

(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2021-56485
(P2021-56485A)

(43) 公開日 令和3年4月8日(2021.4.8)

(51) Int.Cl.	F I	テーマコード (参考)
G 1 0 L 15/22 (2006.01)	G 1 0 L 15/22 4 5 3	5 E 5 5 5
G 1 0 L 15/10 (2006.01)	G 1 0 L 15/10 2 0 0 W	
G 0 6 F 3/16 (2006.01)	G 1 0 L 15/22 2 0 0 Z	
G 0 6 F 3/01 (2006.01)	G 0 6 F 3/16 6 3 0	
	G 0 6 F 3/16 6 5 0	

審査請求 有 請求項の数 12 O L (全 17 頁) 最終頁に続く

(21) 出願番号 特願2019-238252 (P2019-238252)
 (22) 出願日 令和1年12月27日 (2019.12.27)
 (31) 優先権主張番号 201910933027.9
 (32) 優先日 令和1年9月29日 (2019.9.29)
 (33) 優先権主張国・地域又は機関
 中国 (CN)

(71) 出願人 513224353
 バイドゥ オンライン ネットワーク テクノロジー (ベイジン) カンパニー リミテッド
 中華人民共和国、100085 ベイジン ハイディエン ディストリクト、シャンディ 10ティーエイチ ストリート、バイドゥ キャンパス、ナンバー 10、3 /フロア
 (74) 代理人 100118913
 弁理士 上田 邦生
 (74) 代理人 100142789
 弁理士 柳 順一郎
 (74) 代理人 100163050
 弁理士 小栗 真由美

最終頁に続く

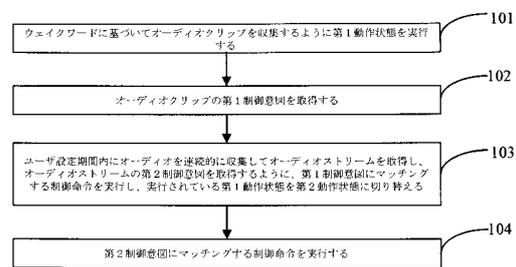
(54) 【発明の名称】 音声制御方法、音声制御装置、電子デバイス、および読み取り可能な記憶媒体

(57) 【要約】 (修正有)

【課題】ユーザは、ウェイクワードを連続的に入力する必要がなく、ユーザ設定期間内にオーディオストリームを連続的に入力するだけで、電子デバイスとの連続的なインタラクションを実現することができ、ユーザの操作を簡素化し、異なるタイプのユーザのニーズを満たし、ユーザ体験を改善する。

【解決手段】ウェイクワードに基づいてオーディオクリップを収集するように第1動作状態を実行し、オーディオクリップの第1制御意図を取得し、次にユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの第2制御意図を取得するように、第1制御意図にマッチングする制御命令を実行し、実行されている前記第1動作状態を第2動作状態に切り替え、次に前記第2制御意図にマッチングする制御命令を実行する。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

ウェイクワードに基づいてオーディオクリップを収集するように第 1 動作状態を実行するステップと、

前記オーディオクリップの第 1 制御意図を取得するステップと、

ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、前記オーディオストリームの第 2 制御意図を取得するように、前記第 1 制御意図にマッチングする制御命令を実行し、実行されている前記第 1 動作状態を第 2 動作状態に切り替えるステップと、

前記第 2 制御意図にマッチングする制御命令を実行するステップと、を含むことを特徴とする、音声制御方法。

10

【請求項 2】

ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、前記オーディオストリームの第 2 制御意図を取得するステップは、

前記第 2 動作状態の構成情報を読み出し、ユーザの操作に応答して設定された前記ユーザ設定期間を取得するステップと、

前記ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、前記オーディオストリームの第 2 制御意図を取得するステップと、

前記ユーザ設定期間内に前記第 2 制御意図が取得されなかった場合、前記第 2 動作状態を前記第 1 動作状態に切り替えるステップと、を含むことを特徴とする、請求項 1 に記載の音声制御方法。

20

【請求項 3】

前記オーディオストリームの第 2 制御意図を取得するステップは、

前記オーディオストリームの音声認識によって得られた情報ストリームを取得するステップと、

前記情報ストリームから各候補意図を取得するステップと、

各候補意図から、現在のシーンにマッチングする第 2 制御意図をスクリーニングするステップと、を含むことを特徴とする、請求項 2 に記載の音声制御方法。

【請求項 4】

前記情報ストリームから各候補意図を取得するステップの後に、

現在のシーンにマッチングしない候補意図への応答を拒否するステップをさらに含むことを特徴とする、請求項 3 に記載の音声制御方法。

30

【請求項 5】

前記第 1 動作状態を第 2 動作状態に切り替えるステップの前に、

前記第 1 制御意図が現在のシーンにマッチングすると決定するステップをさらに含むことを特徴とする、請求項 1 から 4 のいずれかに記載の音声制御方法。

【請求項 6】

ウェイクワードに基づいてオーディオクリップを収集するように第 1 動作状態を実行するように構成される実行モジュールと、

前記オーディオクリップの第 1 制御意図を取得するように構成される取得モジュールと

40

、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、前記オーディオストリームの第 2 制御意図を取得するように、前記第 1 制御意図にマッチングする制御命令を実行し、実行されている前記第 1 動作状態を第 2 動作状態に切り替えるように構成される切り替えモジュールと、

前記第 2 制御意図にマッチングする制御命令を実行するように構成される制御モジュールと、を含むことを特徴とする、音声制御装置。

【請求項 7】

前記切り替えモジュールは、さらに、

前記第 2 動作状態の構成情報を読み出し、ユーザの操作に応答して設定された前記ユー

50

ザ設定期間を取得し、

前記ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、前記オーディオストリームの第2制御意図を取得し、

前記ユーザ設定期間内に前記第2制御意図が取得されなかった場合、前記第2動作状態を前記第1動作状態に切り替えるように構成されることを特徴とする、請求項6に記載の音声制御装置。

【請求項8】

前記切り替えモジュールは、さらに、

前記オーディオストリームの音声認識によって得られた情報ストリームを取得し、

前記情報ストリームから各候補意図を取得し、

各候補意図から、現在のシーンにマッチングする第2制御意図をスクリーニングするように構成されることを特徴とする、請求項7に記載の音声制御装置。

【請求項9】

前記切り替えモジュールは、さらに、

現在のシーンにマッチングしない候補意図への応答を拒否するように構成されることを特徴とする、請求項8に記載の音声制御装置。

【請求項10】

前記装置は、

前記第1制御意図が現在のシーンにマッチングすると決定するように構成される決定モジュールをさらに含むことを特徴とする、請求項6から9のいずれかに記載の音声制御装置。

【請求項11】

少なくとも1つのプロセッサと、

前記少なくとも1つのプロセッサに通信可能に接続されたメモリと、を含む電子デバイスであって、

前記メモリに前記少なくとも1つのプロセッサによって実行可能な命令が記憶され、前記命令は前記少なくとも1つのプロセッサによって実行されることにより、前記少なくとも1つのプロセッサが請求項1～5のいずれかに記載の音声制御方法を実行することを特徴とする、電子デバイス。

【請求項12】

コンピュータ命令が記憶されている非一時的なコンピュータ読み取り可能な記憶媒体であって、

前記コンピュータ命令は、請求項1～5のいずれかに記載の音声制御方法を前記コンピュータに実行させることを特徴とする、非一時的なコンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、音声認識および人工知能の技術分野に関し、特に、音声制御方法、音声制御装置、電子デバイス、および読み取り可能な記憶媒体に関する。

【背景技術】

【0002】

人工知能技術および端末技術の持続的な発展に伴い、人工知能製品、たとえばスマートスピーカなどの電子デバイスが絶えずに普及され、ユーザは、対応する制御命令を実行するように電子デバイスを音声で制御することができる。ユーザが電子デバイスと音声インタラクションを行うとき、電子デバイスの所定のリスリング時間が短すぎるため、ユーザは、電子デバイスとインタラクションできるように、ウェイクワードを繰り返し入力する必要があり、ユーザ体験に影響する。

【発明の概要】

【発明が解決しようとする課題】

10

20

30

40

50

【0003】

本出願は、関連技術における技術的課題の少なくとも1つをある程度解決することを目的とする。

本出願は、ユーザがリスニング時間をカスタマイズすることにより、電子デバイスと音声インタラクションする際に、異なるタイプのユーザのニーズを満たし、ユーザ体験を向上させることができる音声制御方法、音声制御装置、電子デバイス、および読み取り可能な記憶媒体を提供する。

【課題を解決するための手段】

【0004】

本出願の第1態様の実施例は、ウェイクワードに基づいてオーディオクリップを収集するように第1動作状態を実行するステップと、前記オーディオクリップの第1制御意図を取得するステップと、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、前記オーディオストリームの第2制御意図を取得するように、前記第1制御意図にマッチングする制御命令を実行し、実行されている前記第1動作状態を第2動作状態に切り替えるステップと、前記第2制御意図にマッチングする制御命令を実行するステップと、を含む音声制御方法を提供する。

10

【0005】

本出願の実施例の第1の可能な実施例として、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、前記オーディオストリームの音声認識によって得られた第2制御意図を取得するステップは、前記第2動作状態の構成情報を読み出し、ユーザの操作にตอบสนองして設定された前記ユーザ設定期間を取得するステップと、前記ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、前記オーディオストリームの第2制御意図を取得するステップと、前記ユーザ設定期間内に前記第2制御意図が取得されなかった場合、前記第2動作状態を前記第1動作状態に切り替えるステップと、を含む。

20

【0006】

本出願の実施例の第2の可能な実施例として、前記オーディオストリームの第2制御意図を取得するステップは、オーディオストリームの音声認識によって得られた情報ストリームを取得するステップと、前記情報ストリームから各候補意図を取得するステップと、各候補意図から、現在のシーンにマッチングする第2制御意図をスクリーニングするステップと、を含む。

30

【0007】

本出願の実施例の第3の可能な実施例として、前記情報ストリームから各候補意図を取得するステップの後に、現在のシーンにマッチングしない候補意図への応答を拒否するステップをさらに含む。

【0008】

本出願の実施例の第4の可能な実施例として、前記第1動作状態を第2動作状態に切り替えるステップの前に、前記第1制御意図が現在のシーンにマッチングすると決定するステップをさらに含む。

【0009】

本出願の第2態様の実施例は、ウェイクワードに基づいてオーディオクリップを収集するように第1動作状態を実行するように構成される実行モジュールと、前記オーディオクリップの第1制御意図を取得するように構成される取得モジュールと、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、前記オーディオストリームの第2制御意図を取得するように、前記第1制御意図にマッチングする制御命令を実行し、実行されている前記第1動作状態を第2動作状態に切り替えるように構成される切り替えモジュールと、前記第2制御意図にマッチングする制御命令を実行するように構成される制御モジュールと、を含む音声制御装置を提供する。

40

【0010】

本出願の第3態様の実施例は、電子デバイスを提供し、前記電子デバイスは、少なく

50

とも1つのプロセッサと、前記少なくとも1つのプロセッサに通信可能に接続されたメモリとを含み、前記メモリに前記少なくとも1つのプロセッサによって実行可能な命令が記憶され、前記命令は前記少なくとも1つのプロセッサによって実行されることにより、前記少なくとも1つのプロセッサは本出願の第1態様に係る音声制御方法を実行可能である。

【0011】

本出願の第4態様の実施例は、コンピュータ命令の非一時的なコンピュータ読み取り可能な記憶媒体を提供し、前記コンピュータ命令は、本出願の第1態様の実施例に係る音声制御方法をコンピュータに実行させるためのものである。

【0012】

上記出願の実施例は、以下の利点または有益な効果を有する。

ウェイクワードに基づいてオーディオクリップを収集するように第1動作状態を実行し、オーディオクリップの第1制御意図を取得し、かつユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの第2制御意図を取得するように、第1制御意図にマッチングする制御命令を実行し、実行されている第1動作状態を第2動作状態に切り替え、第2制御意図にマッチングする制御命令を実行する。これにより、ユーザは、ウェイクワードを連続的に入力する必要がなく、ユーザ設定期間内にオーディオストリームを連続的に入力するだけで、電子デバイスとの連続的なインタラクションを実現することができ、ユーザの操作を簡素化し、異なるタイプのユーザのニーズを満たし、ユーザ体験を改善することができる。

【0013】

上記の選択可能な実施例が有する他の効果について、具体的な実施例と併せて以下に説明する。

【図面の簡単な説明】

【0014】

図面は、本出願の構成をよりよく理解するためのものであり、本出願を限定するものではない。本出願の上記および/または追加の態様および利点は、図面と併せて実施例の以下の説明により明らかになり、理解されやすくなる。

【図1】本出願の実施例1に係る音声制御方法の概略フローチャートである。

【図2】本出願の実施例2に係る音声制御方法の概略フローチャートである。

【図3】本出願の実施例3に係る音声制御方法の概略フローチャートである。

【図4】本出願の実施例4に係る音声制御装置の概略構成図である。

【図5】本出願の実施例5に係る電子デバイスの概略構成図である。

【発明を実施するための形態】

【0015】

本出願の例示的な実施例について、図面を参照して以下に説明する。理解を容易にするために、その中に本出願の実施例の様々な詳細が含まれ、それらは単なる例示と見なされるべきである。したがって、当業者は、本出願の範囲および精神から逸脱することなく、本明細書に記載の実施例に様々な変更および修正を加えることができることを認識すべきである。また、明確さおよび簡潔さのために、周知の機能および構成の説明は、以下の説明では省略される。

【0016】

本発明の実施例に係る音声制御方法、音声制御装置、電子デバイス、及び読み取り可能な記録媒体について、図面を参照して以下に説明する。

【0017】

図1は、本出願の実施例1に係る音声制御方法の概略フローチャートである。

本発明の実施例において、音声制御方法が音声制御装置で構成されることを例示している。当該音声制御装置は、任意の電子デバイスに適用されることにより、電子デバイスが音声制御機能を実行することができる。

【0018】

10

20

30

40

50

ここで、電子デバイスは、パーソナルコンピュータ（Personal Computer、PCと略称する）、クラウドデバイス、モバイルデバイス、スマートスピーカなどであってもよい。モバイルデバイスは、たとえば、携帯電話、タブレットコンピュータ、携帯情報端末、ウェアラブルデバイス、カーデバイスなどの、さまざまなオペレーティングシステム、タッチスクリーンおよび/またはディスプレイを搭載したハードウェアデバイスであってもよい。

【0019】

図1に示すように、当該音声制御方法は、以下のステップ101～ステップ104を含むことができる。

【0020】

ステップ101において、ウェイクワードに基づいてオーディオクリップを収集するように第1動作状態を実行する。

【0021】

本発明の実施例において、第1動作状態は非リスニング状態であってもよい。当該電子デバイスが当該非リスニング状態にあるとき、ユーザは、ウェイクワードを音声で入力し、ユーザによって入力されたウェイクワードに基づいてオーディオクリップを収集することができる。

【0022】

本出願の実施例では、ウェイクワードは、電子デバイスの内蔵プログラムによって予め設定されてもよく、または、ユーザのパーソナライズニーズを満たすために、ウェイクワードは、自分のニーズに応じてユーザによって設定されてもよく、本出願はこれについて限定しない。たとえば、電子デバイスがスマートスピーカである場合、ウェイクワードは「小度小度（ドちゃんドちゃん）」であってもよい。

【0023】

本出願の実施例では、電子デバイスが第1動作状態にあるとき、電子デバイスは、ユーザがウェイクワードを入力したか否かを検出することができる。ユーザがウェイクワードを入力したと検出された場合、ユーザがウェイクワードに続いて入力したオーディオクリップを収集して、音声認識を行うことができる。

【0024】

たとえば、電子デバイスがスマートスピーカである場合、当該スマートスピーカが当該第1動作状態にあるとき、スマートスピーカのウェイクワード「小度小度」であることを例とする。ユーザが「小度小度、歌Aを再生してください」または「小度小度、歌を聴きたい」という音声を入力したことが検出された場合、電子デバイスは、ウェイクワードに続くオーディオクリップ「歌Aを再生してください」または「歌を聴きたい」を認識することができる。

【0025】

ステップ102において、オーディオクリップの第1制御意図を取得する。

【0026】

本出願の実施例では、制御意図は、電子デバイスの内蔵プログラムによって予め設定されてもよく、または、当該方法の柔軟性および適用可能性を向上させるために、制御意図は、キーワードを介してユーザによって設定されてもよく、これについて限定しない。たとえば、制御意図は、音声ビデオの再生、天気のカエリ、目覚まし時計の設定などであってもよい。

【0027】

なお、後に収集されるユーザによって入力されたオーディオストリームに対応する制御意図と区別するために、電子デバイスが第1動作状態にあるときに取得されたオーディオクリップに対応する制御意図を第1制御意図とする。

【0028】

本出願の実施例では、電子デバイスが第1動作状態にあるとき、ユーザがウェイクワードを入力した後、ユーザによって入力されたウェイクワードに続くオーディオクリップを

10

20

30

40

50

収集して、音声認識を行い、オーディオクリップの第1制御意図を取得する。

【0029】

たとえば、電子デバイスがスマートスピーカであることを例とし、ユーザが音声で「小度小度、明日の午前9時に目覚まし時計を設定してください」または「小度小度、目覚まし時計を設定したい」と入力したことが検出されると、当該スマートスピーカは、ウェイクワードに続く「明日の午前9時に目覚まし時計を設定してください」または「目覚まし時計を設定したい」というオーディオクリップを認識することができる。当該オーディオクリップに対応する第1制御意図は、目覚まし時計を設定することである。

【0030】

ステップ103において、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの第2制御意図を取得するように、第1制御意図にマッチングする制御命令を実行し、実行されている第1動作状態を第2動作状態に切り替える。

【0031】

ここで、ユーザ設定期間は、ユーザの操作に応答して電子デバイスによって設定される任意の時間であってもよく、設定期間は、無限であってもよいし、短いものであってもよく、本明細書では限定されない。たとえば、設定期間は、30秒であってもよいし、3分間などであってもよい。

【0032】

本出願の実施例では、第2動作状態は、リスニング状態であってもよい。電子デバイスが当該リスニング状態にあるとき、ユーザは、ウェイクワードを音声で入力する必要がなく、リアルタイムで音声制御命令を入力して、電子デバイスとインタラクションすることができる。

【0033】

本出願の実施例では、第2制御意図は、電子デバイスの内蔵プログラムによって予め設定されてもよく、または、当該方法の柔軟性および適用性を向上させるために、第2制御意図は、ユーザによって設定されてもよく、本出願はこれについて限定しない。上記の第1制御意図と区別するために、第2動作状態ではオーディオストリームの音声認識によって得られた制御意図を第2制御意図とする。

【0034】

本出願の実施例では、電子デバイスが第1動作状態にあるとき、ユーザがウェイクワードを入力した後に収集されたオーディオクリップに対応する第1制御意図が取得された後、第1制御意図が現在のシーンにマッチングすると決定された場合、第1制御意図にマッチングする制御命令を実行する。ユーザと電子デバイスとのリアルタイムまたは連続的なインタラクションを容易にするために、ウェイクワードを入力する必要はなく、この場合、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの第2制御意図を取得するように、電子デバイスを第1動作状態から第2動作状態に切り替えることができる。

【0035】

たとえば、現在のシーンがゲームシーンであり、電子デバイスが第1動作状態にあるときに取得された第1制御意図は「装備を購入」することであり、当該第1制御意図がゲームシーンにマッチングすると仮定すると、当該第1制御意図に対応する制御命令を実行する。

【0036】

本出願の実施例では、電子デバイスが第2動作状態にあるとき、電子デバイスは、オーディオストリームに対応する第2制御意図を取得するために、オーディオストリームを取得するために、ユーザ設定期間内に連続的に音声を取得してもよい。したがって、ユーザが電子デバイスとのリアルタイムのインタラクションまたは連続的なインタラクションが必要な場合、ユーザ設定期間内に、ユーザは、ウェイクワードを連続的に入力する必要がなく、オーディオストリームを連続的に入力するだけで、電子デバイスとの連続的なイン

10

20

30

40

50

タラクションを実現することができ、ユーザの操作を簡素化し、異なるタイプのユーザのニーズを満たし、ユーザ体験を改善することができる。

【0037】

依然として上記の例を例とし、ユーザ設定期間を30秒にすると、スマートスピーカがリスニング状態にあるとき、ユーザは、第1制御意図にマッチングする制御命令を実行してから30秒内にウェイクワードを入力する必要がなく、オーディオストリームを音声で連続的に入力するだけで、「明日の天気はいかがですか」、「歌を再生しましょう」など、スマートスピーカと連続的にインタラクションして、ユーザによって連続的に入力された音声データに対応する第2制御意図を取得することができる。これにより、ユーザは、電子デバイスがリスニング状態にあるときのリスニング時間を自ら設定することができる、異なるタイプのユーザのニーズを満たすことができる。

10

【0038】

ステップ104において、第2制御意図にマッチングする制御命令を実行する。

本出願の実施例では、電子デバイスが第2動作状態にあるとき、電子デバイスは、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの音声認識を行い、オーディオストリームに対応する第2制御意図を取得した後、第2制御意図にマッチングする制御命令を実行する。

【0039】

なお、「第1」、「第2」の用語は単に説明するためのものであり、比較的な重要性を指示又は暗示するか、又は示された技術的特徴の数を黙示的に指示すると理解してはならない。したがって、「第1」、「第2」が限定された特徴は、少なくとも1つの前記特徴を含むことを明示又は暗示する。

20

【0040】

本発明の実施例に係る音声制御方法は、ウェイクワードに基づいてオーディオクリップを収集するように第1動作状態を実行し、オーディオクリップの第1制御意図を取得し、次にユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの第2制御意図を取得するように、第1制御意図にマッチングする制御命令を実行し、実行されている第1動作状態を第2動作状態に切り替え、次に第2制御意図にマッチングする制御命令を実行する。これにより、ユーザは、ウェイクワードを連続的に入力する必要がなく、ユーザ設定期間内にオーディオストリームを連続的に入力するだけで、電子デバイスとの連続的なインタラクションを実現することができ、ユーザの操作を簡素化し、異なるタイプのユーザのニーズを満たし、ユーザ体験を改善することができる。

30

【0041】

上記の実施例に基づいて、別の可能な場合では、上記のステップ103において、電子デバイスが第2動作状態にあるとき、電子デバイスは、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの音声認識によって得られた第2制御意図を取得する際、ユーザ設定期間内に第2制御意図が取得されなかった場合、第2動作状態を第1動作状態に切り替える。上記のプロセスについて、実施例2と併せて以下に詳細に説明する。

40

【0042】

図2は、本出願の実施例2に係る音声制御方法の概略フローチャートである。

図2に示すように、図1に示す実施例に基づいて、当該音声制御方法は、以下のステップ201～ステップ205をさらに含むことができる。

【0043】

ステップ201において、第2動作状態の構成情報を読み出し、ユーザの操作に応答して設定されたユーザ設定期間を取得する。

本出願の実施例では、電子デバイスが第1制御意図にマッチングする制御命令を実行し、第1状態を第2状態に切り替える場合、第2状態にある電子デバイスの構成情報を読み出して、ユーザ設定期間を取得する。

50

【 0 0 4 4 】

なお、ユーザ設定期間は、ユーザの操作に応答して電子デバイスによって設定される時間であり、ユーザのニーズに応じて設定された任意の時間であってもよく、無限であってもよいし、短いものであってもよく、本明細書では限定されない。

【 0 0 4 5 】

たとえば、第2動作状態がリスニング状態である場合、電子デバイスはリスニング状態にあり、電子デバイスとインタラクションするユーザの習慣は異なる。電子デバイスを常にリスニング状態にしようとするユーザがいれば、電子デバイスのリスニング時間が長過ぎると感じるユーザもいる。したがって、本発明の実施例において、電子デバイスのリスニング時間は、ユーザによって設定することができる。たとえば、異なるタイプのユーザのニーズを満たし、ユーザ体験を向上させるために、リスニング時間を3分間に設定してもよいし、30秒に設定してもよい。

10

【 0 0 4 6 】

ステップ202において、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの第2制御意図を取得する。

【 0 0 4 7 】

本出願の実施例では、電子デバイスが第2動作状態にあるとき、電子デバイスは、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームに対応する第2制御意図を取得することができる。これにより、ユーザが電子デバイスとリアルタイムのインタラクションまたは連続的なインタラクションを実行する必要のある場合、ユーザ設定期間内に、ウェイクワードを連続的に入力する必要はなく、ユーザ設定期間内にオーディオストリームに対応する第2制御意図を連続的に収集するだけで、電子デバイスとの連続的なインタラクションを実現することができ、ユーザの操作を簡素化し、ユーザ使用体験を向上させることができる。

20

【 0 0 4 8 】

ステップ203において、ユーザ設定期間内に第2制御意図が取得されたか否かを判断する。

本出願の実施例では、電子デバイスが第2動作状態にあるとき、音声制御装置は、モニタリング方式で、ユーザ設定期間内に、ユーザが音声データを連続的に入力したか否かをモニタリングすることができる。ユーザ設定期間内に音声データが連続的に収集されてオーディオストリームが取得された場合、ユーザ設定期間内に第2制御意図が取得されたか否かを判断する必要がある。

30

【 0 0 4 9 】

ステップ204において、ユーザ設定期間内に第2制御意図が取得されなかった場合、第2動作状態を第1動作状態に切り替える。

本出願の実施例では、電子デバイスが第2動作状態にあるとき、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの第2制御意図を取得する。ユーザ設定期間内に第2制御意図が取得されなかった場合、電子デバイスを第2動作状態から第1動作状態に切り替える。

40

【 0 0 5 0 】

たとえば、電子デバイスがリスニング状態に入るとき、ユーザ設定期間内にユーザによって入力された音声データが収集されなかった場合、または電子デバイスが認識されたオーディオストリームから第2制御意図を取得できなかった場合、リスニング状態を終了し、非リスニング状態に切り替えることができる。たとえば、ユーザにより設定されたリスニング時間は30秒であり、30秒内に電子デバイスが第2制御意図を得なかった場合、電子デバイスを非リスニング状態に切り替える。この場合、ユーザは、電子デバイスとインタラクションし、ユーザが電子デバイスを制御したい場合は、ユーザは、対応するウェイクワードも入力する必要がある。

【 0 0 5 1 】

本出願の実施例では、ユーザ設定期間内に第2制御意図が取得されなかった場合、電子

50

デバイスを第2動作状態から第1動作状態に切り替える。これにより、ユーザが電子デバイスを制御する意図がない場合、第2動作状態を終了するように電子デバイスを制御することにより、電子デバイスが常にリスニング状態または動作状態にあるという状況を回避し、電子デバイスのエネルギー消費を削減することができる。

【0052】

本出願の実施例の可能な一実現形態として、電子デバイスが第2動作状態にあるとき、電子デバイスの表示インターフェイスにおける第1要素を第2要素に置き換え、第3要素を表示することができる。ここで、第1要素は、電子デバイスが第1動作状態にあることを提示するためのものであり、第2要素は、電子デバイスが第2動作状態にあることを提示するためのものであり、第3要素は、ウェイクワードを提示する、および/または、オーディオまたはビデオを再生するためのものである。

10

【0053】

1つの適用シーンとして、現在のシーンがゲームシーンであることを例とし、電子デバイスが第2動作状態にあるとき、すなわち、電子デバイスがリスニング状態にあるとき、ユーザが電子デバイスの現在の状態情報を知ることを容易にするために、電子デバイスのインターフェイスにおける第1要素を第2要素に置き換えることができる。電子デバイスがユーザ設定期間内に第2制御状態を得ていないため、電子デバイスは第2動作状態を終了する。この場合、ユーザはウェイクワードを再入力する必要がある。したがって、第3要素を表示してウェイクワードを提示する必要がある。

【0054】

ステップ205において、ユーザ設定期間内に第2制御意図が得られた場合、第2制御意図にマッチングする制御命令を実行する。

20

本出願の実施例では、電子デバイスは、ユーザ設定期間内に第2制御意図を得た場合、第2制御意図にマッチングする制御命令を実行する。

【0055】

本発明の実施例に係る音声制御方法は、第2動作状態の構成情報を読み出して、ユーザ設定期間を取得し、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの第2制御意図を取得する。ユーザ設定期間内に第2制御意図が取得されなかった場合、第2動作状態を第1動作状態に切り替え、ユーザ設定期間内に第2制御意図が得られた場合、第2制御意図にマッチングする制御命令を実行する。したがって、ユーザ設定期間内に、ユーザが電子デバイスを制御する意図がない場合、第2動作状態を終了するように電子デバイスを制御することにより、電子デバイスが常にリスニング状態または動作状態にあるという状況を回避し、電子デバイスのエネルギー消費を削減することができる。

30

【0056】

上記の実施例に基づいて、可能な一シーンでは、電子デバイスが第2動作状態にあるとき、ユーザ設定期間内にオーディオストリームの第2制御意図を取得する場合、オーディオストリームの音声認識を行って情報ストリームを取得し、情報ストリームから各候補意図を取得し、各候補意図から現在のシーンにマッチングする第2制御意図をフィルタリングすることができる。これにより、ユーザが電子デバイスを使用する際に、他のシーンに切り替えてユーザの没入感に影響を与える問題が回避される。以上のプロセスについて、実施例3と併せて以下に詳細に説明する。

40

【0057】

図3は、本出願の実施例3に係る音声制御方法の概略フローチャートである。

図3に示すように、ステップ103は、以下のステップ301~ステップ304をさらに含むことができる。

【0058】

ステップ301において、オーディオストリームの音声認識によって得られた情報ストリームを取得する。

本出願の実施例では、電子デバイスが第2動作状態にあるとき、ユーザがユーザ設定期

50

間内に音声データを入力した後、電子デバイスは、オーディオストリームを収集し、当該オーディオストリームの音声認識を行って、対応する情報ストリームを決定することができる。

【0059】

ステップ302において、情報ストリームから各候補意図を取得する。

本出願の実施例では、情報ストリームが取得された後、当該情報ストリームに対してセマンティクス認識を行い、情報ストリームに対応する制御意図を決定することができ、情報ストリームに対応する制御意図から各候補意図をスクリーニングすることができる。

【0060】

ステップ303において、各候補意図から、現在のシーンにマッチングする第2制御意図をスクリーニングする。

本出願の実施例では、情報ストリームから各候補意図が取得された後、各候補意図をスクリーニングして、現在のシーンにマッチングする第2制御意図をスクリーニングして、第2制御意図にマッチングする制御命令を実行する。

【0061】

たとえば、現在のシーンがゲームシーンである場合、情報ストリームから電子デバイスによって取得された各候補意図は、「歌を再生」、「装備を購入」を含み、候補意図をスクリーニングして、ゲームシーンにマッチングする「装備を購入」という第2制御意図を得る。

【0062】

ステップ304において、現在のシーンにマッチングしない候補意図への応答を拒否する。

【0063】

本出願の実施例では、情報ストリームから各候補意図が取得された後、候補意図をスクリーニングし、現在のシーンにマッチングしない候補意図がスクリーニングされた場合、電子デバイスは、現在のシーンにマッチングしない候補意図への応答を拒否する。これにより、現在のシーンでのユーザの没入感が確保される。

【0064】

依然として上記の例を例とし、現在のシーンがゲームシーンである場合、情報ストリームから電子デバイスによって取得された各意図候補は、「歌を再生」、「装備を購入」を含み、意図候補をスクリーニングして、ゲームシーンの制御意図にマッチングしない「歌を再生」という候補意図を得る。この場合、電子デバイスは、「歌を再生」という候補意図への応答を拒否することにより、ゲーム中にユーザが中断されてユーザの没入型ゲーム体験に影響を与えることが回避される。

【0065】

本発明の実施例に係る音声制御方法は、情報ストリームを取得し、情報ストリームから各候補意図を取得し、各候補意図から現在のシーンにマッチングする第2制御意図をスクリーニングし、現在のシーンにマッチングしない候補意図への応答を拒否する。これにより、電子デバイスが第2動作状態にあるとき、ユーザは音声データを入力し続け、現在のシーンにマッチングする制御意図にのみ応答するため、現在のシーンでのユーザの没入感が確保され、ユーザの使用体験が改善される。

【0066】

上記の実施例を実現するために、本出願は、音声制御装置をさらに提供する。図4は、本発明の実施例4に係る音声制御装置の概略構成図である。

【0067】

図4に示すように、当該音声制御装置400は、実行モジュール410、取得モジュール420、切り替えモジュール430および制御モジュール440を含む。

【0068】

実行モジュール410は、ウェイクワードに基づいてオーディオクリップを収集するように第1動作状態を実行するように構成される。

10

20

30

40

50

取得モジュール420は、オーディオクリップの第1制御意図を取得するように構成される。

【0069】

切り替えモジュール430は、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの第2制御意図を取得するように、第1制御意図にマッチングする制御命令を実行し、実行している第1動作状態を第2動作状態に切り替えるように構成される。

制御モジュール440は、第2制御意図にマッチングする制御命令を実行するように構成される。

【0070】

さらに、本出願の実施例の可能な一実施例として、切り替えモジュール430は、第2動作状態の構成情報を読み出し、ユーザの操作に応答して設定されたユーザ設定期間を取得し、ユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの第2制御意図を取得し、ユーザ設定期間内に第2制御意図が取得されなかった場合、第2動作状態を第1動作状態に切り替えるように構成される。

【0071】

本出願の実施例の可能な一実施例として、切り替えモジュール430は、さらに、オーディオストリームの音声認識によって得られた情報ストリームを取得し、情報ストリームから各候補意図を取得し、各候補意図から、現在のシーンにマッチングする第2制御意図をスクリーニングするように構成される。

【0072】

本出願の実施例の可能な一実施例として、切り替えモジュール430は、さらに、現在のシーンにマッチングしない候補意図への応答を拒否するように構成される。

【0073】

本出願の実施例の可能な一実施例として、当該音声制御装置は、第1制御意図が現在のシーンにマッチングすると決定するように構成される決定モジュールをさらに含む。

【0074】

なお、音声制御方法の実施例の前述の説明は、この実施例の音声制御装置にも適用可能であり、ここでは説明を省略する。

【0075】

本発明の実施例に係る音声制御装置は、ウェイクワードに基づいてオーディオクリップを収集するように第1動作状態を実行し、オーディオクリップの第1制御意図を取得し、次にユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの第2制御意図を取得するように、第1制御意図にマッチングする制御命令を実行し、実行している第1動作状態を第2動作状態に切り替え、次に第2制御意図にマッチングする制御命令を実行する。これにより、ユーザは、ウェイクワードを連続的に入力する必要がなく、ユーザ設定期間内にオーディオストリームを連続的に入力するだけで、電子デバイスとの連続的なインタラクションを実現することができ、ユーザの操作を簡素化し、異なるタイプのユーザのニーズを満たし、ユーザ体験を改善することができる。

【0076】

上記の実施例を実現するために、本出願は、電子デバイスをさらに提供し、前記電子デバイスは、少なくとも1つのプロセッサと、少なくとも1つのプロセッサに通信可能に接続されたメモリとを含み、メモリに少なくとも1つのプロセッサによって実行可能な命令が記憶され、命令が少なくとも1つのプロセッサによって実行されることにより、少なくとも1つのプロセッサは、本出願の前述の実施例に係る音声制御方法を実行可能である。

【0077】

上記の実施例を実現するために、本出願は、コンピュータ命令の非一時的なコンピュータ読み取り可能な記憶媒体をさらに提供し、コンピュータ命令は、本出願の前述の実施

10

20

30

40

50

例に係る音声制御方法をコンピュータに実行させるためのものである。

【0078】

本出願の実施例によれば、本出願は、電子デバイスおよび読み取り可能な記憶媒体をさらに提供する。

【0079】

図5に示すように、本発明の一実施例に係る音声制御方法の電子デバイスのブロック図である。電子デバイスは、ラップトップコンピュータ、デスクトップコンピュータ、ワークベンチ、携帯情報端末、サーバ、ブレードサーバ、メインフレームコンピュータ、およびその他の適切なコンピュータなど、さまざまな形式のデジタルコンピュータを表すことを目的としている。電子デバイスは、パーソナルデジタル処理、携帯電話、スマートフォン、ウェアラブルデバイス、その他の同様のコンピューティングデバイスなど、さまざまな形態のモバイルデバイスを表すこともできる。本明細書で示されるコンポーネント、それらの接続および関係、およびそれらの機能は単なる例であり、本明細書で説明および/または必要とされる本出願の実現を限定することを意図するものではない。

10

【0080】

図5に示すように、当該電子デバイスは、1つまたは複数のプロセッサ501、メモリ502、および高速インターフェイスおよび低速インターフェイスを含む様々なコンポーネントを接続するためのインターフェイスを含む。さまざまなコンポーネントは、異なるバスを使用して相互接続され、共通のマザーボードに取り付けられるか、必要に応じて取り付けられてもよい。プロセッサは、外部入力/出力デバイス(たとえば、インターフェイスに結合されたディスプレイデバイス)にGUIのグラフィカル情報を表示するためにメモリに記憶されている命令を含む、電子デバイス内で実行される命令を処理することができる。他の実施例では、必要に応じて、複数のプロセッサおよび/または複数のバスを複数のメモリとともに使用することができる。同様に、複数の電子デバイスを接続し、それぞれの電子デバイスが必要な操作(たとえば、サーバレイ、1グループのブレードサーバ、マルチプロセッサシステム)の一部を提供できます。図5では、1つのプロセッサ501を例に挙げている。

20

【0081】

メモリ502は、本出願によって提供される非一時的なコンピュータ読み取り可能な記憶媒体である。前記メモリは、少なくとも1つのプロセッサによって実行可能な命令を記憶し、その結果として、前記少なくとも1つのプロセッサは、本出願に係る音声制御方法を実行する。本出願の非一時的なコンピュータ読み取り可能な記憶媒体は、本出願によって提供される音声制御方法をコンピュータに実行させるためのコンピュータ命令を記憶する。

30

【0082】

メモリ502は、非一時的なコンピュータ読み取り可能な記憶媒体として、非一時的なソフトウェアプログラム、非一時的なコンピュータ実行可能プログラム、および本出願の実施例の音声制御方法に対応するプログラム命令/モジュール(たとえば、図4に示される実行モジュール410、第1取得モジュール420、切り替えモジュール430、および制御モジュール440)を記憶することができる。プロセッサ501は、メモリ502に記憶されている非一時的なソフトウェアプログラム、命令、およびモジュールを実行することにより、サーバの様々な機能的アプリケーションおよびデータ処理を実行し、すなわち、前述の方法の実施例における音声制御方法を実現する。

40

【0083】

メモリ502は、ストレージプログラム領域およびストレージデータ領域を含むことができ、ストレージプログラム領域は、少なくとも1つの機能に必要なオペレーティングシステムおよびアプリケーションプログラムを記憶することができる。ストレージデータ領域は、電子デバイスの使用に基づいて作成されたデータを記憶することができる。また、メモリ502は、高速ランダムアクセスメモリを含むことができるし、少なくとも1つの磁気ディスク記憶装置、フラッシュメモリ装置、または他の非一時的な固体記憶装置な

50

どの非一時的なメモリも含むことができる。いくつかの実施例では、メモリ502は、プロセッサ501に対して遠隔に配置されたメモリを選択可能に含むことができ、これらの遠隔メモリは、ネットワークを介して電子デバイスに接続され得る。上記のネットワークの例としては、インターネット、イントラネット、ローカルエリアネットワーク、モバイル通信ネットワーク、およびそれらの組み合わせを含むが、これらに限定されない。

【0084】

電子デバイスは、入力装置503および出力装置504をさらに含むことができる。プロセッサ501、メモリ502、入力装置503、および出力装置504は、バスまたは他の方式を介して接続することができ、図5では、バスを介した接続を例に挙げている。

【0085】

入力装置503は、入力された数字または文字情報を受信し、電子デバイスのユーザ設定および機能制御に関するキー信号入力を生成することができ、たとえば、タッチスクリーン、キーパッド、マウス、トラックパネル、タッチパネル、ポインティングスティック、1つまたは複数のマウスボタン、トラックボール、ジョイスティックなどの入力デバイス電子デバイス。出力装置504は、表示装置、補助照明装置（たとえば、LED）、触覚フィードバック装置（たとえば、振動モータ）などを含んでもよい。当該表示装置は、液晶ディスプレイ（LCD）、発光ダイオード（LED）ディスプレイ、およびプラズマディスプレイを含んでもよいが、これらに限定されない。いくつかの実施例では、表示装置はタッチスクリーンであってもよい。

【0086】

本明細書で説明されるシステムおよび技術の様々な実施例は、デジタル電子回路、集積回路システム、特定用途向けASIC（特定用途向け集積回路）、コンピュータハードウェア、ファームウェア、ソフトウェア、および/またはそれらの組み合わせで実現され得る。これらの様々な実施例は、少なくとも1つのプログラマブルプロセッサを含むプログラマブルシステム上で実行可能および/または解釈可能な1つまたは複数のコンピュータプログラムで実装されてもよく、当該プログラマブルプロセッサは、ストレージシステム、少なくとも1つの入力装置、および少なくとも1つの出力装置からデータおよび命令を受信し、データおよび命令を当該ストレージシステム、当該少なくとも1つの入力装置、および当該少なくとも1つの出力装置に送信することができる専用または汎用のプログラマブルプロセッサであってもよい。

【0087】

これらのコンピューティングプログラム（プログラム、ソフトウェア、ソフトウェアアプリケーション、またはコードともいう）は、プログラマブルプロセッサの機械命令を含み、されに高レベルの手続き型および/またはオブジェクト指向プログラミング言語、および/またはアセンブリ/機械言語を用いてこれらのコンピューティングプログラムを実施することができる。本明細書で使用される「機械読み取り可能な媒体」および「コンピュータ読み取り可能な媒体」という用語は、機械命令および/またはデータをプログラマブルプロセッサに提供するための任意のコンピュータプログラム製品、デバイス、および/または装置（たとえば、磁気ディスク、光ディスク、メモリ、プログラマブルロジックデバイス（PLD）を指し、機械読み取り可能な信号としての機械命令を受信する機械読み取り可能な媒体を含む、「機械読み取り可能な信号」という用語は、機械命令および/またはデータをプログラマブルプロセッサに提供するための信号を指す。

【0088】

ユーザとのインタラクションを提供するために、本明細書に記載のシステムおよび技術は、ユーザに情報を表示するための表示装置（たとえば、CRT（陰極線管）またはLCD（液晶ディスプレイ）モニター）と、キーボードおよびポインティング装置（たとえば、マウスまたはトラックボール）とを有するコンピュータで実施されてもよく、ユーザは当該キーボードおよび当該ポインティング装置を介してコンピュータに入力を提供できる。他の種類の装置を使用してユーザとのインタラクションを提供することもできる。たとえば、ユーザに提供されるフィードバックは、任意の形式のセンシングフィードバック（

10

20

30

40

50

たとえば、視覚フィードバック、聴覚フィードバック、または触覚フィードバック)であってもよく、任意の形式(たとえば、音響入力、音声入力、または触覚入力)を用いてユーザからの入力を受信することができる。

【0089】

本明細書に記載のシステムおよび技術の各実施例は、バックエンドコンポーネントを含むコンピューティングシステム(たとえば、データサーバとして)、またはミドルコンポーネントを含むコンピューティングシステム(たとえば、アプリケーションサーバ)、またはフロントエンドコンポーネントを含むコンピューティングシステム(たとえば、グラフィカルユーザインターフェイスまたはWebブラウザを備えたユーザコンピュータ。ユーザは、当該グラフィカルユーザインターフェイスまたは当該Webブラウザを介して、ここで説明されるシステムおよび技術の実施例とインタラクションできる)、またはこのようなバックエンドコンポーネント、ミドルコンポーネント、またはフロントエンドコンポーネントの任意の組み合わせを含むコンピューティングシステムで実施することができる。デジタルデータ通信の任意の形式または媒体(たとえば、通信ネットワーク)によってシステムのコンポーネントを相互接続してもよい。通信ネットワークの例は、ローカルエリアネットワーク(LAN)、ワイドエリアネットワーク(WAN)、およびインターネットを含む。

10

【0090】

コンピュータシステムは、クライアントとサーバとを含むことができる。通常、クライアントとサーバとは互いにリモートであり、通常は通信ネットワークを介してインタラクションする。それぞれのコンピュータで実行され、互いにクライアントとサーバとの関係を持つコンピュータプログラムによってクライアントとサーバとの関係を生成する。

20

【0091】

本出願の実施例の技術的解決策によれば、ウェイクワードに基づいてオーディオクリップを収集するように第1動作状態を実行し、オーディオクリップの第1制御意図を取得し、次にユーザ設定期間内にオーディオを連続的に収集してオーディオストリームを取得し、オーディオストリームの第2制御意図を取得するように、第1制御意図にマッチングする制御命令を実行し、実行されている第1動作状態を第2動作状態に切り替え、次に第2制御意図にマッチングする制御命令を実行する。これにより、ユーザは、ウェイクワードを連続的に入力する必要がなく、ユーザ設定期間内にオーディオストリームを連続的に入力するだけで、電子デバイスとの連続的なインタラクションを実現することができ、ユーザの操作を簡素化し、異なるタイプのユーザのニーズを満たし、ユーザ体験を改善することができる。

30

【0092】

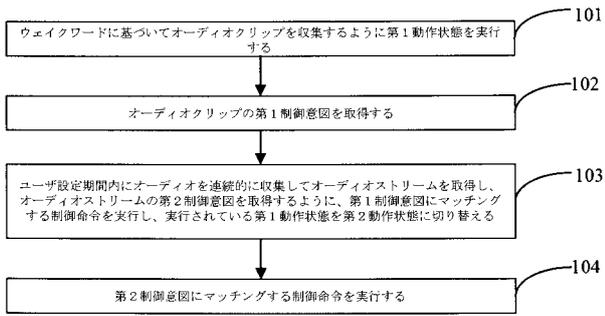
なお、上記のさまざまな形態のプロセスを使用して、ステップを並べ替え、追加、または削除できることを理解されたい。たとえば、本出願で説明されている各ステップは、並列、順次、または異なる順序で実行できる。本出願で開示された技術的解決策の望ましい結果が達成できれば、本明細書は、ここで限定しない。

【0093】

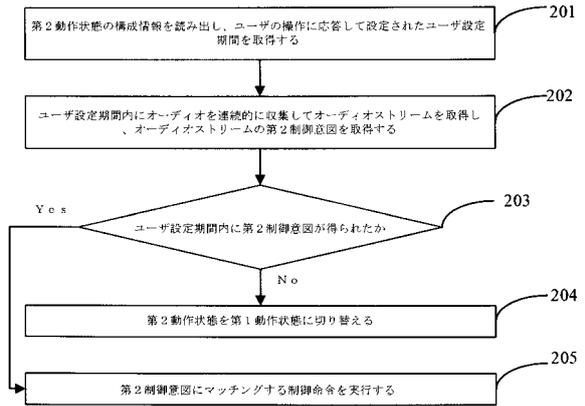
前述した具体的な実施例は、本出願の保護範囲を制限しない。設計要件および他の要因に応じて、さまざまな修正、組み合わせ、サブ組み合わせ、および置き換えを行うことができることを当業者は理解すべきである。本出願の思想および原則の範囲内で行われた修正、同等の置き換え、および改善は、いずれも本出願の保護範囲に含まれるものとする。

40

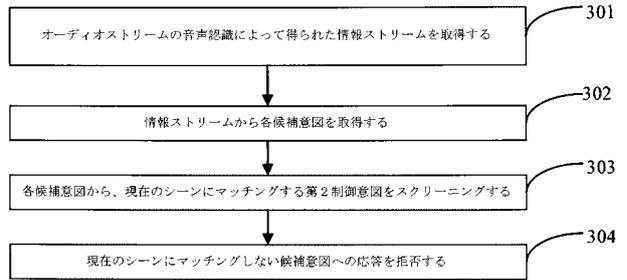
【 図 1 】



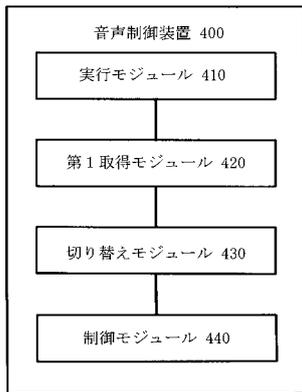
【 図 2 】



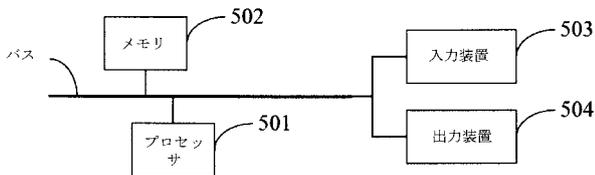
【 図 3 】



【 図 4 】



【 図 5 】



フロントページの続き

(51)Int.Cl.	F I	テーマコード(参考)
	G 0 6 F 3/16 5 2 0	
	G 0 6 F 3/01 5 1 0	

(74)代理人 100201466

弁理士 竹内 邦彦

(72)発明者 ルォ, ヨンシー

中華人民共和国 100085 ベイジン ハイディエン ディストリクト シャンディ 10
イーエイチ ストリート バイドゥ キャンパス ナンバー 10 3/フロア

(72)発明者 ワン, シャシャ

中華人民共和国 100085 ベイジン ハイディエン ディストリクト シャンディ 10
イーエイチ ストリート バイドゥ キャンパス ナンバー 10 3/フロアFターム(参考) 5E555 AA42 AA46 BA04 BB04 BC07 CB56 CB64 CC01 DD07 DD11
EA09 EA23 EA25 FA00