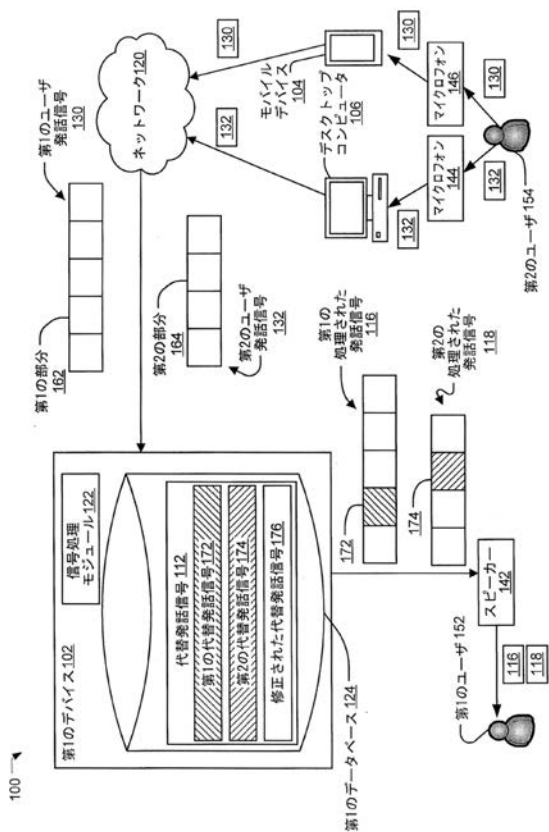
200	システム	
206	サーバ	
220	第3の代替発話信号	
222	第5の代替発話信号	
224	第4の代替発話信号	
244	マイクロフォン	
262	発話信号マネージャ	
264	第2のデータベース	
270	第1の呼プロファイル	
272	訓練発話信号	10
274	第2の呼プロファイル	
276	第1の連絡先リスト	
278	第2の連絡先リスト	
300	ユーザインターフェース	
302	第1のバー	
306	カーソル	
308	追加オプション	
310	削除オプション	
312	自動更新の閾値	
314	第2のバー	20
316	第3のバー	
330	出力信号	
372	第1の複数の発話パラメータ	
400	システム	
402	分析器	
404	検索器	
406	制約分析器	
408	合成器	
410	データベース分析器	
414	検索結果	30
416	選択された結果	
418	検索基準	
422	信号処理モジュール	
424	パラメータ化されたデータベース	
426	エンコーダ	
430	出力信号	
472	第1の発話パラメータ	
474	第2の発話パラメータ	
476	N個の発話パラメータ	
478	N個の代替発話信号	40
500	システム	
526	エンコーダ	
600	システム	
602	分析器	
604	検索器	
606	制約分析器	
608	合成器	
612	検索基準	
614	発話パラメータのセット	
618	第2のデータ	50

622	信号処理モジュール	
626	デコーダ	
630	第1のデータ(たとえば、入力信号)	
700	システム	
704	検索器	
706	制約分析器	
708	合成器	
714	複数の代替発話信号	
722	信号処理モジュール	
726	デコーダ	10
800	システム	
822	アプリケーション	
830	入力信号	
900	システム	
902	置換器	
904	検索器	
922	検索モジュール	
924	データベース	
930	入力信号(たとえば、発話信号)	
1002	MFCC	20
1004	LSP	
1006	LSF	
1010	ケプストラル	
1012	LSI	
1014	DCT	
1016	DFT	
1018	FFT	
1020	フォルマント周波数	
1022	PCMサンプル	
1024	データベース	30
1100	システム	
1102	更新	
1104	局所音声プロファイル	
1106	サーバ	
1108	局所音声プロファイル	
1110	更新要求	
1112	局所代替発話信号	
1120	プロファイル要求	
1128	ディスプレイ	
1130	入力オーディオ信号	40
1132	メモリ	
1134	入力デバイス	
1136	選択	
1138	GUI	
1140	ユーザ入力	
1150	送受信機	
1152	プロセッサ	
1160	プロセッサ	
1162	第1の部分	
1164	第1のセグメント	50

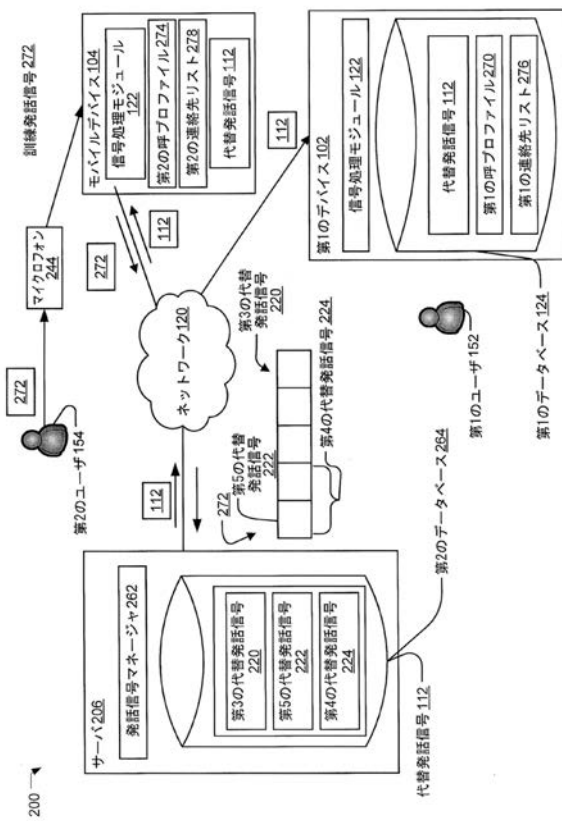
1166	第2のセグメント	
1168	識別子	
1170	発話の内容	
1172	代替発話信号	
1174	遠隔音声プロファイル	
1176	メモリ	
1178	遠隔音声プロファイル	
1180	送受信機	
1200	システム	
1202	更新	10
1204	局所音声プロファイル	
1210	更新要求	
1212	代替発話信号	
1214	局所代替発話信号	
1232	GUI	
1238	入力オーディオ信号	
1240	ユーザ入力	
1244	マイクロフォン	
1274	遠隔音声プロファイル	
1300	システム	20
1302	モード検出器	
1304	領域選択器	
1312	使用モード	
1314	属性領域	
1322	テキスト発話変換器	
1330	発話信号	
1338	訓練発話信号	
1346	分類	
1362	使用モードの設定	
1372	第1の代替発話信号	30
1374	第2の代替発話信号	
1400	方法	
1500	方法	
1600	方法	
1700	方法	
1800	方法	
1900	デバイス	
1902	DAC	
1904	ADC	
1906	プロセッサ	40
1908	発話および音楽コーデック	
1910	プロセッサ(たとえば、DSP)	
1912	エコーキャンセラ	
1922	システムオンチップデバイス	
1926	ディスプレイコントローラ	
1932	メモリ	
1934	コーデック	
1936	ボコーダのエンコーダ	
1938	ボコーダのデコーダ	
1940	ワイヤレスコントローラ	50

- 1942 アンテナ
- 1944 電源
- 1946 マイクロフォン
- 1948 信号処理モジュール
- 1950 RFモジュール
- 1956 命令

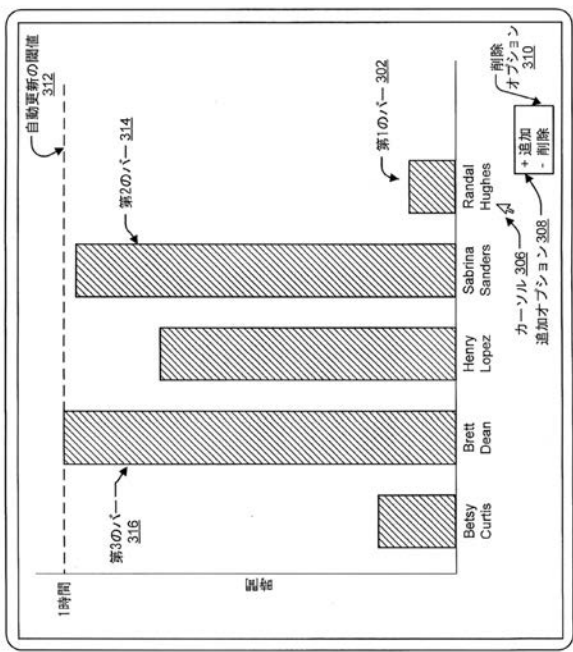
【図 1】



【図 2】

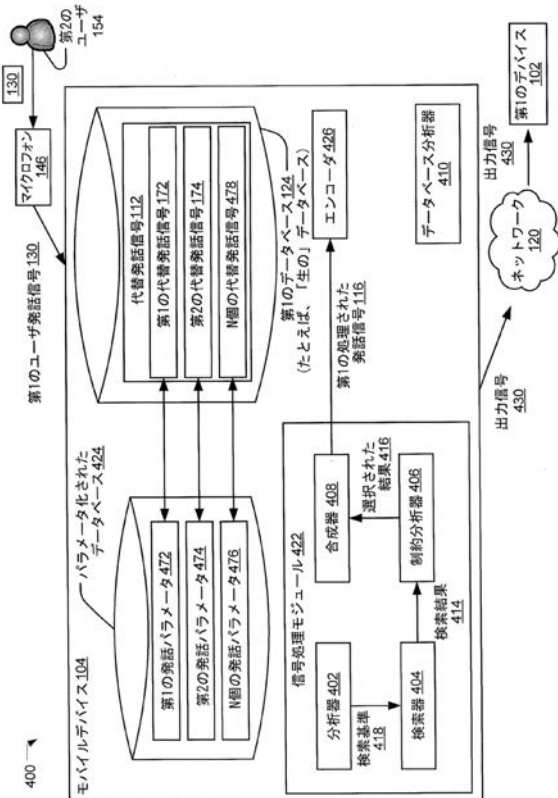


【図 3】



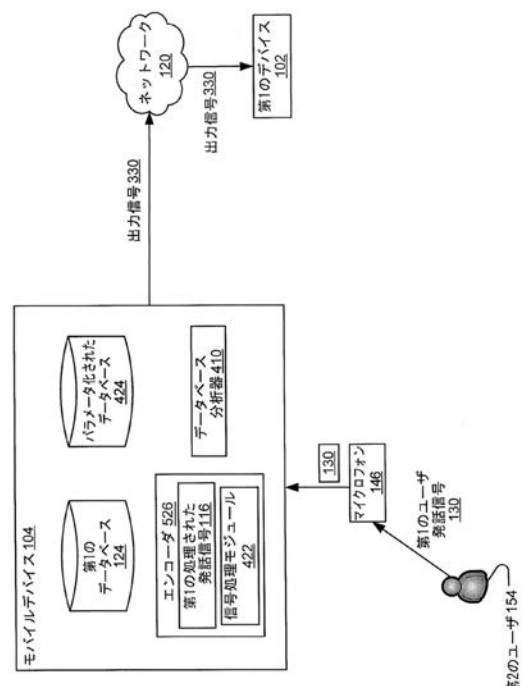
300

【図 4】



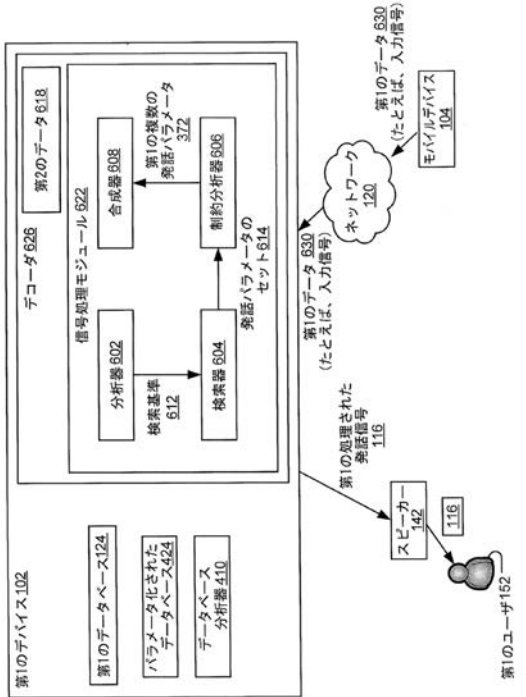
400

【図 5】



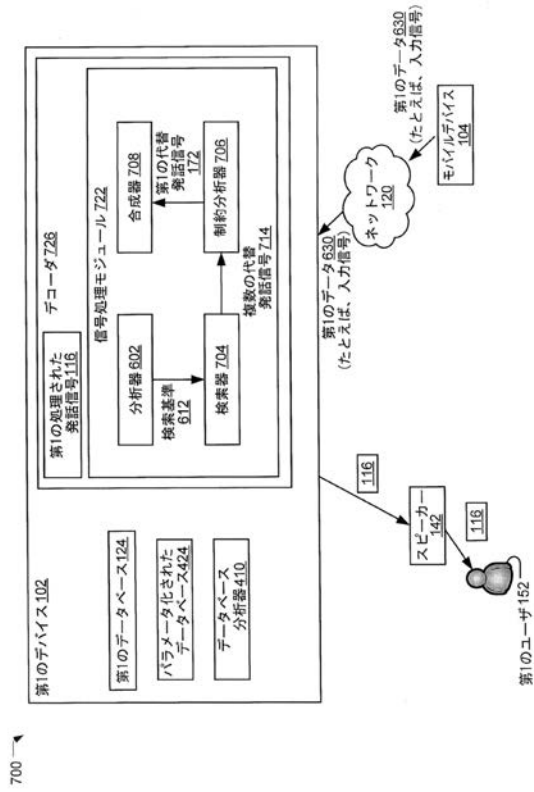
500

【図 6】

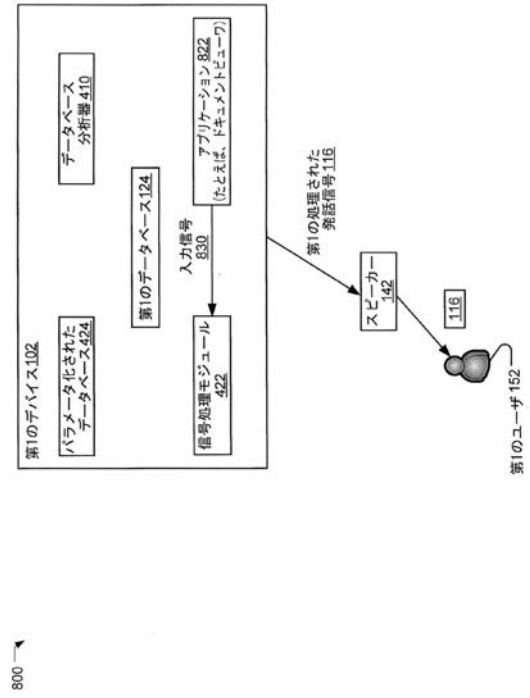


600

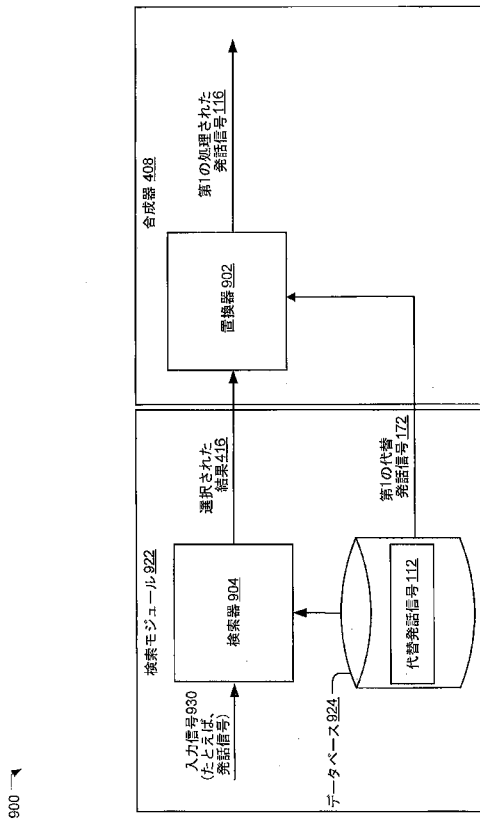
【図7】



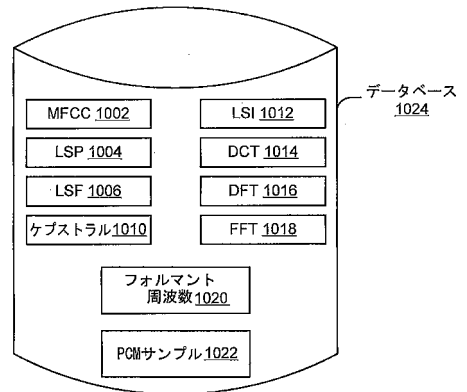
【図8】



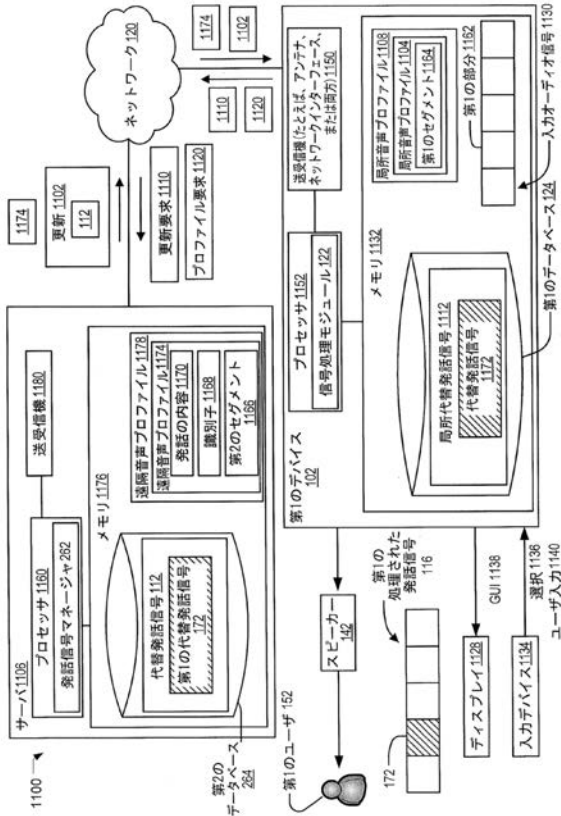
【図9】



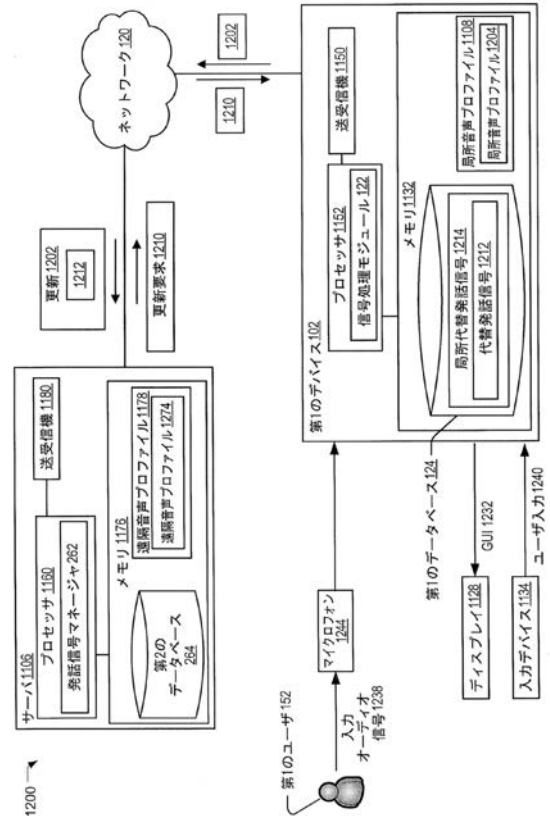
【図10】



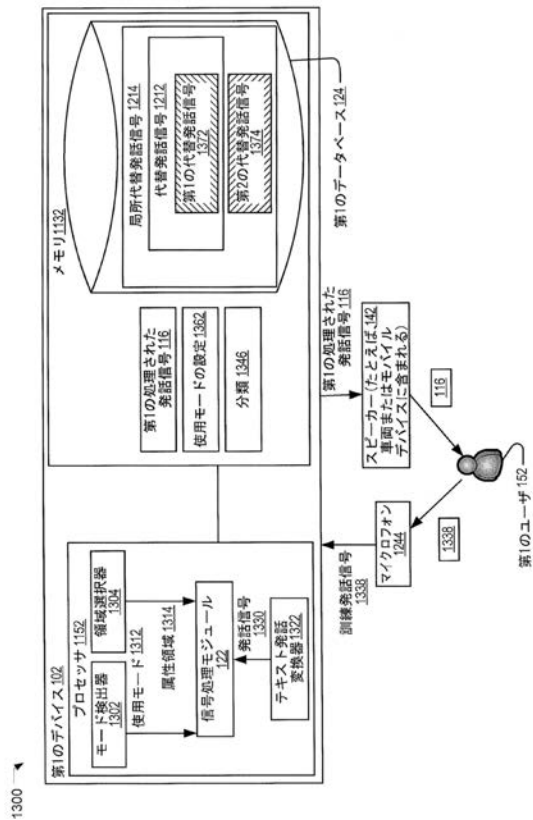
【図 1 1】



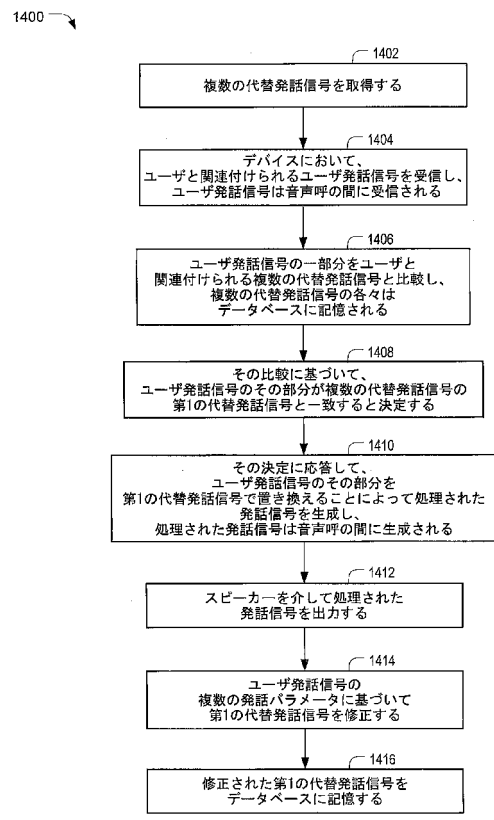
【図 1 2】



【図 1 3】

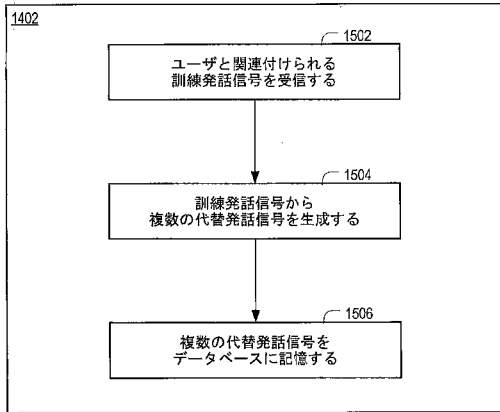


【図 1 4】



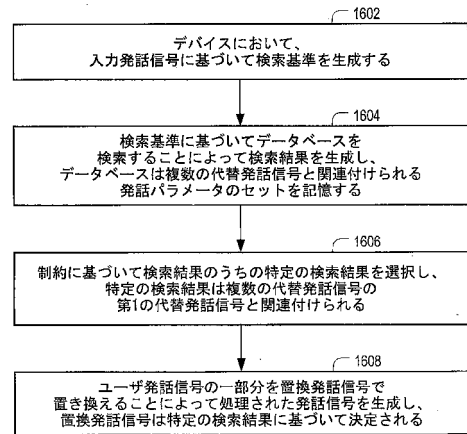
【 図 1 5 】

1500



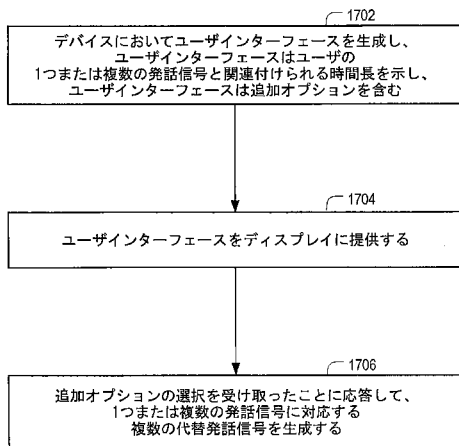
【 図 1 6 】

1600



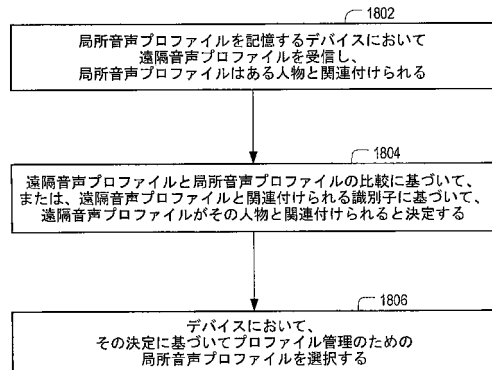
【 図 1 7 】

1700

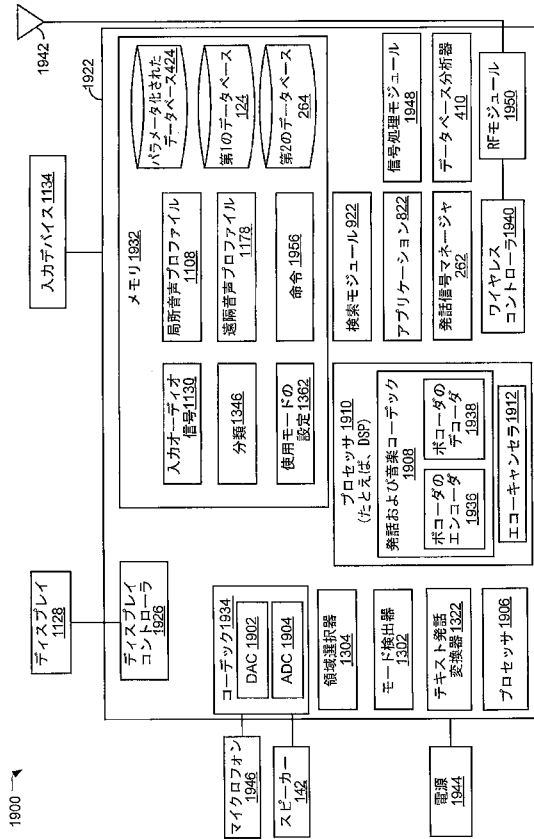


【 図 1 8 】

1800



【図 19】



【手続補正書】

【提出日】平成30年8月14日(2018.8.14)

【手続補正 1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

代替発話信号に対応する代替発話データを記憶するように構成されるメモリと、
プロセッサとを備え、前記プロセッサが、

入力発話信号に対応する入力発話データに基づいて、前記メモリの検索を介して検索
結果データを生成することと、

制約に基づいて特定の検索結果データを選択することであって、前記特定の検索結果
データが特定の代替発話データと関連付けられる、選択することと、

前記入力発話データの一部を置換発話データで置き換えることによって、処理され
た発話データを生成することであって、前記置換発話データが前記特定の検索結果データ
に基づいて決定される、生成することと

を行うように構成される、デバイス。

【請求項 2】

前記入力発話信号が歪んだ信号を備え、前記プロセッサがさらに、前記入力発話データ
が歪みを含むと決定するように構成される、請求項1に記載のデバイス。

【請求項 3】

前記プロセッサがさらに、前記入力発話データに基づいて検索基準データを生成するよ
うに構成され、前記検索基準データが、前記入力発話データと関連付けられる発話パラメ

ータデータを備える、請求項2に記載のデバイス。

【請求項4】

前記検索結果データが、前記入力発話データ、前記検索基準データ、または両方を使用して生成され、前記検索結果データが代替発話データのセットを備え、代替発話データの前記セットが実質的に前記入力発話データと一致する、請求項3に記載のデバイス。

【請求項5】

前記代替発話データがパルス符号変調(PCM)波形を備える、請求項1に記載のデバイス。

【請求項6】

前記メモリがさらに、前記代替発話データと関連付けられる発話パラメータデータを記憶するように構成され、前記発話パラメータデータが、ピッチパラメータ、エネルギーパラメータ、線形予測コーディング(LPC)パラメータ、メル周波数ケプストラル係数(MFCC)、線スペクトル対(LSP)、線スペクトル周波数(LSF)、ケプストラル、線スペクトル情報(LSI)、離散コサイン変換(DCT)パラメータ、パルス符号変調(PCM)、離散フーリエ変換(DFT)パラメータ、高速フーリエ変換(FFT)パラメータ、フォルマント周波数、またはこれらの組合せを含む、請求項1に記載のデバイス。

【請求項7】

前記プロセッサがさらに、前記入力発話データと発話パラメータデータとの比較に基づいて、前記検索結果データを生成するように構成される、請求項1に記載のデバイス。

【請求項8】

前記プロセッサがさらに、

前記入力発話データを複製して複製された発話データを生成し、

前記複製された発話データの特定の位置から前記入力発話データの前記一部分を除去し

、
前記置換発話データおよび前記複製された発話データを前記特定の位置において連結して前記置換発話データを生成する

ように構成される、請求項1に記載のデバイス。

【請求項9】

前記制約が、コストの制約、誤差の制約、または両方を備える、請求項1に記載のデバイス。

【請求項10】

複数の代替発話信号に対応する複数の代替発話データを記憶するように構成されるメモリと、

前記メモリに電氣的に結合されたプロセッサとを備え、前記プロセッサが、

発話データの少なくとも一部分を特定することであって、前記発話データが発話信号に対応する、特定することと、

前記代替発話データに基づいて、前記発話データの前記一部分と一致する前記代替発話データの一部分を選択することと、

前記発話データの少なくとも前記一部分を前記代替発話データの前記選択された一部分で置換して出力発話データを生成することと

を行うように構成される、デバイス。

【請求項11】

前記プロセッサがさらに、前記出力発話データを符号化して符号化された出力発話データを生成するように構成され、前記デバイスが、前記符号化された出力発話データを送信するように構成される送信機をさらに備える、請求項10に記載のデバイス。

【請求項12】

前記符号化された出力発話データがパッケージ化された発話データを備え、前記デバイスが、前記出力発話データを投影するように構成される出力デバイスをさらに備え、前記出力デバイスがスピーカデバイスを備える、請求項11に記載のデバイス。

【請求項13】

符号化された発話データを受信するように構成される受信機をさらに備え、前記プロセ

ッサがさらに、前記符号化された発話データを復号して再構築された発話データを生成するように構成され、前記再構築された発話データが前記発話データに対応する、請求項10に記載のデバイス。

【請求項14】

前記発話データが連続的な発話データフレームを備え、第1の発話データフレームの順番が、前記発話データの前記一部分に対応する第2の発話データフレームの前であり、前記プロセッサがさらに、

前記代替発話データの前記一部分と前記第2の発話データフレームとの間の第1の一致の第1の品質を示す第1の一致値を生成し、

前記第1の一致値が閾値より小さいという決定にตอบสนองして、前記代替発話データの前記一部分、前記第1の発話データフレーム、および前記第2の発話データフレームの間の第2の一致の第2の品質を示す第2の一致値を生成する

ように構成される、請求項10に記載のデバイス。

【請求項15】

前記プロセッサがさらに、前記第2の一致値を第2の閾値と比較するように構成され、前記出力発話データの生成が、前記代替発話データの前記一部分の、前記第1の発話データフレームへの連結に対応する、請求項14に記載のデバイス。

【請求項16】

代替発話データと関連付けられるロケータ値を表現するインジケータを受信するように構成される受信機と、

複数の代替発話信号に対応する前記代替発話データを記憶するように構成されるメモリと、

前記メモリに電氣的に結合されたプロセッサとを備え、前記プロセッサが、

前記ロケータ値に基づいて前記代替発話データの一部分を選択することであって、前記一部分が特定の代替発話信号と関連付けられる、選択することと、

前記代替発話データの前記一部分に基づいて出力発話データを生成することとを行うように構成される、デバイス。

【請求項17】

前記代替発話データが、実質的に歪みのない複数の代替発話データ部分を含む、請求項16に記載のデバイス。

【請求項18】

前記受信機がさらに、発話信号に対応する発話データの表現を受信するように構成され、前記出力発話データが、再構築された発話データの一部分との、前記代替発話データの前記一部分の成分の重複を備え、前記再構築された発話データが前記発話データの前記表現に基づく、請求項16に記載のデバイス。

【請求項19】

前記出力発話データが、前記代替発話データの前記一部分と、発話データに対応する第1の発話フレームとの連結を備える、請求項16に記載のデバイス。

【請求項20】

前記連結が、前記発話データの第2の発話フレームを前記代替発話データの前記一部分で置き換えることと、前記代替発話データの前記一部分を前記第1の発話フレームと接合することとを備え、前記第2の発話フレームが歪んだ発話データを備える、請求項19に記載のデバイス。

【請求項21】

複数の代替発話信号に対応する代替発話データを記憶し、複数のロケータ値を記憶するように構成されるメモリであって、各ロケータ値が前記代替発話データの対応する部分と関連付けられる、メモリと、

発話信号を受信し、前記発話信号を発話データに変換するように構成されるマイクロフォンと、

前記メモリに電氣的に結合されたプロセッサとを備え、前記プロセッサが、

前記マイクロフォンから前記発話データを受信し、
前記発話データの少なくとも一部分に基づいて、前記発話データの前記一部分と一致
する前記代替発話データの特定の部分に対応するロケータ値を特定し、
前記ロケータ値の指示を含む出力データを生成する
ように構成される、デバイス。

【請求項 22】

前記ロケータ値が、前記代替発話データの前記特定の部分が記憶される前記メモリのメ
モリ位置を示すインデックス値に対応する、請求項21に記載のデバイス。

【請求項 23】

前記ロケータ値が、前記代替発話データの前記特定の部分が記憶される前記メモリのメ
モリ位置を示すポインタに対応する、請求項21に記載のデバイス。

【請求項 24】

前記ロケータ値の前記指示を含む前記出力データを送信するように構成される送信機を
さらに備える、請求項21に記載のデバイス。

【請求項 25】

前記プロセッサがさらに、前記代替発話データの第1のオーディオパラメータと前記発
話データの前記一部分の第2のオーディオパラメータとの比較に基づいて、前記発話デー
タの前記一部分と一致する前記代替発話データの前記特定の部分を選択するように構成さ
れる、請求項21に記載のデバイス。

【請求項 26】

前記プロセッサがさらに、
前記代替発話データの前記特定の部分と前記発話データの前記一部分との間の一致の品
質を示す第1の一致値を決定し、
前記第1の一致値を閾値と比較するように構成される、請求項21に記載のデバイス。

【請求項 27】

前記プロセッサがさらに、前記第1の一致値が前記閾値以上であるという決定に回答し
て、前記代替発話データの前記特定の部分を選択するように構成される、請求項26に記載
のデバイス。

【請求項 28】

複数の代替発話信号に対応する代替発話データを記憶するように構成されるメモリと、
前記代替発話データの再生時間長を示すグラフィカルユーザインターフェース(GUI)を
生成するように構成されるプロセッサと、
前記GUIを表示するように構成されるインターフェースとを備える、デバイス。

【請求項 29】

前記再生時間長が、代替発話データのフレームの品質と関連付けられ、前記メモリがさ
らに、前記代替発話データと複数のユーザ識別子との間の関連付けを記憶するように構成
され、各ユーザ識別子が異なる人物に対応し、前記代替発話データが、第1のユーザ識別
子に対応する第1の代替発話データと、第2のユーザ識別子に対応する第2の代替発話デー
タと、第3のユーザ識別子に対応する第3の代替発話データとを備える、請求項28に記載の
デバイス。

【請求項 30】

前記プロセッサがさらに、第1の代替発話データの第1の再生時間長が更新閾値を満たす
という決定に回答して、前記複数のユーザ識別子のうちの特定のユーザ識別子と関連付け
られる追加の代替発話データに対する要求を送信するように構成される、請求項29に記載
のデバイス。

【請求項 31】

前記GUIが、前記第1の代替発話データの削除を示す第1の命令を受け取るように構成さ
れる、請求項29に記載のデバイス。

【請求項 32】

前記GUIが、第1の代替発話データの第1の年齢、第2の代替発話データの第2の年齢、お

よび第3の代替発話データの第3の年齢を示し、前記代替発話データが実質的に歪んでいない発話信号に対応する、請求項29に記載のデバイス。

【請求項33】

代替発話データの年齢が前記代替発話データの取得の日付に対応する、請求項31に記載のデバイス。

【請求項34】

前記GUIが、前記複数のユーザ識別子のうちの第4のユーザ識別子と関連付けられる、第4の代替発話データの追加を示す第2の命令を受け取るように構成される、請求項29に記載のデバイス。

フロントページの続き

(72)発明者 ダニエル・ジャレット・シンダー
アメリカ合衆国・カリフォルニア・9 2 1 2 1 - 1 7 1 4 ・サン・ディエゴ・モアハウス・ドライ
ヴ・5 7 7 5

【外国語明細書】

2018205751000001.pdf