

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6259094号
(P6259094)

(45) 発行日 平成30年1月10日 (2018. 1. 10)

(24) 登録日 平成29年12月15日 (2017. 12. 15)

(51) Int. Cl.

F I

G 1 O L 15/28 (2013. 01)

G 1 O L 15/28 2 3 O K

G 1 O L 15/04 (2013. 01)

G 1 O L 15/04 3 O O Z

G 1 O L 15/25 (2013. 01)

G 1 O L 15/25

請求項の数 20 (全 25 頁)

(21) 出願番号 特願2016-536943 (P2016-536943)
 (86) (22) 出願日 平成26年12月8日 (2014. 12. 8)
 (65) 公表番号 特表2016-540250 (P2016-540250A)
 (43) 公表日 平成28年12月22日 (2016. 12. 22)
 (86) 国際出願番号 PCT/US2014/069110
 (87) 国際公開番号 W02015/088980
 (87) 国際公開日 平成27年6月18日 (2015. 6. 18)
 審査請求日 平成29年4月4日 (2017. 4. 4)
 (31) 優先権主張番号 14/100, 934
 (32) 優先日 平成25年12月9日 (2013. 12. 9)
 (33) 優先権主張国 米国 (US)

早期審査対象出願

(73) 特許権者 507364838
 クアルコム、インコーポレイテッド
 アメリカ合衆国 カリフォルニア 921
 21 サン ディエゴ モアハウス ドラ
 イブ 5775
 (74) 代理人 100108453
 弁理士 村山 靖彦
 (74) 代理人 100163522
 弁理士 黒田 晋平
 (72) 発明者 ケー・ヒュン・パク
 アメリカ合衆国・カリフォルニア・921
 21-1714・サン・ディエゴ・モアハ
 ウス・ドライブ・5775

最終頁に続く

(54) 【発明の名称】 コンピューティングデバイスの音声認識プロセスの制御

(57) 【特許請求の範囲】

【請求項 1】

コンピューティングデバイスの音声認識プロセスを制御する方法であって、
 マイクロフォンからオーディオ信号を受信するステップと、
 前記マイクロフォンから前記オーディオ信号を受信すると同時に、前記コンピュー
 ティングデバイスのユーザの頭部筋肉活動を測定するように構成され、配置された筋肉運動
 検出器から頭部筋肉活動信号を受信するステップと、
 前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマ
 ンドを発信していることを示すオーディオ特性基準を満たすかどうか決定するステップで
 あって、前記オーディオ特性基準が、前記ユーザが前記コンピューティングデバイスにオ
 ーディオコマンドを発信していることを示す第1の2乗平均(RMS)しきい値レベルを含む、
 決定するステップと、

前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話している
 ことを示す筋肉運動基準を満たすかどうかを決定するステップと、

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマ
 ンドを発信していることを示す前記オーディオ特性基準を満たすことと、前記頭部筋肉活
 動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋
 肉運動基準を満たすことの両方を決定することに応答して、ボイス認識アルゴリズムによ
 って前記オーディオ信号を処理するステップであって、前記オーディオ信号が、前記オー
 ディオ信号のRMSレベルが前記第1のRMSしきい値レベルを超えないとき、前記オーディオ

10

20

特性基準を満たす、処理するステップと、

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示す前記オーディオ特性基準を満たさないか、または前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動基準を満たさないかのどちらかの間に、前記ボイス認識アルゴリズムによって前記オーディオ信号を処理しないステップであって、前記オーディオ信号が、前記オーディオ信号の前記RMSレベルが前記第1のRMSしきい値レベルを超えると、前記オーディオ特性基準を満たさない、処理しないステップとを含む、方法。

【請求項2】

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示すオーディオ特性基準を満たすかどうかを決定するステップが、

前記オーディオ信号を音声の特性音色パターンと比較するステップと、

前記オーディオ信号が、音色パターン整合しきい値を超える程度に、音声の前記特性音色パターンに整合するかどうかを決定するステップと、

前記オーディオ信号が、前記音色パターン整合しきい値を超える程度に、音声の前記特性音色パターンに整合すると決定することに応答して、前記オーディオ信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す前記オーディオ特性基準を満たすと決定するステップと

を含む、請求項1に記載の方法。

【請求項3】

前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定するステップが、前記頭部筋肉活動信号が音声を示す筋肉運動信号大きさしきい値を超えるかどうかを決定するステップを含む、請求項1に記載の方法。

【請求項4】

前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定するステップが、

前記頭部筋肉活動信号を前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動信号パターンと比較するステップと、

前記頭部筋肉活動信号が、筋肉活動パターン整合しきい値を超える程度に、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動信号パターンに整合するかどうかを決定するステップと、

前記頭部筋肉活動信号が、前記筋肉活動パターン整合しきい値を超える程度に、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動信号パターンに整合すると決定することに応答して、前記頭部筋肉活動信号が、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動基準を満たすと決定するステップと

を含む、請求項1に記載の方法。

【請求項5】

前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動信号パターンが、前記ユーザによるトレーニングから学習されかつ前記コンピューティングデバイスのメモリ内に記憶された頭部筋肉活動信号を含む、請求項4に記載の方法。

【請求項6】

メモリと、

マイクロフォンと、

筋肉運動検出器と、

前記メモリと、前記マイクロフォンと、前記筋肉運動検出器とに結合されたプロセッサと

10

20

30

40

50

を含む、コンピューティングデバイスであって、前記プロセッサが、

前記マイクロフォンからオーディオ信号を受信することと、

前記マイクロフォンから前記オーディオ信号を受信すると同時に、前記コンピューティングデバイスのユーザの頭部筋肉活動を測定するように構成され、配置された前記筋肉運動検出器から頭部筋肉活動信号を受信することと、

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示すオーディオ特性基準を満たすかどうか決定することであって、前記オーディオ特性基準が、前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示す第1の2乗平均(RMS)しきい値レベルを含む、決定することと、

10

前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定することと、

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示す前記オーディオ特性基準を満たすことと、前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動基準を満たすことの両方を決定することに応答して、ボイス認識アルゴリズムによって前記オーディオ信号を処理することであって、前記オーディオ信号が、前記オーディオ信号のRMSレベルが前記第1のRMSしきい値レベルを超えないとき、前記オーディオ特性基準を満たす、処理することと、

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示す前記オーディオ特性基準を満たさないか、または前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動基準を満たさないかのどちらかの間に、前記ボイス認識アルゴリズムによって前記オーディオ信号を処理しないことであって、前記オーディオ信号が、前記オーディオ信号の前記RMSレベルが前記第1のRMSしきい値レベルを超えると、前記オーディオ特性基準を満たさない、処理しないことと

20

を含む動作を実行するためのプロセッサ実行可能命令で構成される、コンピューティングデバイス。

【請求項7】

30

前記プロセッサが、前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示すオーディオ特性基準を満たすかどうかを決定することが、

前記オーディオ信号を音声の特性音色パターンと比較することと、

前記オーディオ信号が、音色パターン整合しきい値を超える程度に、音声の前記特性音色パターンに整合するかどうかを決定することと、

前記オーディオ信号が、前記音色パターン整合しきい値を超える程度に、音声の前記特性音色パターンに整合すると決定することに応答して、前記オーディオ信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す前記オーディオ特性基準を満たすと決定することと

40

を含むように動作を実行するためのプロセッサ実行可能命令で構成される、請求項6に記載のコンピューティングデバイス。

【請求項8】

前記プロセッサが、前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定することが、前記頭部筋肉活動信号が音声を示す筋肉運動信号大きさをしきい値を超えるかどうかを決定することを含むように動作を実行するためのプロセッサ実行可能命令で構成される、請求項6に記載のコンピューティングデバイス。

【請求項9】

前記プロセッサが、前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユ

50

ーザが発話していることを示す筋肉運動基準を満たすかどうかを決定することが、

前記頭部筋肉活動信号を前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動信号パターンと比較することと、

前記頭部筋肉活動信号が、筋肉活動パターン整合しきい値を超える程度に、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動信号パターンに整合するかどうかを決定することと、

前記頭部筋肉活動信号が、前記筋肉活動パターン整合しきい値を超える程度に、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動信号パターンに整合すると決定することに応答して、前記頭部筋肉活動信号が、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動基準を満たすと決定

10

を含むように動作を実行するためのプロセッサ実行可能命令で構成される、請求項6に記載のコンピューティングデバイス。

【請求項10】

前記プロセッサが、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動信号パターンが、前記ユーザによるトレーニングから学習されかつ前記コンピューティングデバイスの前記メモリ内に記憶された頭部筋肉活動信号を含むように動作を実行するためのコンピュータ実行可能命令で構成される、請求項9に記載のコンピューティングデバイス。

【請求項11】

20

コンピューティングデバイスであって、

マイクロフォンからオーディオ信号を受信するための手段と、

前記マイクロフォンから前記オーディオ信号を受信すると同時に、前記コンピューティングデバイスのユーザの頭部筋肉活動を測定するように構成され、配置された筋肉運動検出器から頭部筋肉活動信号を受信するための手段と、

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示すオーディオ特性基準を満たすかどうか決定するための手段であって、前記オーディオ特性基準が、前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示す第1の2乗平均(RMS)しきい値レベルを含む、決定するための手段と、

30

前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定するための手段と、

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示す前記オーディオ特性基準を満たすことと、前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動基準を満たすことの両方を決定することに応答して、ボイス認識アルゴリズムによって前記オーディオ信号を処理するための手段であって、前記オーディオ信号が、前記オーディオ信号のRMSレベルが前記第1のRMSしきい値レベルを超えないとき、前記オーディオ特性基準を満たす、処理するための手段と、

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示す前記オーディオ特性基準を満たさないか、または前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動基準を満たさないかのどちらかの間に、前記ボイス認識アルゴリズムによって前記オーディオ信号を処理しないための手段であって、前記オーディオ信号が、前記オーディオ信号の前記RMSレベルが前記第1のRMSしきい値レベルを超えると、前記オーディオ特性基準を満たさない、処理しないための手段と

40

を含む、コンピューティングデバイス。

【請求項12】

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示すオーディオ特性基準を満たすかどうかを決定するための手

50

段が、

前記オーディオ信号を音声の特性音色パターンと比較するための手段と、

前記オーディオ信号が、音色パターン整合しきい値を超える程度に、音声の前記特性音色パターンに整合するかどうかを決定するための手段と、

前記オーディオ信号が、前記音色パターン整合しきい値を超える程度に、音声の前記特性音色パターンに整合すると決定することに応答して、前記オーディオ信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す前記オーディオ特性基準を満たすと決定するための手段と

を含む、請求項11に記載のコンピューティングデバイス。

【請求項13】

10

前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定するための手段が、前記頭部筋肉活動信号が音声を示す筋肉運動信号大きさをしきい値を超えるかどうかを決定するための手段を含む、請求項11に記載のコンピューティングデバイス。

【請求項14】

前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定するための手段が、

前記頭部筋肉活動信号を前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動信号パターンと比較するための手段と、

前記頭部筋肉活動信号が、筋肉活動パターン整合しきい値を超える程度に、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動信号パターンに整合するかどうかを決定するための手段と、

20

前記頭部筋肉活動信号が、前記筋肉活動パターン整合しきい値を超える程度に、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動信号パターンに整合すると決定することに応答して、前記頭部筋肉活動信号が、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動基準を満たすと決定するための手段と

を含む、請求項11に記載のコンピューティングデバイス。

【請求項15】

前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動信号パターンが、前記ユーザによるトレーニングから学習されかつ前記コンピューティングデバイスのメモリ内に記憶された頭部筋肉活動信号を含む、請求項14に記載のコンピューティングデバイス。

30

【請求項16】

コンピューティングデバイスのプロセッサに、

マイクロフォンからオーディオ信号を受信することと、

前記マイクロフォンから前記オーディオ信号を受信すると同時に、前記コンピューティングデバイスのユーザの頭部筋肉活動を測定するように構成され、配置された筋肉運動検出器から頭部筋肉活動信号を受信することと、

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示すオーディオ特性基準を満たすかどうか決定することであって、前記オーディオ特性基準が、前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示す第1の2乗平均(RMS)しきい値レベルを含む、決定することと、

40

前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定することと、

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示す前記オーディオ特性基準を満たすことと、前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動基準を満たすことの両方を決定することに応答して、ボイス認識アルゴリズムによ

50

って前記オーディオ信号を処理することであって、前記オーディオ信号が、前記オーディオ信号のRMSレベルが前記第1のRMSしきい値レベルを超えないとき、前記オーディオ特性基準を満たす、処理することと、

前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示す前記オーディオ特性基準を満たさないか、または前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動基準を満たさないかのどちらかの間に、前記ボイス認識アルゴリズムによって前記オーディオ信号を処理しないことであって、前記オーディオ信号が、前記オーディオ信号の前記RMSレベルが前記第1のRMSしきい値レベルを超えると、前記オーディオ特性基準を満たさない、処理しないことと

10

を含む動作を実行させるように構成されたプロセッサ実行可能命令を記憶した、プロセッサ可読記録媒体。

【請求項 17】

前記記憶されたプロセッサ実行可能命令が、前記コンピューティングデバイスに、前記オーディオ信号が前記ユーザが前記コンピューティングデバイスにオーディオコマンドを発信していることを示すオーディオ特性基準を満たすかどうかを決定することが、

前記オーディオ信号を音声の特性音色パターンと比較することと、

前記オーディオ信号が、音色パターン整合しきい値を超える程度に、音声の前記特性音色パターンに整合するかどうかを決定することと、

前記オーディオ信号が、前記音色パターン整合しきい値を超える程度に、音声の前記特性音色パターンに整合すると決定することに応答して、前記オーディオ信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す前記オーディオ特性基準を満たすと決定することと

20

を含むような動作を実行させるように構成される、請求項16に記載のプロセッサ可読記録媒体。

【請求項 18】

前記記憶されたプロセッサ実行可能命令が、前記コンピューティングデバイスに、前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定するステップが、前記頭部筋肉活動信号が音声を示す筋肉運動信号大きさしきい値を超えるかどうかを決定することを含むように動作を実行させるように構成される、請求項16に記載のプロセッサ可読記録媒体。

30

【請求項 19】

前記記憶されたプロセッサ実行可能命令が、前記コンピューティングデバイスに、前記頭部筋肉活動信号が前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定することが、

前記頭部筋肉活動信号を前記コンピューティングデバイスの前記ユーザが発話していることを示す筋肉運動信号パターンと比較することと、

前記頭部筋肉活動信号が、筋肉活動パターン整合しきい値を超える程度に、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動信号パターンに整合するかどうかを決定することと、

40

前記頭部筋肉活動信号が、前記筋肉活動パターン整合しきい値を超える程度に、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動信号パターンに整合すると決定することに応答して、前記頭部筋肉活動信号が、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動基準を満たすと決定することと

を含むように動作を実行させるように構成される、請求項16に記載のプロセッサ可読記録媒体。

【請求項 20】

前記記憶されたプロセッサ実行可能命令が、前記コンピューティングデバイスに、前記コンピューティングデバイスの前記ユーザが発話していることを示す前記筋肉運動信号パ

50

ターンが、前記ユーザによるトレーニングから学習されかつ前記コンピューティングデバイスのメモリ内に記憶された頭部筋肉活動信号を含むように動作を実行させるように構成される、請求項19に記載のプロセッサ可読記録媒体。

【発明の詳細な説明】

【背景技術】

【0001】

消費者に利用可能な装着型コンピューティングデバイスの数は今後4年にわたり5倍に増大することを調査は示している。装着型コンピューティングデバイス用のユーザ入力、ある種の装着型コンピューティングデバイスの成功の主な決定要因になる。装着型コンピューティングデバイスに関する1つのそのようなユーザ入力方法は、ボイス(voice)認識を含む。ボイス認識アルゴリズムは、ユーザの音声を復号して、ユーザが意図するコマンドを決定する。

10

【発明の概要】

【課題を解決するための手段】

【0002】

様々な実施形態の方法およびデバイスは、ユーザが発話していないときに電力と処理リソースとを節約するために、ユーザが発話しているときだけ、マイクロフォンからのオーディオ信号および筋肉運動検出器(muscle movement detector)からの頭部筋肉(head muscle)活動信号を受信するコンピューティングデバイスが、音声認識プロセスをアクティブ化することを可能にする。実施形態の方法およびデバイスは、マイクロフォンからオーディオ信号を受信するステップと、マイクロフォンからオーディオ信号を受信すると同時に、コンピューティングデバイスのユーザの頭部筋肉活動を測定するように構成され、配置された筋肉運動検出器から筋肉活動信号を受信するステップと、オーディオ信号がコンピューティングデバイスのユーザが発話していることを示すオーディオ特性基準を満たすかどうかを決定するステップと、頭部筋肉活動信号がコンピューティングデバイスのユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定するステップとを含み得る。実施形態の方法およびデバイスはまた、オーディオ信号がコンピューティングデバイスのユーザが発話していることを示すオーディオ特性基準を満たすことと、頭部筋肉活動信号がコンピューティングデバイスのユーザが発話していることを示す頭部筋肉運動基準を満たすことの両方を決定することに応答して、ボイス認識アルゴリズムによってオーディオ信号を処理するステップと、オーディオ信号がコンピューティングデバイスのユーザが発話していることを示すオーディオ特性基準を満たさないか、または頭部筋肉活動信号がコンピューティングデバイスのユーザが発話していることを示す筋肉運動基準を満たさないかのどちらかの間に、ボイス認識アルゴリズムによってオーディオ信号を処理しないステップとを含み得る。

20

30

【0003】

いくつかの実施形態の方法およびデバイスでは、オーディオ信号がコンピューティングデバイスのユーザが発話していることを示すオーディオ特性基準を満たすかどうかを決定するステップは、オーディオ信号の2乗平均(RMS)レベルがRMSしきい値レベルを超えるかどうかを決定するステップを含み得る。

40

【0004】

いくつかの実施形態の方法およびデバイスでは、オーディオ信号がコンピューティングデバイスのユーザが発話していることを示すオーディオ特性基準を満たすかどうかを決定するステップは、受信したオーディオ信号を音声の特性音色パターンと比較するステップと、受信したオーディオ信号が、音色パターン整合しきい値を超える程度に、音声の特性音色パターンに整合するかどうかを決定するステップと、受信したオーディオ信号が、音色パターン整合しきい値を超える程度に、音声の特性音色パターンに整合することを決定することに応答して、オーディオ信号がコンピューティングデバイスのユーザが発話していることを示すオーディオ特性基準を満たすこと決定するステップとを含み得る。

【0005】

50

いくつかの実施形態の方法およびデバイスでは、頭部筋肉活動信号がコンピューティングデバイスのユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定するステップは、受信した頭部筋肉活動信号が音声を示す筋肉運動信号大きさしきい値(magnitude threshold)を超えるかどうかを決定するステップを含み得る。

【0006】

いくつかの実施形態の方法およびデバイスでは、頭部筋肉活動信号がコンピューティングデバイスのユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定するステップは、受信した頭部筋肉活動信号をコンピューティングデバイスのユーザが発話していることを示す筋肉運動信号パターンと比較するステップと、受信した頭部筋肉活動信号が、筋肉活動パターン整合しきい値を超える程度に、コンピューティングデバイスのユーザが発話していることを示す特性頭部筋肉活動信号パターンに整合するかどうかを決定するステップと、受信した頭部筋肉活動信号が、筋肉活動パターン整合しきい値を超える程度に、コンピュータデバイスのユーザが発話していることを示す特性頭部筋肉活動信号パターンに整合することを決定することに応答して、受信した頭部筋肉活動信号が、コンピューティングデバイスのユーザが発話していることを示す筋肉運動基準を満たすかどうかを決定するステップとを含み得る。様々な実施形態の方法およびデバイスの筋肉運動信号パターンは、ユーザによるトレーニングから学習され、コンピューティングデバイスのメモリ内に記憶された頭部筋肉活動信号を含み得る。

【0007】

添付の図面は、本開示の実施形態の説明を助けるために提示され、実施形態の限定ではなく、実施形態の例示のためのみに提供される。

【図面の簡単な説明】

【0008】

【図1】様々な実施形態とともに使用するのに適した装着型コンピューティングデバイスの透視図である。

【図2】図1のコンピューティングデバイスを装着しているユーザの図である。

【図3】様々な実施形態とともに使用するのに適した装着型コンピューティングデバイスの概略機能ブロック図である。

【図4】マイクロフォンおよびEMGセンサーから受信したオーディオデータならびにEMGデータのグラフである。

【図5】ボイス認識に関するオーディオ信号を処理するかどうかを決定するための実施形態の方法を示すプロセスフロー図である。

【図6】受信したオーディオ信号がオーディオ特性基準を満たすかどうかを決定するための実施形態の方法を示すプロセスフロー図である。

【図7】受信した頭部筋肉活動が筋肉運動基準を満たすかどうかを決定するための実施形態の方法を示すプロセスフロー図である。

【発明を実施するための形態】

【0009】

様々な実施形態について、添付図面を参照して詳細に説明する。可能な場合には常に、同じ部分または同様の部分を指すために、図面全体を通じて同じ参照番号が使用される。特定の例および実装形態に対して行われる言及は、説明を目的とし、本開示の範囲または特許請求の範囲を限定するものではない。本開示の範囲から逸脱することなく、代替実装形態が考案され得る。加えて、本開示の周知の要素は、本開示の関連する詳細を不明瞭にしないように、詳細には説明されず、または省略される。

【0010】

「例示的」という言葉は、本明細書では、「例、事例、または例示として役立つ」ことを意味するように使用される。「例示的」として本明細書で説明する任意の実装形態は、必ずしも他の実装形態よりも好ましいか、または有利であると解釈されるべきではない。加えて、「第1の」および「第2の」という単語、または同様の言い回しは、様々な説明の要素を区別する明確な目的のためであることが意図されており、特許請求の範囲を要素の

10

20

30

40

50

特定の順序または階層に限定することは意図されていない。

【0011】

ボイス認識は、大量の電力を消費し得る計算集約型プロセスである。ユーザが話していないときにオーディオ入力(たとえば、周囲環境の音楽)に関してボイス認識処理ルーチンを実行することは、エネルギーを浪費し、バッテリー電源式デバイスは、充電されなければならない前に、そのデバイスが動作し得る時間を制限する可能性がある。したがって、背景雑音を含めて、すべての受信したオーディオ入力に関してボイス認識ルーチンを常時処理することは、装着型コンピューティングデバイスを用いたユーザの経験に悪影響を及ぼし得る。

【0012】

様々な実施形態は、受信したオーディオ信号がオーディオ特性基準を満たし、受信した頭部筋肉活動信号が、ユーザが発話していることを示す筋肉運動基準を満たすときだけ、ボイス認識アルゴリズムによってオーディオ信号および/または頭部筋肉活動信号(たとえば、EMG信号)の処理が実行される装着型コンピューティングデバイス内の電力を節約することに関する。様々な実施形態は、不要なボイス認識処理動作を最小限に抑え、それにより、装着型コンピューティングデバイス(たとえば、スマートグラス)の限定されたバッテリー電力を節約する。

【0013】

様々な実施形態では、コンピューティングデバイスは、ユーザが発話している(たとえば、信号強度がしきい値レベルを超えない)ときの筋肉の動きを検出するために、筋電図検査(EMG)センサーなどの1つまたは複数の頭部筋肉運動検出器を使用することができる。また、1つまたは複数のマイクロフォンは、ユーザが発話している(たとえば、受信したオーディオがしきい値未満のRMSレベルを有する)とき、オーディオ信号を受信することができる。マイクロフォンから受信したオーディオ信号と、筋肉運動検出器から受信した頭部筋肉活動信号の両方から行われた決定に基づいて、コンピューティングデバイスは、ボイス認識アルゴリズムをアクティブ化するかどうかを決定することができる。このようにして、デジタル信号プロセッサ(DSP)などのプロセッサは、ユーザが発話している可能性が高いときだけ、計算的に要求の高い音声認識アルゴリズムを使用してオーディオ信号および/または筋肉活動信号を処理する。受信したオーディオ信号と受信した筋肉活動信号の両方からのしきい値が満たされない限り、コンピューティングデバイスは、ボイス認識アルゴリズムを非アクティブ化し、受信したオーディオ信号および/または受信した筋肉活動信号を処理することを控えることができる。

【0014】

「装着型コンピューティングデバイス」または「スマートグラス」という用語は、本明細書では、受信したオーディオ信号をボイス認識アルゴリズムによって処理するように構成されたプロセッサを備えたコンピューティングデバイスを指すために、交換可能に使用される。装着型コンピューティングデバイスは、マイクロフォン、筋肉運動センサー、入力機構、トランシーバ、ディスプレイ、およびカメラなど、1つまたは複数の他の構成要素を有し得る。装着型コンピューティングデバイスは、スマートグラスまたはスマートウォッチなど、装着可能なデバイスを含むが、様々な実施形態の方法に関して、ワイヤードまたはワイヤレスの筋肉運動検出器を用いてオーディオ信号を受信することができる任意のコンピューティングデバイスが想定される。たとえば、説明した方法の実施形態では、遠隔筋肉運動センサーを備えたモバイルフォンを使用することができる。

【0015】

本明細書で使用する場合、「マイクロフォン」という用語は、直近の周囲環境からの音を拾い上げ、それを電気信号に変換するコンピューティングデバイスの入力トランスデューサを指す。

【0016】

本明細書で使用する場合、「頭部筋肉活動」または「筋肉活動」という用語は、人間の音声に関連する(顔面筋を含む)何らかの頭部筋肉の活動を指す。人間の音声に関連する1

10

20

30

40

50

つの頭部筋肉は、収縮するとき、顎を動かす側頭筋(temporal muscle)すなわち側頭筋(temporalis)である(たとえば、下顎が上がる時側頭筋は収縮する)。人間の音声に関連するもう1つの筋肉は、やはり下顎を上げるのを助ける咬筋である。様々な実施形態では、頭部筋肉運動検出器は、側頭筋または咬筋のどちらか、ならびに人間の音声に関連する、他の頭部筋肉、顔面筋、およびの顎筋の動きを検出するように配置され得る。

【0017】

本明細書で使用する場合、「頭部筋肉運動検出器」または「筋肉運動検出器」という用語は、頭部筋肉活動、特に、発話しているユーザまたはコンピューティングデバイスの装着者に関連する筋肉の動き(たとえば、側頭筋の動き)を検出することができるセンサーを指す。筋肉運動検出器は、頭部筋肉および/または頭の動きに関連する表皮の動きの表示である頭部筋肉運動活動を受信することが可能である。様々な実施形態では、筋肉運動検出器は、特に、発話に関連する筋肉の運動を検出するのに適切であり得、かつ/またはそのように位置し得る。一実施形態による例示的な筋肉運動検出器は筋電図検査(EMG)センサーである。EMGは、骨格筋によって生み出される電気的活動を評価し、記録するための技法である。

10

【0018】

EMGセンサーは、それらの細胞が電氣的または神経学的にアクティブ化されたときに筋肉細胞によって生成される電位の形で信号を検出することができる。信号を分析して、人物の発話に対応する顎の動きなど、人間の生体力学を検出することができる。EMGセンサーは、筋肉繊維が収縮するとき、筋肉繊維によって生成される小さな電氣的刺激を検出し、それを増幅することによって、筋肉運動活動を測定することができる。

20

【0019】

もう1つの形態の筋肉運動検出器は、筋肉の動き、組織変位、および/または電極変形によって引き起こされる変化を検出することができる、皮膚に接触するように配置された、1つもしくは複数の導電性組織電極(conductive textile electrodes)を含み得る。さらなる筋肉運動検出器は、特に、装着者の顎における、またはその付近の表皮変化を検出するように構成された圧力センサーであり得る。もう1つの筋肉運動検出器は、骨を含めて、装着者の組織を通して伝導される音を検出するように構成されたマイクロフォンであり得る。

【0020】

30

図1は、本実施形態の方法を実装するのに適したスマートグラスの形態の装着型コンピューティングデバイス100の透視図を示す。スマートグラスが示されるが、任意のモバイルコンピューティングデバイスが本実施形態の方法を実装することができる。しかしながら、明快および簡潔のために、限定せずに、これらの実施形態は、図1に示すスマートグラスなどの装着型コンピューティングデバイスを参照して説明される。

【0021】

装着型コンピューティングデバイス100は、音を受信して、オーディオ信号に変換するための、1つまたは複数のマイクロフォン104a、104bを有し得る。マイクロフォン104は、ユーザのボイスからの音と、背景で演奏している音楽または騒々しい雑談など、ユーザのボイスからではない音とを含めて、周囲環境内の音を受信する。装着型コンピューティングデバイスは、ユーザからオーディオを直接受信するように配置され、構成されたマイクロフォン104bを含み得るが、異なるマイクロフォン104aが周囲環境からの(すなわち、ユーザからではない)オーディオを受信するように配置され、構成され得る。様々な実施形態の装着型コンピューティングデバイスは、1つまたは複数のバッテリーなど、電源を含み得る。

40

【0022】

装着型コンピューティングデバイス100はまた、筋肉の動きを検出し、頭部筋肉活動信号を生成するように構成された、1つまたは複数の頭部筋肉運動検出器102a、102bを含み得る。筋肉運動検出器は、筋電図検査(EMG)センサーであり得る。冗長性を提供するため、および/または異なる筋群を感知するために、2つ以上の筋肉運動検出を使用することが

50

できる。筋肉運動検出器102a、102bは、装着者の皮膚の一部を、頭部筋肉の動き、特に、発話に関連する筋肉の動きを検出するのに適したロケーションに係合させるように構成され得る。このようにして、筋肉運動検出器102a、102bは、装着者が発話しているとき、頭部筋肉の動きを検出するように構成され得る。

【0023】

装着型コンピューティングデバイス100は、1つまたは複数の入力機構106(たとえば、ボタン)を含み得る。入力機構は、ユーザが入力機構を押下している間、ユーザが、オーディオ信号および/または筋肉運動信号の処理を無効にして、音声認識処理を直接アクティブ化することを可能にし得る。このようにして、特に、ユーザがボイス処理のためのコマンドを目的として、装着型コンピューティングデバイスに音声を導いているとき、ユーザは、デバイスが、発話している間に、ユーザのオーディオおよび/または頭部筋肉活動を学習することを可能にし得る。

10

【0024】

入力機構は、ユーザが、意思決定プロセスを補正すること、または音声を示すオーディオ基準もしくは筋肉運動基準を修正することを可能にし得る。たとえば、ユーザは、音声がボイス認識アルゴリズムによって処理されずに(たとえば、オーディオ特性基準または筋肉運動基準が満たされなかった)、装着型コンピューティングデバイス100に対して発話することができる。未登録音声の直後に、ユーザは、オーディオ入力または筋肉活動入力を処理しない判定は不正確であったことをプロセッサにシグナリングする入力機構を用いる(たとえば、ボタンを押下する)ことができる。装着型コンピューティングデバイスのプロセッサは、ユーザの音声からのオーディオを処理しないという誤りを明らかにして、補正するために、オーディオ特性基準または筋肉運動基準のどちらか、あるいは両方を更新/修正することができる。

20

【0025】

装着型コンピューティングデバイス100は、ワイヤレストランシーバ108を含み得る。ワイヤレストランシーバ108は、ワイヤレス広域ネットワーク(WWAN)(たとえば、LTEまたは任意のセルラー接続)およびワイヤレスローカルエリアネットワーク(WLAN)(たとえば、Wi-Fi)など、ワイヤレス通信ネットワークとの間でワイヤレス信号を送受信することが可能であり得る。装着型コンピューティングデバイスは、ボイス認識データ(たとえば、音声を示すRMSレベル、音色パターン、EMG信号強度、およびEMG信号パターン)を送信ならびに/または受信するためにネットワークサーバに接続することができる。ボイス認識データは、受信したオーディオ信号および頭部筋肉活動信号の処理を更新するためのユーザによるコマンドに基づいて、周期的に受信されるか、または取り出されることが可能である。

30

【0026】

装着型コンピューティングデバイス100は、様々な実施形態の動作を実行するためのプロセッサ実行可能命令を用いて構成された、1つもしくは複数のプロセッサ、コア、またはコプロセッサ110をさらに含み得る。プロセッサは、ボイス認識アルゴリズムによるオーディオ信号および/もしくは筋肉運動信号の処理をアクティブ化または非アクティブ化するかどうかを決定するために、オーディオ信号を受信するための1つもしくは複数のマイクロフォン104a、104bと、筋肉運動信号を受信するための筋肉運動検出器102a、102bとに結合され得る。プロセッサ110はまた、入力機構106と、トランシーバ108と、カメラ103とに結合され得る。

40

【0027】

プロセッサ110は、音声を示すオーディオ信号および/または筋肉運動活動信号の特性パターンが記憶され得るメモリ112に結合され得る。そのような記憶されたパターンは、ユーザが発話しているときをより良好に区別するために、下で説明するように、プロセッサ110によって、受信した信号と比較され得る。

【0028】

図2は、図1に示す装着型コンピューティングデバイス100を装着しているユーザ201を示す。示すように、筋肉運動検出器に関する1つの考えられるロケーションは、側頭筋の筋

50

肉の動きを検出するために、ユーザ201の耳の真上である。筋肉運動検出器はスマートグラスに直接接続されることとして示されるが、筋肉運動検出器は、音声の間に収縮するいずれの筋肉から筋肉の動きを検出し、スマートグラスもしくは他のコンピューティングデバイスにワイヤレスで接続するために、ユーザの頭または顔の上のどこかに配置された別個のデバイスであり得る。

【0029】

図3は、装着型コンピューティングデバイス300の電子構成要素の例示的な概略機能ブロック図を示す。上で論じたように、装着型コンピューティングデバイスは、入力されたオーディオ信号を受信するための(図1に示した、1つまたは複数のマイクロフォン104a、104bを含み得る)マイクロフォン310を含む。たとえば、マイクロフォン310は、装着者自身の口からの音をより容易に拾い上げるために(ユーザの口に近接してなど)装着型コンピューティングデバイス上に配置され得る第1のマイクロフォン104bを含み得るのに対して、第2のマイクロフォン104aは、主に周囲音を拾い上げるように、(ユーザの口から離してなど)デバイス上に配置され得る。マイクロフォン310によって表される1つまたは複数のマイクロフォンが存在するかどうかにかかわらず、マイクロフォン310は、アナログ/デジタル(A/D)変換器315を通して処理され得る音を受信し、検出した/受信した音のデジタル信号を生成する。受信したオーディオ信号は、ボイス活動分析器320によって処理され得る。

【0030】

ボイス活動分析器320は、受信したオーディオ信号内にボイスが存在するかどうかを検出するユニットであり得る。一実施形態では、ボイス活動分析器は、オーディオ信号の測定された2乗平均(RMS)レベルを、音声を示す、メモリ内に記憶されたRMSレベルと比較することができる。音声を示すRMSレベルは、デバイス上に前に記憶されている可能性があり(すなわち、人間の音声に関連する一般的なRMSレベル)、コンピューティングデバイスを使用して、個人の音声パターンを反映するために、経時的に学習され得る。ユーザが発話していることを示すRMSレベルは、特定のデシベル限界の形態であってよい。たとえば、ユーザは、食卓の向こうの人物に話すよりも穏やかなレベルでその装着型コンピューティングデバイスにコマンドを発話する可能性があり、したがって、ユーザが、ボイスコマンドを示すために、その装着型コンピューティングデバイスにおいて声を上げる(たとえば、高いデシベルレベル)可能性は低い。ボイス活動分析器は、受信したオーディオ信号の測定されたRMS値をRMSしきい値レベルと比較することによって、ユーザが装着型コンピューティングデバイスに対して発話していること、またはそのデバイスに音声を導いていることを検出することができる。受信したオーディオ信号がしきい値レベルを上回るRMSレベルを有する場合、ボイス活動検出は、ユーザがコンピューティングデバイスに対して発話していないと決定することができる。

【0031】

一実施形態では、ボイス活動分析器320は、ユーザの音声パターンを特徴付ける、1つまたは複数の基本周波数を認識するために、受信したオーディオ信号に周波数分析を適用することによって、装着者のボイスを特に検出するように構成され得る。したがって、ボイス活動分析器320は、ある時間期間にわたってユーザを聴取することによって、デバイスが学習した音色パターンなど、受信した音を示すデジタル化された表示と、人間の音声と相関された周波数パターン、または音色パターンの1つもしくは複数のセットとを比較することによって、ボイス検出器として機能し得る。たとえば、ボイス活動分析器320および装着型コンピューティングデバイスのセットアップの一環として、ユーザは、声を出して段落を読むこと、または何らかの言い回しを繰り返すことによって、デバイスをトレーニングするように要求されてよく、コンピューティングデバイスは、この入力を使用して、メモリ内に記憶され得るユーザの音声の周波数パターン、音色パターン、および/またはRMSレベルを識別し、ユーザによるボイスコマンドを認識するために音響信号を分析する際に後で使用することができる。代替的に、1つもしくは複数の指定された言語、単語、および/または、さらに音節に関連する音の同期パターンならびに調和構造の存在を使用して、ボイス活動を識別することができる。このようにして、ボイス活動分析器320は

、オーディオ信号パターンによって表された特性など、受信したオーディオ信号の少なくとも一部が装着者によって生成された音声に関連する第1の音色パターンの同様の特性に整合するかどうかを決定することができる。オーディオ信号と音色パターンの整合は、(各々がオーディオ信号を表す)2つのパターンが実質的に同様であることを意味する。加えて、ボイス活動分析器320または雑音パターン分析器350は、コンピューティングデバイスがボイス認識処理を適用する必要がない(集合的に、「雑音」と呼ばれる)あらかじめ定義された音、望ましくない雑音、および/またはパターンを識別するフィルタとして機能し得る。雑音と識別される、受信したオーディオ信号の部分は、無視されてよいが、または、装着型コンピューティングデバイスがそのユーザが発話していると決定する場合、判定ルーチン370において、受信したオーディオ信号から減じられ得る雑音パターンとして標示されてもよい。このようにして、「雑音」と識別された部分は、装着型コンピューティングデバイスによって(人物の環境のオーディオ記録を捕捉するなど)記録目的のために記録され得るが、増幅または減衰されるとは限らない。人間の音声を含む入力されたオーディオ信号の部分は、その何らかのサブ部分が人間の音声および/または装着者の音声として識別されているかどうかについての何らかの表示とともに、判定ルーチン370に転送され得る。オーディオ信号の「雑音」と識別される部分は、判定ルーチン370によって減衰されてよく、または雑音パターン分析器350および/もしくはボイス活動分析器320によってフィルタリングされて完全に排除されてもよい。

【0032】

具体的には、音声に関連する頭部筋肉から頭部筋肉活動信号を受信するように配置された装着型コンピューティングデバイスの筋肉運動検出器330は、電圧差を測定するための表面電極を含み得るEMGセンサーであり得る。筋肉運動検出器330は、デバイスが装着されているとき、装着者の皮膚に直接接触するように配置され得る。たとえば、筋肉運動検出器330は、その動きが発話に関連するユーザ/装着者の頭または顔の部分(たとえば、図2に示すようにユーザの耳の後ろ)に接触する、装着型コンピューティングデバイス100の外側部分上に配置され得る。

【0033】

筋肉運動検出器330は、動きが検出されることを確実にする目的で、筋肉運動のパターンを検出/分化するために、および/または冗長性を提供するために、2つ以上の筋肉運動検出器を含み得る。たとえば、第1の筋肉運動検出器は、装着型コンピューティングデバイスの第1の部分上に配置されてよいのに対して、第2の筋肉運動検出器は、装着型コンピューティングデバイス100の第2の部分上の第1の筋肉運動検出器から離して、または主な装着型コンピューティングデバイス本体からすら離して配置されてよい。

【0034】

筋肉運動検出器330は、デバイスプロセッサによるデジタル処理のためにアナログ/デジタル(AD)変換器335によって処理され得る頭部筋肉活動の特徴付ける信号を生成する。受信された頭部筋肉活動は、増幅器フィルタ340によって、1つまたは複数の信号として処理され得る。

【0035】

増幅器フィルタ340は、ボイス活動分析器320が装着型コンピューティングデバイスのマイクロフォンからのオーディオを処理すると同時に、ユーザの頭部筋肉活動信号を処理することができる。増幅器フィルタが、ユーザの発話の動きに対応する筋肉運動信号を認識するとき、増幅器フィルタは高い増幅信号を筋肉運動分析器360に送ることができる。同様に、増幅器フィルタ340がその筋肉運動信号はユーザの音声に対応しないと決定したとき、増幅器フィルタ340は低い増幅信号を筋肉運動分析器360に送ることができる。これらの高い増幅信号または低い増幅信号はまた、装着型コンピューティングデバイスのプロセッサが、オーディオ信号と筋肉運動信号の両方でない場合、少なくとも受信したオーディオ信号に関してボイス認識プロセッサ380をアクティブ化するかどうかを判定することができるように、判定ルーチン370が、ユーザが発話していることに対応する可能性が最も高い、受信したオーディオ信号内の時間期間を決定するのに役立ち得る。

【0036】

筋肉運動分析器360は、元の筋肉運動信号(たとえば、元のEMG信号)に加えて、増幅器フィルタ340からオーディオ信号を受信することができる。筋肉運動分析器360は、筋肉運動信号を処理して、受信したEMG信号内の「雑音」を区別することができる。たとえば、ユーザが、発話するのにまさに先立って、および発話したまさにその後に、自らの口(または、顎)を動かすとき、筋肉運動検出器330は、筋肉運動信号(たとえば、EMG信号)を生成することができる。筋肉運動分析器360は、同時オーディオ区分、すなわち、並行オーディオ区分を伴わない筋肉運動信号が「雑音」と見なされ、同時オーディオ区分、すなわち、並行オーディオ区分を伴う筋肉運動信号と区別され得るように、受信した、ユーザの筋肉運動を特定のオーディオ区分と関連させることができる。たとえば、筋肉運動検出器330は、ユーザがガムを噛んでいるとき、筋肉運動信号を生成することができるが、ユーザは話していないため、マイクロフォン310はオーディオ信号を受信することができず、したがって、筋肉運動分析器360は同時オーディオ信号を受信していないため、筋肉運動分析器360はその筋肉運動信号が「雑音」とであると決定することができる。筋肉運動分析器360は、処理した筋肉運動信号を判定ルーチン370に送ることができる。

10

【0037】

一実施形態では、筋肉運動分析器360および装着型コンピューティングデバイス100のセットアップの一環として、コンピューティングデバイスがユーザの音声の筋肉運動信号強度および/または筋肉運動パターンを解析する間、ユーザは、声を出して段落を読むこと、または何らかの言い回しを繰り返すことによって、デバイスをトレーニングするように要求されてよい。検出された筋肉運動信号パターンは、ユーザによるボイスコマンドを認識するために音響信号および/または頭部筋肉活動信号を分析する際に後で使用するためにメモリ内に記憶され得る。代替的に、1つもしくは複数の指定された言語、単語、および/または、さらに音節に関連する同期筋肉運動パターンの存在を使用して、ボイス活動を識別することができる。これにより、筋肉運動分析器360は、そのユーザに一意の筋肉運動信号強度および/または筋肉運動パターンを学習することができる。

20

【0038】

判定ルーチン370で、コンピューティングデバイスのプロセッサは、頭部筋肉活動信号とオーディオ信号の両方の区分が、ユーザが発話していることに対応するかどうかを決定することができる。装着型コンピューティングデバイスが、頭部筋肉活動信号とオーディオ信号の両方が、ユーザが発話していることに対応すると決定する場合、オーディオ信号は、ボイス認識アルゴリズムによるボイス認識処理のためにボイス認識プロセッサに送られ得る。一実施形態では、ボイス認識処理がアクティブ化されたとき、頭部筋肉活動信号および筋肉運動信号は、オーディオ信号とともに、ボイス認識プロセッサ380に送られ得る。一実施形態では、判定ルーチン370を実行するプロセッサが、頭部筋肉活動信号とオーディオ信号の組合せから、そのユーザが発話していると決定したとき、頭部筋肉活動信号によって示され得る、そのユーザが発話しているときに対応する、オーディオデータの断片だけが、処理のためにボイス認識プロセッサ380に送られ得る。このようにして、筋肉運動信号を使用して、ボイス認識のために処理されるオーディオデータ381(および、オプションで、頭部筋肉活動データ382)をゲート制御して、バッテリー電力を節約して、他のタスクのために処理リソースを解放し得る、ボイス認識アルゴリズムによって処理される音データ(および、オプションで、筋肉活動データ)量の低減を行い、背景雑音、付近の他者のボイス、およびユーザが発話していることに属さない筋肉運動信号を除外することによって、エラーの削減を行うことができる。

30

40

【0039】

一実施形態では、判定ルーチン370において、プロセッサは、そのユーザに対して一意に識別可能なRMSレベルパターンおよび/または筋肉運動信号強度パターンを学習することができる。たとえば、各ユーザは、別のユーザとは異なる平均RMSレベルおよび/または筋肉運動信号強度を有し得る。耳の不自由な1人のユーザは、発話しているとき、耳が不自由でないもう1人のユーザよりも、その装着型コンピューティングデバイスにおいてより

50

大きな声で話し、顎をより大きく動かす可能性がある。したがって、プロセッサによって実装される判定ルーチン370は、そのユーザに関する個人化されたRMSしきい値および/もしくは筋肉運動信号強度しきい値を調整または作成することができる。一実施形態では、判定ルーチン370を実行するコンピューティングデバイスプロセッサが、受信したオーディオ信号および頭部筋肉活動信号が音声に対応すると決定するたびに、プロセッサは、RMSレベルおよび/または筋肉運動信号強度レベルをメモリ内に記憶することができる。別の実施形態では、ユーザは、入力機構(たとえば、図1の入力機構106)を押下することによって、受信したオーディオおよび/または受信した頭部筋肉活動信号の誤った解釈(たとえば、デバイスプロセッサがRMSレベルは音声を示さないと決定すること)を補正することができる。それに対して、プロセッサは、そのようなエラーの直後に、音声を示すとして、受信した信号を誤って解釈する際に使用した、記憶されたオーディオRMSしきい値および/もしくは筋肉活動信号しきい値を削除または調整することによって反応することができる。

【0040】

一実施形態では、判定ルーチン370を実装するプロセッサは、そのユーザに対して一意の音色パターンおよび/または筋肉運動パターンを学習することができる。一実施形態では、プロセッサが、受信したオーディオ信号および頭部筋肉活動信号がユーザの音声に対応すると判断するたびに、プロセッサは、そのオーディオ信号および頭部筋肉活動信号を使用して、ユーザの音声を認識するためにプロセッサが使用する、メモリ内に記憶されたオーディオ信号パターンおよび頭部筋肉活動パターンを精錬することができる。このようにして、コンピューティングデバイスは、ユーザの音声を認識するために学習し、それによって、発話されたコマンドを正確に認識するその能力を改善する。

【0041】

ボイス活動分析器320、雑音パターン分析器350、増幅器フィルタ340、筋肉運動分析器360、および判定ルーチン370の動作は、デジタル信号プロセッサ(DSP)など、単一のプロセッサ341内で、または複数のプロセッサ内で実装され得ることに留意されたい。明快および簡潔のために、これらの実施形態の方法は、ボイス認識プロセッサ380を計算集約型ボイス認識プロセスの実行からゲート制御する、単一のプロセッサ341によって実行されているとして説明される。しかしながら、特許請求の範囲は、特に明記されていない限り、そのようなハードウェア実装に限定されることを意図しない。

【0042】

図4は、オーディオ信号のボイス認識処理をアクティブ化する目的で、ユーザの音声を検出する精度を改善するために、2つの信号セットの相関をどのように使用できるかを示すために、マイクロフォン310から受信したオーディオ信号と筋肉運動検出器330から受信した頭部筋肉活動信号(たとえば、EMG信号)とを同時に示す。上部のグラフは、マイクロフォン310によって検出されたアナログオーディオ信号を示す。下部のグラフは、この例では、EMGセンサーからである、筋肉運動検出器330によって検出されたアナログ頭部筋肉活動を示す。上部のグラフおよび下部のグラフは同時に実行され、すなわち、これらのグラフは、同じ時間期間に生じる別個の入力を表し、したがって、オーディオ信号とEMG信号とが経時的にどのように相関され得るかを示す。

【0043】

オーディオ信号およびEMG信号が同時に受信されるとき、プロセッサ341によって実装される判定ルーチン370は、図3に示したように、両方の信号を使用して、ボイス認識プロセッサ380内でボイス認識処理をいつアクティブ化するかを決定することができる。たとえば、ユーザが発話していることに一致するオーディオ信号またはEMG信号のどちらも存在しないとき、プロセッサ341は、装着型コンピューティングデバイスのユーザは発話していないと決定することができ、ボイス認識処理、またはボイス認識プロセッサ380を非アクティブ化状態に残すことができる。しかしながら、ユーザが発話していることに一致するオーディオ信号と筋肉運動信号の両方が検出されたとき、プロセッサ341は、そのユーザが発話していると決定し、ボイス認識処理および/またはボイス認識プロセッサをアクティブ化することができる。このプロセスは、図4に例示し、下で説明する、4つの時間期

間 $T_1 \sim T_4$ の形で示される。

【 0 0 4 4 】

第1の例では、第1の時間期間 T_1 の間、マイクロフォン310は、音を受信して、オーディオ信号区分 A_1 を生み出す。同じ時間期間の間、筋肉運動検出器330は、頭部筋肉活動信号区分 M_1 内の信号の欠如によって表されるように、ユーザの頭部筋肉活動を検出しない。この状況は、装着型コンピューティングデバイスのマイクロフォン310が、ユーザから発生しない、周囲環境からの音(たとえば、装着型コンピューティングデバイスを装着しているユーザの近傍の人物からのボイス)を受信するときに生じ得る。したがって、マイクロフォン310は、他の人物からのボイスを拾い上げるが、そのユーザは話していないため、EMGセンサーから、ユーザが話していることに対応する頭部筋肉活動を拾い上げない。この情報を
10 用いて、プロセッサ341は、時間期間 T_1 内で受信したオーディオ信号区分 A_1 内の音は、ユーザが発話していることに対応しないという結論を下し、非アクティブ化するように(または、非アクティブ化状態に留まるように)ボイス認識プロセッサ380に命令し、それにより、オーディオ信号区分 A_1 内で音を不要に処理することを回避することができる。

【 0 0 4 5 】

第2の例は、ユーザが食事をしているとき、またはガムを噛んでいるときなど、時間期間 T_2 の間に、ユーザが頭部または顔面の筋肉を運動させているが、発話していないときに受信され得る信号を示す。この状況では、装着型コンピューティングデバイスの筋肉運動検出器330は、頭部筋肉活動信号区分 M_2 内で、音声に対応しない頭部筋肉活動をユーザから検出する。したがって、筋肉運動検出器330は、ユーザから頭部筋肉活動を拾い上げる
20 が、ユーザは話しておらず、マイクロフォン310の近傍には何の音も存在しないため、マイクロフォン310からの背景音だけである。この情報を
用いて、時間期間 T_2 において対応するオーディオ信号区分 A_2 を伴わない頭部筋肉活動区分 M_2 だけが存在するため、判定ルーチン370を実行するプロセッサ341は、そのユーザは発話していないという結論を下すことができる。応答して、プロセッサ341は、非アクティブ化するように(または、非アクティブ化状態に留まるように)ボイス認識プロセッサ380に命令し、それにより、オーディオ信号区分 A_2 内で音を不要に処理することを回避することができる。

【 0 0 4 6 】

第3の例は、第3の時間期間 T_3 の間、マイクロフォン310は、オーディオ信号区分 A_3 内に反映された大きな音を受信するが、ユーザは発話しておらず、したがって、筋肉運動検出器330は区分 M_3 内でいずれの頭部筋肉活動もユーザから検出しないという点で第1の例と同様である。したがって、第1の例と同様、判定ルーチン370を実行するプロセッサ341は、何の対応する頭部筋肉活動信号も存在しないため、時間期間 T_3 内のオーディオ信号区分 A_3 はユーザが発話していることに対応しないと決定する。応答して、プロセッサ341は、非
30 アクティブ化するように(または、非アクティブ化状態に留まるように)ボイス認識プロセッサ380に命令し、それにより、オーディオ信号区分 A_3 内で音を不要に処理することを回避することができる。

【 0 0 4 7 】

第4の例では、時間期間 T_4 の間、マイクロフォン310は、オーディオ信号区分 A_4 を生み出す音を受信し、筋肉運動検出器330は、頭部筋肉活動を検出し、頭部筋肉活動信号区分 M_4
40 を生み出す。EMGセンサー330は、言葉を形成するためにユーザの顎が動く頭部筋肉活動を検出することができるため、これは、装着型コンピューティングデバイスのマイクロフォン310が発話しているユーザから(たとえば、装着型コンピューティングデバイスにコマンドを与える)音を受信するときに生じ得る。オーディオ区分 A_4 内にかなりのオーディオ信号と、頭部筋肉活動信号セグメント M_4 内のかなりの筋肉活動との両方を受信することに対応して、判定ルーチン370を実行するプロセッサ341は、ユーザが発話している可能性が高いと決定し、ボイス認識プロセス380内のボイス認識処理をアクティブ化することができる。

【 0 0 4 8 】

図5は、マイクロフォンと頭部筋肉活動センサーの両方からの信号に基づいて、装着型

10

20

30

40

50

コンピューティングデバイスなど、コンピューティングデバイスの音声認識プロセスを制御するための一実施形態の方法を例示するプロセスフロー図を示す。この実施形態の方法500は、装着型コンピューティングデバイスの1つまたは複数のプロセッサによって実装され得る。簡潔および明瞭のために、方法500は、ボイス認識プロセッサから分離された単一のプロセッサによって実装されるとして説明される。しかしながら、実施形態は、様々なプロセッサアーキテクチャによって実装可能であり、したがって、特に明記されていない限り、特許請求の範囲は、そのような実施形態に限定されることを意図しない。

【0049】

ブロック502で、装着型コンピューティングデバイスのプロセッサは、図1および図3に示したマイクロフォン104a、104b、310など、1つまたは複数のマイクロフォンからオーディオ信号を受信する。

10

【0050】

オプションのブロック504で、無視されてよい、雑音および/または容易に定義される区分を除去するために、受信したオーディオ信号を増幅またはフィルタリングすることができる。たとえば、受信したオーディオ信号が(デシベルレベルまたは他の特性によって)人間の音声と明確に識別可能でないほどゆがんでいる場合、プロセッサは、オーディオ信号のこれらの部分を識別可能な音声である可能性が高い他の部分から分離することができる。

【0051】

同時に、ブロック506で、プロセッサは、図1および図3に示したEMGセンサー102a、102b、330など、1つまたは複数の筋肉運動検出器から頭部筋肉活動信号(たとえば、EMG信号)を受信する。受信した筋肉活動信号または筋肉運動信号は、それらが、ユーザが発話していることに関連する筋肉運動パターンに対応するかどうかを決定するために、プロセッサによって分析され得る、筋肉運動検出器からの未処理の信号であり得る。

20

【0052】

決定ブロック508で、プロセッサは、受信したオーディオ信号が、ユーザが発話していることに関連するオーディオ特性基準を満たすかどうかを決定することができる。この動作で、オーディオ信号のいくつかの特性をしきい値またはパターンと比較して、それらが音声特性基準を満たすかどうかを決定することができる。たとえば、音量の測定値であり得るRMS値を、ユーザが発話していることに一致するしきい値RMS値に対して比較することができる。別の例として、オーディオ信号を、トレーニングルーチンを介して学習され、メモリ内に記憶されたユーザの音声のパターンなど、音声に一致する、記憶された音色パターンおよび側頭筋パターンと比較することができる。プロセッサがオーディオ信号はオーディオ特性基準を満たさないと決定し(すなわち、決定ブロック508=「No」)、オーディオがユーザのボイスからであり得る可能性が低いことを示す場合、プロセッサは、ブロック512で、ボイス認識処理および/またはボイス分析プロセッサを非アクティブ化すること(または、非アクティブ化状態に残すこと)ができる。その後、音がマイクロフォンによって拾い上げられ、頭部筋肉運動活動が筋肉運動検出器によって検出されるにつれて、プロセッサによって方法500の動作を連続的な形で繰り返すことができる。

30

【0053】

プロセッサがオーディオ信号は音声特性基準を満たすと決定し(すなわち、決定ブロック508=「Yes」)、RMS音量レベルおよび/または音色パターンなど、様々な要因に基づいて、ユーザが発話している可能性があることを示す場合、プロセッサは、決定ブロック510で、受信した頭部筋肉活動信号が発話に伴う筋肉運動(speech muscle activity)特性基準を満たすかどうかを決定することができる。決定ブロック508のオーディオ信号の分析と同様に、プロセッサは、受信した頭部筋肉活動信号を、ユーザが発話していることに一致することが知られているRMS信号レベルおよび/または信号パターンデータと比較することができる。一例として、受信した頭部筋肉活動信号を、筋肉運動活動信号が記録された間に、ユーザが定義されたスクリプトを読むトレーニングプロセスの間にメモリ内に記憶された信号と比較することができる。プロセッサがオーディオ信号は筋肉運動基準を満たさ

40

50

ないと決定し(すなわち、決定ブロック510=「No」)、ユーザが発話している可能性が低いことを示す場合、プロセッサは、ブロック512で、ボイス認識処理および/またはボイス分析プロセッサを非アクティブ化すること(もしくは、非アクティブ化状態に残すこと)ができる。その後、音がマイクロフォンによって拾い上げられ、頭部筋肉運動活動が筋肉運動検出器によって検出されるにつれて、プロセッサによって方法500の動作を連続的な形で繰り返すことができる。

【0054】

プロセッサが頭部筋肉活動信号は発話に伴う筋肉運動特性基準を満たすと決定する場合(すなわち、決定ブロック510=「Yes」)、プロセッサは、ユーザが発話している可能性は高いという結論を下し、ブロック514で、ボイス認識アルゴリズムによってオーディオ信号のボイス認識処理をアクティブ化することができる。オプションのブロック516で、ボイス認識アルゴリズムは、ボイス認識プロセスの精度を高めるために、音声認識分析の一環として、筋肉運動検出器から受信した頭部筋肉活動信号を使用することもできる。

【0055】

したがって、方法500に示すように、プロセッサがオーディオ信号の音声認識処理をアクティブ化または可能にする前に、オーディオ信号と頭部筋肉活動信号は両方とも音声特性基準を満たさなければならない。

【0056】

図6は、図5を参照して上で説明した方法500の決定ブロック508の動作を実行するための装着型コンピューティングデバイス100など、コンピューティングデバイス上で実装され得る実施形態の方法508aを例示する。

【0057】

方法508aにおいて、オプションの決定ブロック608で、プロセッサは、オプションで、受信したオーディオ信号のRMSレベルが、ユーザが発話していることに一致するRMSしきい値レベルを満たすかどうかを決定することができる。装着型コンピューティングデバイスのプロセッサが、受信したオーディオ信号のRMSレベルがRMSしきい値レベルを満たさないと決定する場合(すなわち、決定ブロック608=「No」)、プロセッサは、上で説明した方法500のブロック512で、ボイス認識アルゴリズムによってオーディオ信号のボイス認識処理を非アクティブ化すること(または、非アクティブ化状態に残すこと)ができる。たとえば、プロセッサが受信したオーディオ信号は装着型コンピューティングデバイスにオーディオコマンドを与えるために使用されるユーザ音声よりも高いデシベルレベルを有すると決定する場合、プロセッサは、ユーザは発話していない(すなわち、音は環境からのものである)、またはユーザはボイスコマンドとして認識されるべきであることが意図された形で話していないという結論を下すことができる。その場合、プロセッサは、オーディオ信号のボイス認識処理を停止するか、またはそれを開始しない。

【0058】

プロセッサが受信したオーディオ信号のRMSレベルが音声の特性を示すRMSしきい値レベルを満たすと決定する場合(すなわち、決定ブロック608=「Yes」)、オプションのブロック609で、プロセッサは、オーディオ信号が、ユーザが発話している音色パターン特性に整合するかどうかを決定するためにオプションの動作の対を実行することができる。ブロック610で、プロセッサは、受信したオーディオ信号を、メモリ内に記憶された音声の特性音色パターンと比較することができる。一実施形態では、プロセッサは、受信したオーディオ信号を、一般的な人間の音声に関連する、(特定の方言または地方のアクセントを含む)前に記憶した音色パターンと比較することができる。これらの前に記憶されたパターンは、コンピューティングデバイスをインターネットに接続することによって、時々(自動的にまたは手動で)更新され得る。別の実施形態では、プロセッサは、WWANまたはWLANを使用して、インターネットを介して、受信したオーディオ信号を遠隔サーバから音色パターンと比較することができる。さらなる実施形態では、プロセッサは、受信したオーディオ信号を、ユーザトレーニングセッションの間に取得されたユーザの記憶された音色パターンと比較することができ、デバイスまたは遠隔サーバのメモリ内に記憶することが

できる。代替的に、ユーザの音色パターンは、それらのパターンをデバイス上のメモリ内に記憶され、ユーザが発話しているときを認識するために、オーディオ信号と比較され得る、一意のユーザ発話パターンに統合して、ユーザの発話を聴取することから経時的に学習可能である。

【0059】

決定ブロック611で、プロセッサは、受信したオーディオ信号が、音色パターン整合しきい値を超える程度に、音声の特性音色パターンに整合するかどうかを決定することができる。すなわち、パターン特徴の事前に定義された割合の整合範囲内であるなど、信号がかなりの程度、そのパターンに整合するという条件で、受信したオーディオ信号は、ユーザが発話していることをプロセッサに示すために、音声の特性音色パターンの完全な整合でなくてもよい。たとえば、プロセッサは、受信したオーディオ信号が一定の割合のパターン特徴範囲内で特性音色パターンに整合する(たとえば、86%整合)かどうかを決定することができる。プロセッサが、受信したオーディオ信号が音声の特性音色パターンに整合する程度は音色パターン整合しきい値以上でないと決定する場合(すなわち、決定ブロック611=「No」)、プロセッサは、上で説明した方法500のブロック512で、ボイス認識アルゴリズムによってオーディオ信号のボイス認識処理を非アクティブ化すること(または、非アクティブ化状態に残すこと)ができる。たとえば、しきい値が特性音色パターンの少なくとも80%の整合を要求しているとき、プロセッサが受信したオーディオ信号は音声を示す特性音色パターンに45%だけ整合すると決定する場合、プロセッサは、ユーザは発話していないという結論を下し、したがって、受信したオーディオ信号の音声認識処理を停止するか、またはアクティブ化しない。

【0060】

プロセッサが、受信したオーディオ信号が音声の特性音色パターンに整合する程度が音色パターン整合しきい値以上であると決定する場合(すなわち、決定ブロック611=「Yes」)、図5を参照して上で説明した方法500の決定ブロック510で、プロセッサは、受信した頭部筋肉活動が筋肉運動基準を満たすかどうかを決定することができる。たとえば、しきい値が特性音色パターンの少なくとも80%整合を要求するときに、オーディオ信号が、音声を示す、記憶された音色パターンに90%整合するとき、プロセッサは、受信したオーディオ信号がユーザの音声を含むと決定することができる。

【0061】

一実施形態では、方法508aで、オーディオ信号がオーディオ特性基準を満たすかどうかを決定することは、決定ブロック608におけるオプションの動作を含み得るが、ブロック609におけるオプションの動作は含まない。別の実施形態では、方法508aで、オーディオ信号がオーディオ特性基準を満たすかどうかを決定することは、オプションの決定608における動作を実行せずに、ブロック609において動作の対を含み得る。

【0062】

図7は、図5を参照して上で説明した方法500の決定ブロック510の動作を実行するための装着型コンピューティングデバイス100などのコンピューティングデバイス上で実装され得る実施形態の方法510aを示す。図7に示すように、この決定は、筋肉運動活動信号がしきい値レベルを超えるかどうか(すなわち、活動の振幅または大きさ)、ユーザが発話していることを示す活動のパターンに整合するかどうか、またはしきい値を超えることと、パターンに整合することとの両方であるかどうかを決定することを含み得る。

【0063】

方法500の決定ブロック508で、受信したオーディオ信号がオーディオ特性基準を満たすかどうかを決定した後、その前、またはそれと並行して、オプションの決定ブロック709で、プロセッサは、オプションで、頭部筋肉活動信号が筋肉運動信号強度しきい値を超えるかどうかを決定することができる。プロセッサが受信した頭部筋肉活動信号は発話しているユーザの筋肉運動信号強度しきい値特性を超えないと決定する場合(すなわち、決定ブロック709=「No」)、プロセッサは、上で説明した方法500のブロック512のボイス認識アルゴリズムによって、オーディオ信号および/または頭部筋肉活動信号のボイス認識処

理を非アクティブ化することができる。たとえば、受信した頭部筋肉活動信号(たとえば、EMG信号)が、ユーザが発話しているときに対応する頭部筋肉活動のレベル未満である場合、マイクロフォンによって、かなりの音が感知されている(たとえば、決定ブロック608 = 「Yes」)としても、プロセッサは、ユーザは発話していないという結論を下すことができる。

【0064】

プロセッサが受信した頭部筋肉活動信号は発話しているユーザの筋肉運動信号強度しきい値特性を超えると決定する場合(すなわち、決定ブロック709= 「No」)、オプションのブロック710で、プロセッサは、その信号を、ユーザが発話している可能性が高いときの音声の頭部筋肉活動信号パターン特性と比較することができる。そのような特性パターンは、ユーザが定義された文を読む間に筋肉活動信号を記録し、それらの信号を分析して信号内の特性パターンを識別し、識別された特性パターンをメモリ内に記憶することによって、取得され得る。このようにして、コンピューティングデバイスのプロセッサは、そのプロセッサが、ユーザが、あくび、食事、体操、またはガムを噛むなど、他のことを行っているときに検出された頭部筋肉活動の信号から、ユーザが発話しているときに検出された頭部筋肉活動の信号を区別することを可能にさせるパターンを学習することができる。

【0065】

他の頭部筋肉運動から発話に伴う筋肉運動活動を区別するために、ブロック711で、プロセッサは、受信した頭部筋肉活動信号を、音声を示す、記憶された筋肉運動信号特性パターンと比較することができる。この比較動作は、割合整合、または、受信した信号が特性パターンにどの程度密に対応するかに関する他の基準など、受信した頭部筋肉活動信号が記憶された筋肉運動信号特性パターンと整合する程度を決定し得る。決定ブロック712で、プロセッサは、受信した頭部筋肉活動信号が記憶された筋肉運動信号特性パターンに整合する程度が筋肉運動信号パターン整合しきい値を超えるかどうか、またはさもなければ、整合基準を満たすかどうかを決定することができる。プロセッサが受信した頭部筋肉活動信号が記憶された筋肉運動信号特性パターンに整合する程度が筋肉運動信号パターン整合しきい値を超えない、または別の整合基準を満たさないと決定する場合(すなわち、決定ブロック712= 「No」)、プロセッサは、上で説明した方法500のブロック512で、ボイス認識アルゴリズムによってオーディオ信号のボイス認識処理を非アクティブ化することができる。たとえば、しきい値が筋肉運動信号パターンに対して少なくとも80%の整合を要求するとき、プロセッサが受信した頭部筋肉活動信号は、音声を示す、記憶された特性筋肉運動信号パターンに45%だけ整合すると決定する場合、プロセッサは、ユーザは発話していないという結論を下すことができる。

【0066】

プロセッサが受信した頭部筋肉活動信号が記憶された筋肉運動信号特性パターンに整合する程度が筋肉運動信号パターン整合しきい値を超えるか、またはさもなければ、整合基準を満たすと決定する場合(すなわち、決定ブロック712= 「Yes」)、プロセッサは、上で説明した方法500のブロック514で、ボイス認識アルゴリズムによってオーディオ信号のボイス認識処理をアクティブ化することができる。たとえば、しきい値が筋肉運動信号パターンに対して少なくとも80%の整合を要求するとき、プロセッサが受信した頭部筋肉活動信号は、音声を示す筋肉運動信号パターンに94%整合すると決定できる場合、プロセッサは、ユーザは発話しているという結論を下すことができる。

【0067】

一実施形態では、方法510aで、頭部筋肉活動が筋肉運動基準を満たすかどうかを決定することは、オプションの決定ブロック709におけるオプションの決定を含み得るが、オプションのブロック710におけるオプションの動作は含まない。別の実施形態では、方法510aで、頭部筋肉活動が筋肉運動基準を満たすかどうかを決定することは、頭部筋肉運動活動信号を、オプションのブロック710で、音声の特性を有する、記憶されたパターンと比較するオプションの動作を含み得るが、オプションの決定ブロック709における決定は含まない。

10

20

30

40

50

【0068】

図1を参照して上で説明したように、装着型コンピューティングデバイスプロセッサ110は、1つまたは複数のマイクロフォン104a、104b、筋肉運動検出器102a、102b、および1つまたは複数の入力機構106(たとえば、ボタン)からオーディオ信号を受信するためのプロセッサ実行可能命令で構成され得る。これらのセンサー(すなわち、マイクロフォン、頭部筋肉運動検出器、および入力機構)は、信号および/または表示を受信するための手段として使用され得る。プロセッサ110は、上で説明した様々な実施形態の方法に従って、パターンが整合するかどうかなど、条件/トリガを決定するための手段、または、別個のプロセッサ上で、同じプロセッサ上で、もしくはそれらの何らかの組合せの上で計算集約型動作(たとえば、ボイス認識処理)をウェイクさせるかどうかを決定するための手段として使用され得る。プロセッサ110は、特性パターンデータを記憶するための手段として使用され得る、1つまたは複数の内部メモリ112に結合され得る。内部メモリ112は、揮発性メモリもしくは不揮発性メモリであり得るし、また、セキュアメモリおよび/もしくは暗号化メモリ、または非セキュアメモリおよび/もしくは非暗号化メモリ、あるいはそれらの任意の組合せであり得る。プロセッサ110は、上で説明した様々な実施形態の機能を含む、様々な機能を実行するようにソフトウェア命令(すなわち、アプリケーション)によって構成され得る、任意のプログラマブルマイクロプロセッサ、マイクロコンピュータ、または1つもしくは複数の多重プロセッサチップであり得る。1つまたは複数の機能専用の1つのプロセッサ、および他のアプリケーション/機能の実行専用の1つまたは複数のその他のプロセッサなど、複数のプロセッサ110がコンピューティングデバイス内に含まれ得る。典型的には、ソフトウェアアプリケーションは、それらがアクセスされプロセッサ内にロードされる前に内部メモリ内に記憶され得る。プロセッサ110は、アプリケーションソフトウェア命令を記憶するのに十分な内部メモリを含み得る。本明細書では、メモリへの一般的な言及は、内部メモリ112と、装着型コンピューティングデバイスに差込み可能なリムーバブルメモリ(図示せず)と、プロセッサ内のメモリとを含む、プロセッサ110によってアクセス可能なメモリを指す。

【0069】

1つまたは複数の例示的实施形態では、説明した機能が、ハードウェア、ソフトウェア、ファームウェア、またはそれらの任意の組合せの形で実装されてよい。ソフトウェアで実装される場合、機能は、非一時的コンピュータ可読記憶媒体または非一時的プロセッサ可読記憶媒体上に1つもしくは複数の命令またはコードとして記憶され得る。方法またはアルゴリズムのステップを、非一時的コンピュータ可読記憶媒体または非一時的プロセッサ可読記憶媒体に存在し得る、プロセッサ実行可能ソフトウェアモジュール内で実施することができる。非一時的なコンピュータ可読またはプロセッサ可読の記憶媒体は、コンピュータまたはプロセッサによってアクセスされ得る任意の記憶媒体であり得る。限定ではなく例として、そのような非一時的なコンピュータ可読またはプロセッサ可読の媒体は、RAM、ROM、EEPROM、FLASHメモリ、CD-ROMもしくは他の光ディスクストレージ、磁気ディスクストレージもしくは他の磁気ストレージデバイス、または命令もしくはデータ構造の形態で所望のプログラムコードを記憶するために使用され得、コンピュータによってアクセスされ得る任意の他の媒体を含む場合がある。ディスク(disk)およびディスク(disc)は、本明細書において使用されるときに、コンパクトディスク(disc)(CD)、レーザディスク(disc)、光ディスク(disc)、デジタル多用途ディスク(disc)(DVD)、フレキシブルディスク(disk)およびブルーレイディスク(disc)を含み、ディスク(disk)は通常、データを磁気的に再生し、一方、ディスク(disc)は、レーザを用いてデータを光学的に再生する。上記の組合せは、また、非一時的コンピュータ可読媒体および非一時的プロセッサ可読媒体の範囲内に含まれる。加えて、方法またはアルゴリズムの動作は、コンピュータプログラム製品に組み込まれる場合がある、非一時的なプロセッサ可読媒体および/またはコンピュータ可読媒体上のコードならびに/あるいは命令の1つもしくは任意の組合せまたはセットとして存在する場合がある。

10

20

30

40

50

【 0 0 7 0 】

前述の方法の説明とプロセスフロー図とは、単に例示的な例として提供され、様々な実施形態のブロックが提示される順序で実行されなければならないことを要求しまたは暗示することは意図されていない。当業者が了解するように、前述の実施形態のブロックの順序を、任意の順序で実行することができる。

【 0 0 7 1 】

「その後」、「次いで」、「次に」などの用語は、ブロックの順序を限定するものではなく、これらの用語は、単に、方法の説明を通じて読者を導くために使用される。さらに、たとえば、冠詞「a」、「an」または「the」を使用する単数形での請求要素へのいかなる言及も、その要素を単数形に限定するものとして解釈されるべきではない。加えて、本明細書、特に請求項で使用する場合、「含む(comprising)」は、1つもしくは複数の追加の不特定の要素、ステップ、および態様がさらに含まれ、かつ/または存在し得るように、自由な意味を有する。

10

【 0 0 7 2 】

諸実施形態に関連して説明された様々な例示的な論理ブロック、モジュール、回路、およびプロセスフロー図ブロックを、電子ハードウェア、コンピュータソフトウェア、またはその両方の組合せとして実施することができる。ハードウェアおよびソフトウェアのこの交換可能性を明瞭に示すために、様々な例示的な構成要素、ブロック、モジュール、回路、およびブロックを、上では一般にその機能に関して説明した。そのような機能がハードウェアまたはソフトウェアのどちらとして実施されるのかは、特定の応用例とシステム全体に課せられる設計制約とに依存する。当業者は、説明された機能を、各特定の応用例のために様々な形で実施することができるが、そのような実装決定が、本発明の範囲からの逸脱を引き起こすと解釈されてはならない。

20

【 0 0 7 3 】

開示した実施形態の前述の説明は、任意の当業者が本発明を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。これらの実施形態に対する様々な修正は当業者には容易に明らかとなり、本明細書で定義した一般的な原理は、本発明の要旨または範囲から逸脱することなく他の実施形態に適用され得る。したがって、本発明は、本明細書に示す実施形態に限定されることを意図しておらず、以下の特許請求の範囲、ならびに本明細書で開示する原理および新規の特徴と一致する最も広い範囲を与えられるべきである。

30

【 符号の説明 】

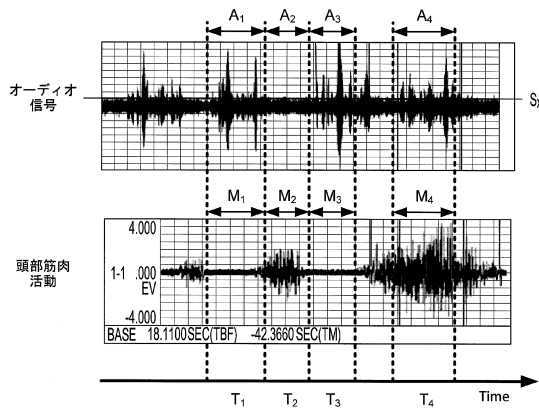
【 0 0 7 4 】

- 100 装着型コンピューティングデバイス
- 102a 筋肉運動検出器
- 102b 筋肉運動検出器
- 103 カメラ
- 104 マイクロフォン
- 104a マイクロフォン
- 104b マイクロフォン
- 106 入力機構
- 108 ワイヤレストランシーバ、トランシーバ
- 110 プロセッサ、コア、またはコプロセッサ
- 112 メモリ、内部メモリ
- 201 ユーザ
- 300 装着型コンピューティングデバイス
- 310 マイクロフォン
- 315 アナログ/デジタル(A/D)変換器
- 320 ボイス活動分析器
- 330 筋肉運動検出器
- 335 アナログ/デジタル(A/D)変換器

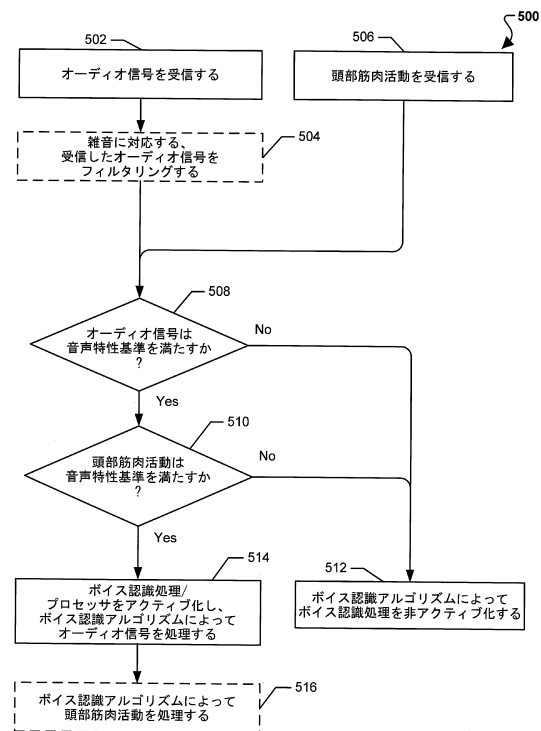
40

50

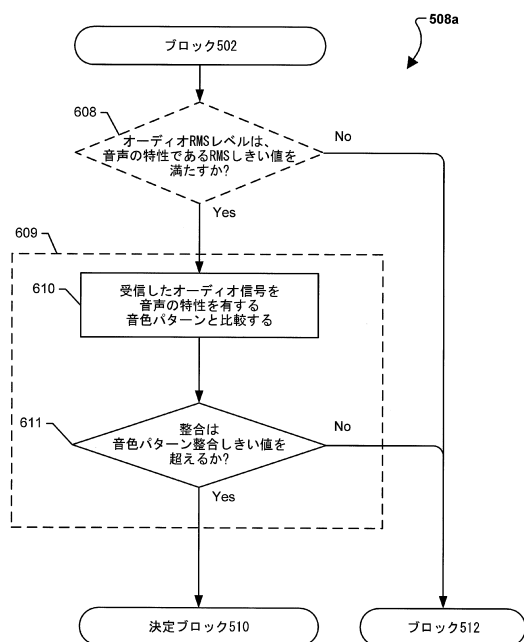
【図4】



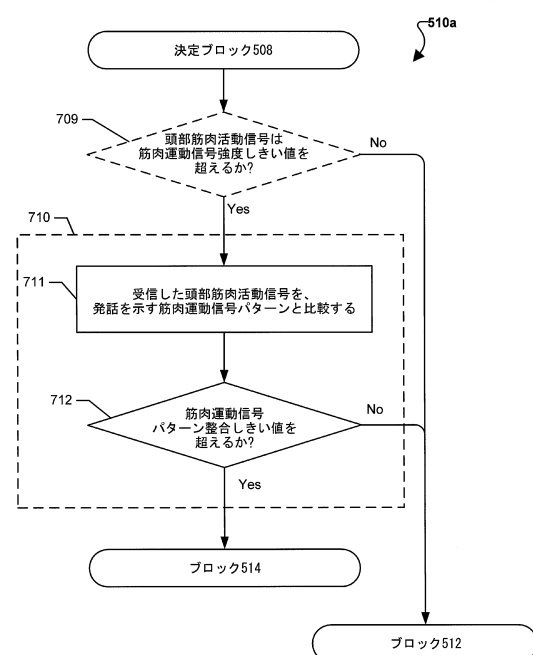
【図5】



【図6】



【図7】



フロントページの続き

- (72)発明者 シェン・シュ・チェン
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ
ヴ・5775
- (72)発明者 アラヴィンド・サンカラン
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ
ヴ・5775
- (72)発明者 パラシュラム・カダディ
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ
ヴ・5775

審査官 菊池 智紀

- (56)参考文献 特開2013-025605(JP,A)
特開2006-208751(JP,A)
特開2006-171226(JP,A)
特開2007-156076(JP,A)

- (58)調査した分野(Int.Cl., DB名)
G10L 15/00-15/34,
25/00-25/93