

(19) 日本国特許庁(JP)

## (12) 公表特許公報(A)

(11) 特許出願公表番号

特表2017-535809

(P2017-535809A)

(43) 公表日 平成29年11月30日(2017.11.30)

(51) Int.Cl.	F 1	テーマコード (参考)
<b>G 10 L 15/06</b>	(2013.01)	G 10 L 15/06 300 Y
<b>G 10 L 17/04</b>	(2013.01)	G 10 L 15/06 500 P
<b>G 10 L 17/24</b>	(2013.01)	G 10 L 17/04
<b>G 10 L 15/20</b>	(2006.01)	G 10 L 17/24
		G 10 L 15/20 360 Z

審査請求 未請求 予備審査請求 有 (全 47 頁)

(21) 出願番号	特願2017-521507 (P2017-521507)	(71) 出願人	595020643 クアアルコム・インコーポレイテッド QUALCOMM INCORPORATED
(86) (22) 出願日	平成27年10月2日 (2015.10.2)	(74) 代理人	100108855 弁理士 蔡田 昌俊
(85) 翻訳文提出日	平成29年6月20日 (2017.6.20)	(74) 代理人	100109830 弁理士 福原 淑弘
(86) 國際出願番号	PCT/US2015/053665	(74) 代理人	100158805 弁理士 井関 守三
(87) 國際公開番号	W02016/064556	(74) 代理人	100112807 弁理士 岡田 貴志
(87) 國際公開日	平成28年4月28日 (2016.4.28)		
(31) 優先権主張番号	62/067,322		
(32) 優先日	平成26年10月22日 (2014.10.22)		
(33) 優先権主張国	米国(US)		
(31) 優先権主張番号	14/682,009		
(32) 優先日	平成27年4月8日 (2015.4.8)		
(33) 優先権主張国	米国(US)		

最終頁に続く

(54) 【発明の名称】サウンド検出モデルを生成するためのサウンドサンプル検証

## (57) 【要約】

電子デバイスにおいてサウンド検出モデルを生成する際に使用される少なくとも1つのサウンドサンプルを検証するための方法は、第1のサウンドサンプルを受け取ることと、第1のサウンドサンプルから第1の音響特徴を抽出することと、第2のサウンドサンプルを受け取ることと、第2のサウンドサンプルから第2の音響特徴を抽出することと、第2の音響特徴が第1の音響特徴に類似するかどうかを決定することと、を含む。

【選択図】 図3

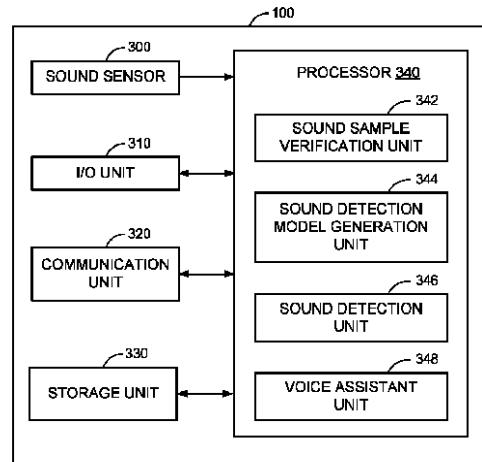


FIG. 3

**【特許請求の範囲】****【請求項 1】**

サウンド検出モデルを生成する際に使用される少なくとも1つのサウンドサンプルを検証するために、電子デバイスにおいて実行される方法であって、前記方法は、

第1のサウンドサンプルを受け取ることと、

前記第1のサウンドサンプルから第1の音響特徴を抽出することと、

第2のサウンドサンプルを受け取ることと、

前記第2のサウンドサンプルから第2の音響特徴を抽出することと、

前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定することとを備える、方法。

10

**【請求項 2】**

前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、前記第1のサウンドサンプルまたは前記第2のサウンドサンプルのうちの少なくとも1つに基づいて、前記サウンド検出モデルを生成することをさらに備える、請求項1に記載の方法。

**【請求項 3】**

前記第1のサウンドサンプルの信号対雑音比(SNR)を決定することと、

前記第2のサウンドサンプルのSNRを決定することと、

前記第1および第2のサウンドサンプルの前記SNRに基づいて、前記第1のサウンドサンプルまたは前記第2のサウンドサンプルのうちの少なくとも1つを選択することと、

20

前記第2の音響特徴が前記第1の音響特徴に類似することを決定すると、前記選択された少なくとも1つのサウンドサンプルに基づいて、前記サウンド検出モデルを生成することと、

選択されていないサウンドサンプルに基づいて、前記生成されたサウンド検出モデルを調整することと

をさらに備える、請求項1に記載の方法。

**【請求項 4】**

前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、前記第1の音響特徴および前記第2の音響特徴に基づいて、組み合わせられた音響特徴を決定することをさらに備える、請求項1に記載の方法。

30

**【請求項 5】**

第3のサウンドサンプルを受け取ることと、

前記第3のサウンドサンプルから第3の音響特徴を抽出することと、

前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定することと

をさらに備える、請求項4に記載の方法。

**【請求項 6】**

前記第1の音響特徴および前記第2の音響特徴の各々は、スペクトル特徴または時間領域特徴のうちの少なくとも1つを含む、請求項1に記載の方法。

40

**【請求項 7】**

前記第1の音響特徴および前記第2の音響特徴の各々は、サブワードのシーケンスを含む、請求項1に記載の方法。

**【請求項 8】**

前記サブワードは、音、音素、トライフォン、または音節のうちの少なくとも1つを含む、請求項7に記載の方法。

**【請求項 9】**

前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルを受け取ることと、

前記新規のサウンドサンプルから新規の音響特徴を抽出することと、

前記新規の音響特徴が前記第1の音響特徴に類似するかどうかを決定することと

50

をさらに備える、請求項 1 に記載の方法。

【請求項 1 0】

前記新規の音響特徴が前記第 1 の音響特徴に類似することを決定することに応答して、前記第 1 のサウンドサンプルまたは前記新規のサウンドサンプルのうちの少なくとも 1 つに基づいて、前記サウンド検出モデルを生成することをさらに備える、請求項 9 に記載の方法。

【請求項 1 1】

前記サウンド検出モデルを生成することは、

前記第 1 のサウンドサンプルまたは前記新規のサウンドサンプルのうちの少なくとも 1 つに基づいて、前記サウンド検出モデルの閾値を決定することと、

10

前記第 2 のサウンドサンプルに基づいて、前記閾値を調整することと  
を備える、請求項 1 0 に記載の方法。

【請求項 1 2】

前記第 2 の音響特徴が前記第 1 の音響特徴に類似するかどうかを決定することに基づいて、前記第 2 のサウンドサンプルが前記第 1 のサウンドサンプルに類似するかどうかの指示を出力することと、

前記サウンド検出モデルを生成する際に使用される前記第 1 のサウンドサンプルまたは前記第 2 のサウンドサンプルのうちの少なくとも 1 つを示す入力を受け取ることと

をさらに備える、請求項 1 に記載の方法。

20

【請求項 1 3】

前記第 1 のサウンドサンプルおよび前記第 2 のサウンドサンプルの各々は、前記電子デバイスをアクティブ化するためのコマンド、あるいは前記電子デバイスにおけるアプリケーションまたは機能を制御するためのコマンドのうちの少なくとも 1 つを示す、請求項 1 に記載の方法。

【請求項 1 4】

音声入力を受け取ることと、

前記サウンド検出モデルに基づいて、前記音声入力からキーワードまたはユーザのうちの少なくとも 1 つを認識することと

をさらに備える、請求項 2 に記載の方法。

30

【請求項 1 5】

サウンド検出モデルを生成する際に使用される少なくとも 1 つのサウンドサンプルを検証するための電子デバイスであって、

第 1 のサウンドサンプルおよび第 2 のサウンドサンプルを受け取るように構成されたサウンドセンサと、

前記第 1 のサウンドサンプルから第 1 の音響特徴を抽出し、前記第 2 のサウンドサンプルから第 2 の音響特徴を抽出し、前記第 2 の音響特徴が前記第 1 の音響特徴に類似するかどうかを決定するように構成されたサウンドサンプル検証ユニットと

を備える、電子デバイス。

【請求項 1 6】

前記第 2 の音響特徴が前記第 1 の音響特徴に類似することを決定することに応答して、前記第 1 のサウンドサンプルまたは前記第 2 のサウンドサンプルのうちの少なくとも 1 つに基づいて、前記サウンド検出モデルを生成するように構成されたサウンド検出モデル生成ユニットをさらに備える、請求項 1 5 に記載の電子デバイス。

40

【請求項 1 7】

前記サウンドサンプル検証ユニットは、前記第 2 の音響特徴が前記第 1 の音響特徴に類似することを決定することに応答して、前記第 1 の音響特徴および前記第 2 の音響特徴に基づいて、組み合わせられた音響特徴を決定するように構成される、請求項 1 5 に記載の電子デバイス。

【請求項 1 8】

前記サウンドセンサは、第 3 のサウンドサンプルを受け取るように構成され、

50

前記サウンドサンプル検証ユニットは、前記第3のサウンドサンプルから第3の音響特徴を抽出し、前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定するように構成される、

請求項17に記載の電子デバイス。

【請求項19】

前記サウンドセンサは、前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルを受け取るように構成され、

前記サウンドサンプル検証ユニットは、前記新規のサウンドサンプルから新規の音響特徴を抽出し、前記新規の音響特徴が前記第1の音響特徴に類似するかどうかを決定するように構成される、

10

請求項15に記載の電子デバイス。

【請求項20】

前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定することに基づいて、前記第2のサウンドサンプルが前記第1のサウンドサンプルに類似するかどうかの指示を出力するように構成された出力ユニットと、

前記サウンド検出モデルを生成する際に使用される前記第1のサウンドサンプルまたは前記第2のサウンドサンプルのうちの少なくとも1つを示す入力を受け取るための入力ユニットと

をさらに備える、請求項15に記載の電子デバイス。

【請求項21】

命令を備える非一時的コンピュータ可読記憶媒体であって、前記命令は、電子デバイスの少なくとも1つのプロセッサに、

第1のサウンドサンプルを受け取り、

前記第1のサウンドサンプルから第1の音響特徴を抽出し、

第2のサウンドサンプルを受け取り、

前記第2のサウンドサンプルから第2の音響特徴を抽出し、

前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定する動作を実行させる、非一時的コンピュータ可読記憶媒体。

20

【請求項22】

前記電子デバイスの前記少なくとも1つのプロセッサに、前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、前記第1のサウンドサンプルまたは前記第2のサウンドサンプルのうちの少なくとも1つに基づいて、サウンド検出モデルを生成する動作を実行させる命令をさらに備える、請求項21に記載の非一時的コンピュータ可読記憶媒体。

30

【請求項23】

前記電子デバイスの前記少なくとも1つのプロセッサに、前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、前記第1の音響特徴および前記第2の音響特徴に基づいて、組み合わせられた音響特徴を決定する動作を実行させる命令をさらに備える、請求項21に記載の非一時的コンピュータ可読記憶媒体。

40

【請求項24】

前記電子デバイスの前記少なくとも1つのプロセッサに、

第3のサウンドサンプルを受け取り、

前記第3のサウンドサンプルから第3の音響特徴を抽出し、

前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定する動作を実行させる命令をさらに備える、請求項23に記載の非一時的コンピュータ可読記憶媒体。

【請求項25】

前記電子デバイスの前記少なくとも1つのプロセッサに、

前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルを受け取り、

50

前記新規のサウンドサンプルから新規の音響特徴を抽出し、  
前記新規の音響特徴が前記第1の音響特徴に類似するかどうかを決定する  
動作を実行させる命令をさらに備える、請求項21に記載の非一時的コンピュータ可読  
記憶媒体。

【請求項26】

サウンド検出モデルを生成する際に使用される少なくとも1つのサウンドサンプルを検  
証するための電子デバイスであって、

第1のサウンドサンプルを受け取るための手段と、

前記第1のサウンドサンプルから第1の音響特徴を抽出するための手段と、

第2のサウンドサンプルを受け取るための手段と、

前記第2のサウンドサンプルから第2の音響特徴を抽出するための手段と、

前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定するための手段と  
を備える、電子デバイス。

10

【請求項27】

前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、  
前記第1のサウンドサンプルまたは前記第2のサウンドサンプルのうちの少なくとも1つ  
に基づいて、前記サウンド検出モデルを生成するための手段をさらに備える、請求項26  
に記載の電子デバイス。

【請求項28】

前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、  
前記第1の音響特徴および前記第2の音響特徴に基づいて、組み合わせられた音響特徴を  
決定するための手段をさらに備える、請求項26に記載の電子デバイス。

20

【請求項29】

第3のサウンドサンプルを受け取るための手段と、

前記第3のサウンドサンプルから第3の音響特徴を抽出するための手段と、

前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定する  
ための手段と

をさらに備える、請求項28に記載の電子デバイス。

【請求項30】

前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、新  
規のサウンドサンプルを受け取るための手段と、

前記新規のサウンドサンプルから新規の音響特徴を抽出するための手段と、

前記新規の音響特徴が前記第1の音響特徴に類似するかどうかを決定するための手段と  
をさらに備える、請求項26に記載の電子デバイス。

30

【発明の詳細な説明】

【関連出願の相互参照】

【0001】

[0001] 本願は、2014年10月22日に提出された「SOUND SAMPLE VERIFICATION  
FOR GENERATING SOUND DETECTION MODEL」と題された米国仮特許出願第62/067,3  
22号、および2015年4月8日に提出された「SOUND SAMPLE VERIFICATION FOR GENE  
RATING SOUND DETECTION MODEL」と題された米国特許出願第14/682,009号から  
の優先権の利益を主張するものであり、その全体の内容が参照により本明細書に組み込ま  
れている。

40

【技術分野】

【0002】

[0002] 本願は、一般に、電子デバイスにおいてサウンドサンプルを検証することに關  
し、より具体的には、電子デバイスにおいてサウンド検出モデルを生成するためにサウン  
ドサンプルを検証することに關する。

【背景技術】

【0003】

50

[0003] 近年、スマートフォン、タブレットコンピュータ、ウェアラブル電子デバイスなどのような電子デバイスが、消費者の間で急速に普及してきている。これらのデバイスは通常、ワイヤレスまたは有線ネットワークを介してボイスおよび／またはデータ通信機能を提供している。加えて、このような電子デバイスは、ユーザからのボイスコマンドを受け取りかつ認識するための音声認識機能 (speech recognition function) を含むことが多い。このような機能は、ユーザからのボイスコマンドが受け取られ、認識されるときに、ボイスコマンド (例えば、キーワード) に関連付けられた機能を電子デバイスが実行することを可能にする。例えば、電子デバイスは、ユーザからのボイスコマンドに応答して、ボイスアシスタントアプリケーションをアクティブ化し、オーディオファイルを再生し、あるいは写真撮影を行い得る。

10

#### 【0004】

[0004] 音声認識特徴を有する電子デバイスにおいて、製造業者またはキャリアは、予め定められたキーワードおよび関連付けられたサウンドモデルをデバイスに備えることが多く、それらは、入力サウンド中のキーワードを検出する際に使用され得る。これらのサウンドモデルは通常、様々なサウンド環境における異なる話者からのキーワードの、非常に多くのサウンドレコーディング (例えば、数千以上の音声サンプル) に基づいて生成される。製造業者またはキャリアによって提供されるこのようなサウンドモデルは、電子デバイスのユーザによって洗練され (refined) 得る。例えば、電子デバイスは、予め定められたキーワードのうちの数個の発話 (utterances) をユーザから受け取り、その発話を使用して、予め定められたキーワードに関連付けられる予め記憶された (pre-stored) サウンドモデルを訓練 (train) し得る。いくつかの電子デバイスはまた、あるキーワードをユーザがボイスコマンドに指定することも可能にし得る。この場合、電子デバイスは、指定されたキーワードのうちの数個の発話をユーザから受け取り、指定されたキーワードについてのサウンドモデルをその発話を生成し得る。

20

#### 【0005】

[0005] 一般に、サウンドモデルの検出性能は、サウンドモデルが生成または訓練された発話の数と品質に関連する。サウンドモデルの検出性能は、発話の数が増えるにつれて改善され得る。しかしながら、キーワードについての所定数の発話に対し、発話の品質は話者間または発話間で異なり得る。例えば、意図されないユーザからの発話が受けられた場合、その発話から生成されたキーワードモデルの検出性能は、意図されたユーザからの入力サウンド中のキーワードを検出する際に低くなり得る。一方、ユーザがキーワードを2つの発話で異なって発音する場合、その発話から生成されたサウンドモデルからのキーワードは、正しく検出されない可能性がある。さらに、いくつかの発話は、騒音のある環境で受け取られる可能性があり、そのため、サウンドモデルを生成するのに十分な品質を提供しない可能性がある。よって、このような発話から生成または訓練されたサウンドモデルは、適切な検出性能を実現しない可能性がある。

30

#### 【発明の概要】

#### 【0006】

[0006] 本開示は、電子デバイスにおいてサウンド検出モデルを生成するためのサウンドサンプルを検証することに関する。

40

#### 【0007】

[0007] 本開示の1つの態様によると、サウンド検出モデルを生成する際に使用される少なくとも1つのサウンドサンプルを検証するための方法が開示される。方法は、電子デバイスにおいて実行され得る。この方法では、第1のサウンドサンプルが受け取られ得る。第1の音響特徴は、第1のサウンドサンプルから抽出され得る。さらに、第2のサウンドサンプルが受け取られ得る。第2の音響特徴は、第2のサウンドサンプルから抽出され得る。第2の音響特徴が第1の音響特徴に類似するかどうかが、決定される。本開示はまた、この方法に関連する装置、デバイス、システム、手段の組み合わせ、およびコンピュータ可読媒体を説明する。

#### 【0008】

50

[0008] 本開示の別の態様によると、サウンド検出モデルを生成する際に使用される少なくとも1つのサウンドサンプルを検証するための電子デバイスが開示される。電子デバイスは、サウンドセンサとサウンドサンプル検証ユニットとを含み得る。サウンドセンサは、第1のサウンドサンプルおよび第2のサウンドサンプルを受け取るように構成され得る。サウンドサンプル検証ユニットは、第1のサウンドサンプルから第1の音響特徴を抽出し、第2のサウンドサンプルから第2の音響特徴を抽出し、第2の音響特徴が第1の音響特徴に類似するかどうかを決定するように構成され得る。

【0009】

[0009] 本開示の発明の態様の実施形態は、添付の図面とともに読むとき、下記の詳細な説明に関連して理解されるだろう。

10

【図面の簡単な説明】

【0010】

【図1】[0010] 図1は、本開示の1つの実施形態に従った、サウンド検出モデルを生成するために1つ以上のサウンドサンプルを検証するように構成された電子デバイスを図示する。

【図2】[0011] 図2は、本開示の別の実施形態に従った、サウンド検出モデルを生成するために1つ以上のサウンドサンプルを検証するように構成された電子デバイスを図示する。

20

【図3】[0012] 図3は、本開示の1つの実施形態に従った、サウンド検出モデルを生成するために1つ以上のサウンドサンプルを検証するように構成された電子デバイスのブロック図を図示する。

【図4】[0013] 図4は、本開示の1つの実施形態に従った、1つ以上のサウンドサンプルが、サウンド検出モデルを生成する際に使用され得るかどうかを決定するように構成された、電子デバイス中のサウンドサンプル検証ユニットのブロック図を図示する。

20

【図5】[0014] 図5は、本開示の1つの実施形態に従った、サウンド検出モデルを生成する際に使用される1つ以上のサウンドサンプルを検証するために、電子デバイスにおいて実行される例示的な方法のフローチャートである。

30

【図6】[0015] 図6は、本開示の1つの実施形態に従った、組み合わせられた音響特徴を決定するために、電子デバイスにおいて実行される例示的な方法のフローチャートである。

【図7】[0016] 図7は、本開示の別の実施形態に従った、サウンド検出モデルを生成する際に使用される1つ以上のサウンドサンプルを検証するために、電子デバイスにおいて実行される例示的な方法のフローチャートである。

30

【図8】[0017] 図8は、本開示の1つの実施形態に従った、一括モード(batch mode)でサウンド検出モデルを生成する際に使用される1つ以上のサウンドサンプルを検証するために、電子デバイスにおいて実行される例示的な方法のフローチャートである。

40

【図9】[0018] 図9は、本開示の1つの実施形態に従った、1つ以上のサウンドサンプルの音響特徴が類似するかどうかを決定するために、電子デバイスにおいて実行される例示的な方法のフローチャートである。

【図10】[0019] 図10は、本開示のいくつかの実施形態に従った、サウンド検出モデルを生成する際に使用される1つ以上のサウンドサンプルを検証するための方法および装置が実装され得る電子デバイスのブロック図を図示する。

40

【図11】[0020] 図11は、いくつかの実施形態に従って実装された、先に説明されたサーバのいずれか1つであり得るサーバシステムを図示するブロック図を図示する。

【詳細な説明】

【0011】

[0021] 様々な実施形態に対して詳細に参照がなされ、その例が添付の図面で図示される。下記の詳細な説明では、本願の主題の完全な理解を提供するために、多くの特定の詳細が記載される。しかしながら、これらの特定の詳細がなくとも本願の主題が実現されることは、当業者にとって明らかだろう。他の事例では、様々な実施形態の態様を余計に

50

曖昧にしないように、周知の方法、プロジェクト、システム、および構成要素は、詳細には説明されていない。

【0012】

[0022] 図1は、本開示の1つの実施形態に従った、サウンド検出モデルを生成するために複数のサウンドサンプルS1、S2、S3、S4、およびS5を検証するように構成された電子デバイス100を図示する。電子デバイス100は、スマートフォン、セルラーフォン、パーソナルコンピュータ、ラップトップコンピュータ、タブレットコンピュータ、スマートテレビジョン、ゲームデバイス、マルチメディアプレーヤなどのような、サウンドキャプチャ能力および処理能力を備えた任意の適切なデバイスであり得る。電子デバイス100は、サウンド検出モデルを生成する際の使用のためのサウンドサンプルS1、S2、S3、S4、およびS5をユーザ120から受け取り得る。

10

【0013】

[0023] 電子デバイス100は、5つのサウンドサンプルS1、S2、S3、S4、およびS5を受け取るように図示されているが、任意の適切な数のサウンドサンプルがサウンド検出モデルを生成するために受け取られおよび処理され、また、サウンドサンプルの数は、ユーザの便宜、好み、性能要求のような多数の要素に基づいて、予め定められ得る。1つの実施形態では、サウンドサンプルS1、S2、S3、S4、およびS5が次々に受け取られ、次に、一括モードで処理され得る。別の実施形態では、サウンドサンプルS1、S2、S3、S4、およびS5は、下記の図2を参照してさらに詳細に説明されるように、1つずつ受け取られ、処理され得る。

20

【0014】

[0024] 本明細書で使用される際に、「サウンド検出モデル」という用語は、電子デバイス100によって受け取られた入力サウンドのサウンド信号またはデータにおいて、キーワードおよび/または特定のユーザを検出する際に使用されるモデルまたはデータベースを指し得、キーワードおよび/またはユーザを示す1つ以上の音響特徴(acoustic features)または特性を含み得る。例えば、サウンド検出モデルは、キーワードに関連付けられるまたはキーワードを示すキーワード検出モデルであり得、ユーザから受け取った入力音声中のキーワードを検出するように適合され得る。「キーワード」という用語は、電子デバイス100中の機能またはアプリケーションをアクティブ化、動作、または制御するために使用され得る、1つ以上のワードの任意のデジタルまたはアナログサウンド表現を指し得る。追加的にまたは代替的に、サウンド検出モデルは、入力サウンドからユーザを認識するように適合されたスピーカ検証モデルであり得る。

30

【0015】

[0025] 1つの実施形態では、音響特徴は、スペクトル特徴、時間領域特徴などを含み得、それらは、キーワードまたは特定のユーザを示し得る。スペクトル特徴は、メル周波数ケプストラム係数(MFCC:mel frequency cepstral coefficients)、ケプストラム差分係数(デルタMFCC)、線スペクトルペア(LSP:line spectral pair)係数などを含み得る。時間領域特徴は、ゼロ交差率(zero crossing rate)、フレームエネルギーなどを含み得る。音響特徴はまた、平均値、中央値、モード、変数、標準偏差、共分散、最大値、最小値、尖度、高次運動量(high order momentum)などのような、スペクトル特徴または時間領域特徴の統計的尺度も含み得、それらは、キーワードまたは特定のユーザを示し得る。別の実施形態では、音響特徴は、キーワードまたは特定のユーザを示し得るサブワード(subwords)のシーケンスを含み得る。本明細書で説明されるような「サブワード」または「サブワードユニット」という用語は、音(phone)、音素(phoneme)、トライフォン(triphone)、音節などのような、基本サウンドユニットを指し得る。追加的にまたは代替的に、音響特徴は、サウンド強度レベル(sound intensity level)、信号対雑音比(SNR)、または残響時間(RT)を含み得、それらは、サウンド品質を示し得る。

40

【0016】

[0026] 電子デバイス100は、1つ以上の関連付けられた機能またはアプリケーショ

50

ンをアクティブ化または実行するための 1 つ以上のキーワードを電子デバイス 100 中に記憶するように構成され得る。キーワードは、1 つ以上の予め定められたキーワード（例えば、製造業者によって割り当てられたキーワード）および / または 1 つ以上のユーザによって定義された（user-defined）キーワードであり得る。予め定められたキーワードについてのキーワード検出モデルは、電子デバイス 100 によって受け取られた入力サウンド中の予め定められたキーワードを検出する際に使用され得る。1 つの実施形態によると、複数の予め定められたキーワードについての複数のキーワード検出モデルは、電子デバイス 100 のサードパーティプロバイダまたは製造業者によって生成され、電子デバイス 100 中に予め記憶される、および / または、外部サーバまたはデバイス（図示されない）からダウンロードされ得る。

10

## 【0017】

[0027] 図示される実施形態では、電子デバイス 100 は、予め定められたキーワード「やあ、スナップドラゴン（Hey Snapdragon）」130 を示すサウンドサンプル S1、S2、S3、S4、および S5 をユーザ 120 から受け取り得る。予め定められたキーワードに関連付けられる予め記憶されたキーワード検出モデルは、ユーザ 120 からのサウンドサンプル S1、S2、S3、S4、および S5 に基づいてキーワード検出モデルを訓練することによって、ユーザ 120 に適合され得る。電子デバイス 100 は次に、入力サウンド中の予め定められたキーワードを検出するために、ユーザ 120 に適合された、訓練されたキーワード検出モデルを使用し得る。

20

## 【0018】

[0028] 追加的なまたは代替的な実施形態では、ユーザ 120 はまた、電子デバイス 100 においてアクティブ化または実行される 1 つ以上の機能またはアプリケーションに関連付けられる、1 つ以上のユーザによって定義されたキーワードを定義し得る。例えば、電子デバイス 100 は、ユーザによって定義されたキーワードを示すサウンドサンプルとして、ユーザ 120 からサウンドサンプル S1、S2、S3、S4、および S5 を受け取り得る。本明細書で説明されるような「ユーザによって定義されたキーワード」という用語は、電子デバイス 100 の機能またはアプリケーションをアクティブ化または実行するために、ユーザ 120 によって定義または指定され得るキーワードを指し得る。受け取ったサウンドサンプル S1、S2、S3、S4、および S5 に基づいて、電子デバイス 100 は、ユーザ 120 に適合された、ユーザによって定義されたキーワードについてのキーワード検出モデルを生成し得る。電子デバイス 100 は次に、入力サウンド中のユーザによって定義されたキーワードを検出するために、ユーザ 120 に適合された、生成されたキーワード検出モデルを使用し得る。

30

## 【0019】

[0029] 追加的にまたは代替的に、サウンド検出モデルは、入力サウンドからユーザ 120 を認識するように適合されたスピーカ検証モデルを含み得る。ユーザ 120 を認識するためのスピーカ検証モデルを生成するために、電子デバイス 100 は、予め定められたキーワードまたはユーザによって定義されたキーワードのようなキーワードを、予め定められ得る特定の回数話すようユーザ 120 に要求し得る。ユーザ 120 が、要求された数のサウンドサンプル（例えば、サウンドサンプル S1、S2、S3、S4、および S5）を提供するとき、電子デバイス 100 は、提供されたサウンドサンプルに基づいて、ユーザ 120 についてのスピーカ検証モデルを生成し得る。電子デバイス 100 は次に、ユーザ 120 に関連付けられているような入力サウンドを認識するために、ユーザ 120 に適合された、生成されたスピーカ検証モデルを使用し得る。

40

## 【0020】

[0030] サウンド検出モデルを生成するためのサウンドサンプル S1、S2、S3、S4、および S5 が一旦受け取られると、電子デバイス 100 は、受け取ったサウンドサンプル S1、S2、S3、S4、および S5 の各々がサウンド検出モデルを生成する際に使用され得るかどうかを決定し得る。この処理では、サウンドサンプル S1、S2、S3、S4、および S5 の各々から 1 つ以上の音響特徴が抽出され得る。1 つの実施形態では、

50

電子デバイス 100 は、関連付けられる抽出された音響特徴を比較することによって、サウンドサンプル S1、S2、S3、S4、および S5 の各ペアの音響特徴が類似するかどうかを決定し得る。少なくとも 2 つのサウンドサンプルの音響特徴が互いに類似すると決定される場合、電子デバイス 100 は、その少なくとも 2 つのサウンドサンプルが類似することを決定し、それらは、サウンド検出モデルを生成する際に使用され得る。一方、あるサウンドサンプルの音響特徴が、少なくとも 2 つの他のサウンドサンプルの音響特徴に類似しないと決定される場合、そのサウンドサンプルは、サウンド検出モデルを生成する際に使用されない可能性がある。

#### 【0021】

[0031] いくつかの実施形態では、サウンドサンプル S1、S2、S3、S4、および S5 からサウンドサンプルのペアが選択され得、選択されたサウンドサンプルの各々から音響特徴が抽出され得る。電子デバイス 100 は次に、選択されたサウンドサンプルの音響特徴が互いに類似するかどうかを決定し得る。例えば、電子デバイス 100 は、サウンドサンプル S1 および S2 を選択し、サウンドサンプル S1 および S2 の音響特徴が互いに類似するかどうかを決定し得る。サウンドサンプル S1 と S2 との音響特徴が類似すると決定される場合、電子デバイス 100 は、サウンドサンプル S1 と S2 とが類似することを決定し、それらは、サウンド検出モデルを生成する際に使用され得る。

10

#### 【0022】

[0032] 1 つの実施形態では、電子デバイス 100 は、サウンドサンプル S1 および S2 の音響特徴に基づいて、残りのサウンドサンプル S3、S4、および S5 の各々がサウンド検出モデルを生成する際に使用され得るかどうかを決定し得る。例えば、サウンドサンプル S1 および S2 の音響特徴は、サウンドサンプル S1 および S2 についての組み合わせられた音響特徴を生成するために組み合わせられ得る。電子デバイス 100 は、サウンドサンプル S3 を選択し得、サウンドサンプル S3 から音響特徴を抽出し得る。サウンドサンプル S1 および S2 の組み合わせられた音響特徴は次に、サウンドサンプル S3 から抽出された音響特徴と比較され得る。組み合わせられた音響特徴とサウンドサンプル S3 の音響特徴とが類似すると決定される場合、電子デバイス 100 は、サウンドサンプル S3 がサウンド検出モデルを生成する際に使用され得ることを決定し得る。電子デバイス 100 は、2 つ以上のサウンドサンプルの組み合わせられた音響特徴に基づいて、サウンドサンプル S4 および S5 がサウンド検出モデルを生成する際に使用され得るかどうかを決定し得る。例えば、サウンドサンプル S4 は、サウンドサンプル S1 および S2 の組み合わせられた音響特徴、またはサウンドサンプル S1、S2、および S3 の組み合わせられた音響特徴に基づいて処理され得る。

20

#### 【0023】

[0033] 本明細書で使用されるような「類似する音響特徴」という用語、またはそのバリエーションに相当する語句は、スペクトル特徴、時間領域特徴、統計的尺度、サブワードなどのようなパラメータまたは特徴値中の指定された許容値または閾値あるいは割合内で、音響特徴が同じかまたは実質的に同じであることを意味し得る。例えば、音響特徴中のサブワードの 2 つのシーケンスのケースについて、電子デバイス 100 は、シーケンス中の同一サブワードの割合が閾値を超える場合に、2 つのシーケンスが類似することを決定し得る。1 つの実施形態では、2 つの音響特徴は、それらが 30 % 未満だけ異なるときに、類似すると決定され得る。

30

#### 【0024】

[0034] 1 つの実施形態では、電子デバイス 100 は、サウンドサンプル S1、S2、S3、S4、および S5 の各々のサウンド強度レベル（例えば、平均のサウンド強度値）を音響特徴として決定し得る。サウンドサンプルのサウンド強度レベルが閾値サウンドレベルよりも低いと決定される場合、それは、そのサウンドサンプルが優れた（good）サウンドサンプルではないことを示し得る。従って、電子デバイス 100 は、サウンド検出モデルを生成する際にそのサウンドサンプルが使用されない可能性があることを決定し得る。

40

50

## 【0025】

[0035] 追加的にまたは代替的に、電子デバイス100は、サウンドサンプルS1、S2、S3、S4、およびS5の各々のSNRを音響特徴として決定し得る。サウンドサンプルのSNRが閾値SNRよりも低いと決定される場合、それは、そのサウンドサンプルが非常に多くのノイズを有することを示し得る。よって、電子デバイス100は、サウンド検出モデルを生成する際にそのサウンドサンプルが使用されない可能性があることを決定し得る。

## 【0026】

[0036] サウンドサンプルS1、S2、S3、S4、およびS5がサウンド検出モデルを生成する際に使用され得るかどうかを決定すると、電子デバイス100は、サウンドサンプルS1、S2、S3、S4、およびS5がサウンド検出モデルを生成するために使用され得るかどうかの指示を出力し得る。図示される実施形態では、予め定められたキーワード「やあ、スナップドラゴン」130に関して、電子デバイス100は、キーワードについての5つのサウンドサンプルS1、S2、S3、S4、およびS5をユーザ120から受け取り得る。サウンドサンプルS1、S2、S3、S4、およびS5の各々について、電子デバイス100は、サウンドサンプルがサウンド検出モデルを生成する際に使用され得るかどうかを決定し得る。例えば、電子デバイス100は、図1で示されるように、サウンドサンプルS1、S2、S4、およびS5がサウンド検出モデルを生成する際に使用され得ることを決定し得る。

10

## 【0027】

[0037] 一方、電子デバイス100は、サウンドサンプルS3がサウンド検出モデルを生成する際に使用されない可能性があることを決定し得る。例えば、サウンドサンプルS3の複数のスペクトル特徴がサウンドサンプルS1、S2、S4、およびS5の各々の対応するスペクトル特徴に類似しないことを決定すると、電子デバイス100は、サウンドサンプルS3に関連付けられたユーザと、サウンドサンプルS1、S2、S4、およびS5に関連付けられたユーザとが異なることを決定し得、そのため、サウンドサンプルS3がサウンド検出モデルを生成する際に使用されない可能性があることを決定し得る。

20

## 【0028】

[0038] 1つの実施形態では、電子デバイス100は、サウンドサンプルS1、S2、S3、S4、およびS5の各々がサウンド検出モデルを生成する際に使用され得るかどうかを示す、サウンドサンプルS1、S2、S3、S4、およびS5のリスト136をディスプレイスクリーン110上に表示し得る。例えば、リスト136は、サンプルS1、S2、S4、およびS5の各々をチェックマークとともにボックス内に表示することによって、サウンドサンプルS1、S2、S4、およびS5が、サウンド検出モデルを生成するために使用され得る優れたサウンドサンプルであることを示し得る。一方、優れたサウンドサンプルではない可能性があるサウンドサンプルS3は、リスト136に丸印で表示され得る。さらに、サウンドサンプルS3が異なるユーザに関連付けられることを示すメッセージ138が、ディスプレイスクリーン110上に表示され得る。追加的にまたは代替的に、サウンドサンプルS3が優れたサウンドサンプルではないためサウンド検出モデルを生成する際に使用されない可能性があることを示す不機嫌な顔(frowning face)132が、サウンドサンプルS3に関する追加のサウンドサンプルを入力するようユーザ120に要求するメッセージ134とともに、ディスプレイスクリーン100上に表示され得る。

30

40

## 【0029】

[0039] いくつかの実施形態では、メッセージ138は、サウンドサンプルがサウンド検出モデルを生成する際に使用されない可能性がある1つ以上の理由の簡単な説明を含み得る。例えば、メッセージ138は、サウンドサンプルの話者特性(speaker characteristics)が他のサウンドサンプルの話者特性とは異なること、サウンドサンプルの認識されたキーワードが他のサウンドサンプルの認識されたキーワードとは異なること、サウンドサンプルのサウンド強度レベルが閾値サウンド強度レベルよりも低いこと、サウンドサ

50

ンブルの S N R が閾値 S N R よりも低いことなどを示す、1つ以上の理由を含み得る。

【0030】

[0040] 電子デバイス 100 はまた、関連付けられた音響特徴に基づいて、サウンドサンブル S1、S2、S3、S4、および S5 の間で最も高い品質のサウンドサンブルを決定するように構成され得る。最も高い品質のサウンドサンブルは、サウンド検出モデルのためサブワードのシーケンスを生成するために使用され得る。この実施形態では、アイコン 140 ( 例えば、クリック可能なバー (clickable bar) ) は、最も高い品質のサウンドサンブルを最も優れたサウンドサンブルとして再生するためのコマンドに関連付けられ、ディスプレイスクリーン 110 上に表示され得る。ユーザ 120 がアイコン 140 を押下するとき、電子デバイス 100 は、最も優れたサウンドサンブル再生し、それは、ユーザ 120 によって1つ以上の追加のサウンドサンブルを入力するために参照され得る。追加的にまたは代替的に、電子デバイス 100 は、サウンドサンブル S1、S2、S3、S4、および S5 のうちの1つを最も優れたサウンドサンブルであると識別するユーザ 120 からの入力を受け取り、識別されたサウンドサンブルに基づいてサウンド検出モデルのためのサブワードのシーケンスを生成し得る。電子デバイス 100 は、最も優れたサウンドサンブルを再生するためのアイコン 140 を表示するように構成されるが、それはまた、ユーザ 120 がリスト 136 中のサウンドサンブル S1 ~ S5 の間の任意のサウンドサンブルを選択し、選択されたサウンドサンブルを再生することを可能にするようにも構成され得る。

10

【0031】

[0041] 1つの実施形態では、電子デバイス 100 は、1つ以上の代替のまたは追加のサウンドサンブルを入力するために、複数のアイコン 142 および 144 を表示し得る。アイコン「前のやり直し (Redo Previous)」142 は、前に受け取ったサウンドサンブルを再度入れることをユーザ 120 に示し得る。ユーザ 120 がアイコン 142 を押下するとき、電子デバイス 100 は、前に受け取ったサウンドサンブルを廃棄し、その廃棄されたサウンドサンブルに対する代替サンブル (replacement sample) として、新規のサウンドサンブルをユーザ 120 から受け取り得る。一方、アイコン「始めからやり直す (Start Over)」144 は、廃棄され得るサウンドサンブル S1 ~ S5 に対して新規のサウンドサンブルを入れることを示し得る。例えば、サウンド検出モデルを生成する際に使用されると決定されるサウンドサンブル S1 ~ S5 がない場合、またはユーザ 120 がサウンドサンブル S1 ~ S5 に対して新規のサウンドサンブルを入れることを希望する場合、ユーザ 120 は、アイコン 144 を押下することによって新規のサウンドサンブルを入力し得る。新規のサウンドサンブルを受け取ると、電子デバイス 100 は、新規のサウンドサンブルがサウンド検出モデルを生成する際に使用され得るかどうかを決定し、新規のサウンドサンブルについての結果をディスプレイスクリーン 110 上に表示し得る。

30

【0032】

[0042] いくつかの実施形態によると、電子デバイス 100 はまた、サウンド検出モデルを生成する際に使用され得る1つ以上のサウンドサンブル S1、S2、S3、S4、および S5 を選択する、ユーザ 120 からの入力を受け取るように構成され得る。例えば、ユーザ 120 は、サウンド検出モデルを生成する際の使用のために、サウンドサンブル S1、S2、および S4 を選択し得る。追加的にまたは代替的に、電子デバイス 100 は、新規のサウンドサンブルと置き換えられる1つ以上のサウンドサンブルを選択する、ユーザ 120 からの入力を受け取り得る。例えば、ユーザ 120 は、新規のサウンドサンブルと置き換えられる、電子デバイス 100 によってサウンド検出モデルを生成する際に使用されると決定されている可能性があるサウンドサンブル S4 を選択し得る。このように、電子デバイス 100 は、ユーザ 120 による選択に基づいて、サウンドサンブル S1、S2、S3、S4、および S5 に対する1つ以上の新規のサウンドサンブルを受け取り、処理し得る。

40

【0033】

[0043] サウンドサンブル S1、S2、S3、S4、および S5 がサウンド検出モデル

50

を生成する際に使用されると決定されるとき、電子デバイス100は、サウンドサンプルS1、S2、S3、S4、またはS5のうちの少なくとも1つに基づいて、サウンド検出モデルを生成し得る。例えば、電子デバイス100は、サウンドサンプルS1、S2、S3、S4、およびS5の音響特徴に基づいて、サウンド検出モデルを生成し得る。この場合、音響特徴は、サウンドサンプルがサウンド検出モデルを生成する際に使用され得るかどうかを決定するために使用されている音響特徴と同じ音響特徴であり得る。追加的にまたは代替的に、電子デバイス100は、サウンド検出モデルを生成するように適合され、または生成するのに適切なサウンドサンプルから音響特徴を抽出し得る。

#### 【0034】

[0044] 1つの実施形態では、電子デバイス100は、除外または廃棄されているか、あるいは低いS/Nのために新規のサウンドサンプルと置換られている可能性がある少なくとも1つのサウンドサンプルを、生成されたサウンド検出モデルを調整または修正する際に、ノイズが含まれたサウンドサンプル(*noise-embedded sound sample*)として使用し得る。例えば、人工的に生成されたノイズサンプルを使用することに加えてまたはその代わりに、除外または廃棄されている可能性があるサウンドサンプルは、生成されたサウンド検出モデルを調整する際に、ノイズが含まれたサウンドサンプルとして使用され得る。1つ以上のこののようなサウンドサンプルを使用することは、実質的に騒音のない環境および騒音のある環境を含む様々なサウンド環境において、サウンド検出モデルの検出性能を改善し得る。

#### 【0035】

[0045] 追加的にまたは代替的に、電子デバイス100は、低いS/N、低いサウンド強度レベルなどのために廃棄されている可能性がある少なくとも1つのサウンドサンプルを、サウンド検出モデルについての閾値(例えば、検出または類似性閾値)を調整するために使用し得る。例えば、キーワードを示す入力サウンドが受け取られるとき、キーワードに関連付けられたサウンド検出モデルを閾値に基づいて識別するために、電子デバイス100中に記憶された複数のサウンド検出モデルがアクセスされ得る。サウンド検出モデルを識別するために、複数のサウンド検出モデル中の入力サウンドとキーワードの各々との間の類似性の度合いが決定され得る。電子デバイス100は次に、非常に高い類似性を有し、その類似性の度合いが閾値以上であるサウンド検出モデルを、キーワードについてのサウンド検出モデルとして識別し得る。1つの実施形態では、サウンド検出モデルについての閾値は、サウンド検出モデルを生成する際に使用されると決定されているサウンドサンプルのうちの少なくとも1つに基づいて決定され得る。サウンド検出モデルについての検出精度を強化するために、サウンド検出モデルについての閾値は、低いS/N、低いサウンド強度レベルなどのために廃棄されている可能性がある、少なくとも1つのサウンドサンプルに基づいて調整され得る。例えば、電子デバイス100は、サウンドサンプルS1、S2、S3、S4、およびS5の話者特性とは異なる話者特性、予め定められた閾値S/Nより低いS/Nなどを有し得る1つ以上のサウンドサンプルを使用して、サウンド検出モデルについての閾値を調整し得る。

#### 【0036】

[0046] 図2は、本開示の別の実施形態に従った、サウンド検出モデルを生成するための1つ以上のサウンドサンプルS1、S2、S3、S4、およびS5を検証するように構成され得る電子デバイス100を図示する。この実施形態では、サウンドサンプルS1、S2、S3、S4、およびS5は、1つずつ順に受け取られ、処理され得る。1つの実施形態では、電子デバイス100は、サウンドサンプルが受け取られるときに各サウンドサンプルから音響特徴を抽出し、サウンドサンプルのサウンド品質が、サウンド検出モデルを生成する際の使用のための閾値品質以上であるかどうかを、サウンド品質を示す音響特徴に基づいて検証し得る。サウンドサンプルのサウンド品質が閾値品質よりも低いことが決定される場合、電子デバイス100は、サウンドサンプルのサウンド品質が閾値品質以上であると決定されるまで、代替のサウンドサンプルとしてユーザ120から新規のサウンドサンプルを受け取り得る。電子デバイス100は次に、サウンドサンプルの音響特徴

10

20

30

40

50

と1つ以上の他のサウンドサンプルとが類似するかどうかに基づいて、サウンドサンプルがサウンド検出モデルを生成する際に使用され得るかどうかを決定することを進め得る。

【0037】

[0047] 電子デバイス100は始めに、ユーザ120からサウンドサンプルS1を受け取り、サウンドサンプルS1から音響特徴を抽出し得る。電子デバイス100は、S NR、サウンド強度レベルなどのようなサウンド品質を示す音響特徴に基づいて、サウンドサンプルS1のサウンド品質が、サウンド検出モデルを生成する際の使用のための閾値品質以上であるかどうかを決定し得る。サウンドサンプルS1のサウンド品質が閾値品質よりも低いことを電子デバイス100が決定する場合、サウンドサンプルS1のサウンド品質がサウンド検出モデルを生成する際の使用のための閾値品質以上であると決定されるまで、サウンドサンプルS1を再度入れるようユーザ120に促すメッセージ134を表示し得る。

10

【0038】

[0048] 電子デバイス100は、サウンドサンプルS1のサウンド品質がサウンド検出モデルを生成する際の使用のための閾値品質以上であることを決定するとき、ユーザ120から次のサウンドサンプルS2を受け取り、サウンドサンプルS2から音響特徴を抽出し得る。電子デバイス100は、サウンド検出モデルを生成する際の使用のために、サウンドサンプルS1とS2との音響特徴が類似するかどうかを決定し、ディスプレイスクリーン110上にその結果を出力し得る。サウンドサンプルS1とS2との音響特徴が類似すると決定される場合、電子デバイス100は、サウンドサンプルS1およびS2の音響特徴に基づいて、組み合わせられた音響特徴を生成し得る。サウンドサンプルS1とS2との音響特徴が類似しないと決定される場合、電子デバイス100は、サウンドサンプルS1とS2との音響特徴が類似すると決定されるまで、サウンドサンプルS1および/またはS2に対する代替のサウンドサンプルとして、新規のサウンドサンプルを入力するようユーザ120に要求し得る。

20

【0039】

[0049] サウンドサンプルS1およびS2の音響特徴が類似することを決定すると、電子デバイス100は、ユーザ120から次のサウンドサンプルS3を受け取り、サウンドサンプルS3から音響特徴を抽出し得る。電子デバイス100は次に、サウンドサンプルS3の音響特徴がサウンドサンプルS1およびS2の組み合わせられた音響特徴に類似するかどうかを決定し得る。追加的にまたは代替的に、電子デバイス100は、サウンドサンプルS1、S2、およびS3の各ペアの音響特徴が類似するかどうかを決定し得る。例えば、電子デバイス100は、サウンドサンプルS1とS3との音響特徴が類似するかどうかを決定する、および/またはサウンドサンプルS2とS3との音響特徴が類似するかどうかを決定し得る。サウンドサンプルS1、S2、およびS3についての結果は、図2で示されるように、ディスプレイスクリーン110上に表示され得る。このように、電子デバイス100はまた、サウンドサンプルS4およびS5を順に受け取ることを進め、サウンドサンプルS4およびS5の各々が、複数の音響特徴または前に受け取ったサウンドサンプルに関連付けられた音響特徴の組み合わせに基づいて、サウンド検出モデルを生成する際に使用され得るかどうかを決定し得る。

30

【0040】

[0050] 図示される実施形態で示されるように、電子デバイス100は、サウンドサンプルS1、S2、およびS3の各々がサウンド検出モデルを生成する際に使用され得るかどうかをディスプレイスクリーン110上に示す、サウンドサンプルS1、S2、およびS3のリスト200を表示し得る。例えば、リスト200は、サンプルS1およびS2の各々をチェックマークとともにボックス内に表示することによって、サウンドサンプルS1およびS2が、サウンド検出モデルを生成する際に使用され得る優れたサウンドサンプルであることを示し得る。一方、優れたサウンドサンプルではない可能性があるサウンドサンプルS3は、リスト200に丸印で表示され得る。さらに、サウンドサンプルS3が異なるユーザに関連付けられることを示すメッセージ138もまた、ディスプレイスクリ

40

50

ーン 110 上に表示され得る。追加的にまたは代替的に、サウンドサンプル S3 が優れたサウンドサンプルではなく、そのためサウンド検出モデルを生成する際に使用されない可能性があることを示す不機嫌な顔 132 は、サウンドサンプル S3 に関する別のサウンドサンプルを入力するようユーザ 120 に促すメッセージ 134 とともに、ディスプレイスクリーン 110 上に表示され得る。さらに、ディスプレイスクリーン 110 上のアイコン 202 は、サウンドサンプル S1、S2、および S3 から選択され得るサウンドサンプルを再生するためのコマンドに関連付けられ得る。サウンドサンプル S3 が選択されたときにユーザ 120 がアイコン 202 を押下するとき、電子デバイス 100 は、サウンドサンプル S3 を再生し得る。

【0041】

10

[0051] 電子デバイス 100 はまた、1 つ以上の代替のまたは追加のサウンドサンプルを入力するために、複数のアイコン 142 および 144 も表示し得る。アイコン「前のやり直し」142 は、前に受け取ったサウンドサンプル（例えば、サウンドサンプル S3）を再度入れることを示し得る。ユーザ 120 がアイコン 142 を押下するとき、電子デバイス 100 は、廃棄されたサウンドサンプルに対する代替サンプルとして、新規のサウンドサンプルを受け取り得る。一方、アイコン「始めからやり直す」144 は、受け取ったサウンドサンプル S1、S2、および S3 に対して、新規のサウンドサンプルを入れることを示し得る。例えば、サウンド検出モデルを生成する際に使用されるサウンドサンプル S1、S2、および S3 がないと決定される場合、またはユーザ 120 がサウンドサンプル S1、S2、および S3 に対して新規のサウンドサンプルを入れることを希望する場合、ユーザ 120 は、アイコン 144 を押下することによって、新規のサウンドサンプルを入力し得る。新規のサウンドサンプル S1、S2、S3、S4、および S5 の各々を順に受け取ると、電子デバイスは、各新規のサウンドサンプルがサウンド検出モデルを生成する際に使用され得るかどうかを決定し、新規のサウンドサンプルについての決定結果をディスプレイスクリーン 110 上に表示し得る。

20

【0042】

[0052] 図示される実施形態では、サウンドサンプル S4 および S5 は、これらがまだ受け取られていないことを示すように、点線の円で図示され得る。電子デバイス 100 は、サウンドサンプル S3 に関して説明されているが、図 1 を参照して上述される方法でサウンド検出モデルを生成するために、サウンドサンプル S1、S2、S4、および S5 の各々もまた処理し、サウンドサンプルおよび前に受け取ったサウンドサンプルについての情報を累積的に表示し得る。

30

【0043】

[0053] 電子デバイス 100 は、次に続くサウンドサンプル S4 および S5 を受け取り、サウンドサンプル S1、S2、S3、S4、および S5 の各々がサウンド検出モデルを生成する際に使用され得ることを決定し得る。サウンドサンプル S1～S5 についての結果は、ディスプレイスクリーン 110 上に累積的に表示され得る。サウンドサンプル S1、S2、S3、S4、および S5 がサウンド検出モデルを生成する際に使用され得ると決定されたとき、電子デバイス 100 は、サウンドサンプル S1、S2、S3、S4、または S5 のうちの少なくとも 1 つに基づいて、サウンド検出モデルを生成し得る。例えば、電子デバイス 100 は、サウンドサンプル S1、S2、S3、S4、および S5 の音響特徴に基づいてサウンド検出モデルを生成し得る。この場合、音響特徴は、サウンドサンプルがサウンド検出モデルを生成する際に使用され得るかどうかを決定するために使用されている音響特徴と同じ音響特徴であり得る。追加的にまたは代替的に、電子デバイス 100 は、サウンド検出モデルを生成するように適合されたまたは生成するのに適切なサウンドサンプルから、音響特徴を抽出し得る。いくつかの実施形態によると、電子デバイス 100 はまた、サウンドサンプルの各々がサウンド検出モデルを生成する際に使用され得るかどうかを示す、ユーザ 120 からの入力を受け取るように構成され得る。この場合、サウンドサンプルがサウンド検出モデルを生成する際に使用されない可能性があると電子デバイス 100 が決定したとしても、ユーザ 120 によって選択された 1 つ以上のサウンド

40

50

サンプルは、サウンド検出モデルを生成する際に使用され得る。

【0044】

[0054] 図3は、本開示の1つの実施形態に従った、サウンド検出モデルを生成するために1つ以上のサウンドサンプルを検証するように構成された電子デバイス100のブロック図を図示する。電子デバイス100は、サウンドセンサ300、入力／出力(I/O)ユニット310、通信ユニット320、記憶ユニット330、およびプロセッサ340を含み得る。I/Oユニット310は、ディスプレイスクリーン110およびスピーカ(図示されない)を含み得る。ディスプレイスクリーン110は、ユーザからのタッチ入力を受け取るように構成されたタッチディスプレイスクリーンであり得る。プロセッサ340は、サウンドサンプル検証ユニット342、サウンド検出モデル生成ユニット344、サウンド検出ユニット346、およびボイスアシスタントユニット348を含み得る。プロセッサ340は、電子デバイスについての命令を実行するまたは動作を行い得る、中央処理ユニット(CPU)、アプリケーションプロセッサ、マイクロプロセッサなどのような任意の適切な処理ユニットを使用して実装され得る。

10

【0045】

[0055] サウンドセンサ300は、ユーザからの入力サウンドストリームを受け取るように構成され得る。本明細書で使用される場合、「サウンドストリーム」という用語は、1つ以上のサウンド信号またはサウンドデータのシーケンスを指し得る。サウンドセンサ300は、電子デバイス100へのサウンド入力を受け取り、キャプチャし、感知し、および／または検出するために使用され得る、1つ以上のマイクロフォンまたは任意の他のタイプのサウンドセンサを含み得る。加えて、サウンドセンサ300は、このような機能を実行するための任意の適切なソフトウェアおよび／またはハードウェアを用い得る。ユーザの受け取られた入力サウンドストリームは、記憶ユニット330中に記憶され得る。サウンドセンサ300は、処理のために、プロセッサ340にユーザの受け取ったサウンドストリームを提供し得る。

20

【0046】

[0056] 1つの実施形態では、サウンドセンサ300は、1つ以上のサウンドサンプルを含む入力サウンドストリームをユーザから受け取り得る。入力サウンドストリームを受け取ると、サウンドセンサ300は、任意の適切な終点検出アルゴリズムを使用して入力サウンドストリーム中の独立した発話またはサウンドサンプルの各々の始点と終点とを検出することによって、入力サウンドストリームからサウンドサンプルの各々を検出し得る。検出されたサウンドサンプルが抽出され、記憶ユニット330中に記憶され得る。抽出されたサウンドサンプルは、処理のために、プロセッサ340に提供され得る。

30

【0047】

[0057] 記憶ユニット330は、サウンドセンサ300、I/Oユニット340、通信ユニット320、およびプロセッサ340を動作するためのデータおよび命令を記憶するように構成され得る。記憶ユニット330はまた、サウンドセンサ300によって受け取られた入力サウンドストリーム、または入力サウンドストリームから抽出された1つ以上のサウンドサンプルを記憶し得る。記憶ユニット330は、ランダムアクセスメモリ(RAM)、読み取り専用メモリ(ROM)、電気的消去可能プログラマブル読取専用メモリ(EEPROM(登録商標))、フラッシュメモリ、またはソリッドステートドライブ(SSD)のような、任意の適切な記憶装置またはメモリデバイスを使用して実装され得る。

40

【0048】

[0058] 記憶ユニット330はまた、記憶ユニット330中に予め記憶され得る、および／または通信ユニット320を介して外部サーバまたはデバイス(図示されない)からダウンロードされ得る少なくとも1つのサブワードモデルを記憶し得る。いくつかの実施形態では、サブワードモデルは、サブワードモデルによって表わされるサブワードのタイプに従った、限定はされないが、音ベースモデル、音素ベースモデル、トライフォンベースモデル、音節ベースモデルなどであり得、サブワードユニットのリストおよび各サブワ

50

ードユニットについての 1 つ以上の音響特徴を含み得る。

【 0 0 4 9 】

[0059] 記憶ユニット 330 はまた、入力サウンド中のサウンド信号またはデータにおいて、キーワードおよび / または特定のユーザを検出する際に使用される 1 つ以上のサウンド検出モデルを記憶し得る。例えば、サウンド検出モデルは、入力サウンドからユーザを認識するように適合されたスピーカ検証モデルを含み得る。追加的にまたは代替的に、サウンド検出モデルは、予め定められたキーワードを検出するための 1 つ以上のキーワード検出モデル、および / またはユーザによって定義されたキーワードを検出するための 1 つ以上のキーワード検出モデルを含み得る。予め定められたキーワードを検出するためのキーワード検出モデルは、記憶ユニット 330 中に予め記憶され得るか、または通信ユニット 320 を介して外部サーバまたはデバイス（図示されない）からダウンロードされ得る。1 つの実施形態では、キーワード検出モデルは、複数の部分を含むサブワードのシーケンス（すなわち、複数のサブワードまたはサブワードユニット）を含み得、それらは、キーワードを示す 1 つ以上のサウンドサンプルから決定され得る。キーワード検出モデルはまた、サブワードのシーケンス中の複数のサブワードの各々に関連付けられたモデルパラメータ、およびキーワードを検出するための閾値を含み得る。

10

【 0 0 5 0 】

[0060] 別の実施形態では、キーワード検出モデルは、サブワードネットワークを含み得る。サブワードネットワークは、複数のノードのうちの少なくとも 2 つのノードを接続し得る複数のノードおよび複数の線を含み得る。キーワード検出モデルはまた、隠れマルコフモデル（HMM : hidden Markov model）、セミマルコフモデル（SMM : semi-Markov model）などのような少なくとも 1 つのグラフィカルモデルを含み得、それは、サブワードネットワークのノードに対応する。グラフィカルモデルは、推移確率、状態出力確率などのような、多数の状態およびパラメータを含み得る。

20

【 0 0 5 1 】

[0061] プロセッサ 340 中のサウンドサンプル検証ユニット 342 は、サウンドセンサ 300 または記憶ユニット 330 から 1 つ以上のサウンドサンプルを受け取るように構成され得、受け取ったサウンドサンプルの各々がサウンド検出モデルを生成する際に使用され得るかどうかを決定する。追加的にまたは代替的に、サウンドサンプル検証ユニット 342 は、サウンドセンサ 300 または記憶ユニット 330 から 1 つ以上のサウンドサンプルを含む入力サウンドストリームを受け取り、任意の適切な終点検出アルゴリズムを使用して、入力サウンドストリームからサウンドサンプルの各々を抽出し得る。この場合、サウンドサンプル検証ユニット 342 は、抽出されたサウンドサンプルを記憶ユニット 330 中に記憶し得る。

30

【 0 0 5 2 】

[0062] 一旦サウンドサンプルが受け取られるかまたは抽出されると、サウンドサンプル検証ユニット 342 は、サウンドサンプルの各々がサウンド検出モデルを生成する際に使用され得るかどうかを決定するために、サウンドサンプルの各々から 1 つ以上の音響特徴を抽出し得る。1 つの実施形態では、サウンドサンプル検証ユニット 342 は、サウンド検出モデルを生成する際の使用のために、サウンドサンプルの音響特徴が類似するかどうかを決定し得る。別の実施形態では、サウンドサンプル検証ユニット 342 は、サウンドサンプルの各々のサウンド品質がサウンド検出モデルを生成する際の使用のための閾値品質以上であるかどうかを、サウンド品質（例えば、S N R、サウンド強度レベルなど）を示す音響特徴に基づいて決定し得る。いくつかの実施形態では、サウンドサンプル検証ユニット 342 は、関連付けられた音響特徴に基づいて、サウンドサンプル間で最も高い品質のサウンドサンプルを決定するように構成され得る。決定された最も高い品質のサウンドサンプルは、サウンド検出モデルについてのサブワードのシーケンスを生成するために、サウンド検出モデル生成ユニット 344 に提供され得る。

40

【 0 0 5 3 】

[0063] 追加的にまたは代替的に、サウンドサンプル検証ユニット 342 は、I / O ユ

50

ニット310を介して、サウンドサンプルのうちの1つを最も優れたサウンドサンプルとして識別する、ユーザからの入力を受け取り得る。この場合、識別されたサウンドサンプルは、サウンド検出モデルについてのサブワードのシーケンスを生成するために、サウンド検出モデル生成ユニット344に提供され得る。いくつかの実施形態では、サウンドサンプル検証ユニット342は、I/Oユニット310のスピーカを通して、最も優れたサウンドサンプルを再生するようI/Oユニット310に指示し得る。追加的にまたは代替的に、サウンドサンプル検証ユニット342は、I/Oユニット310を介して再生される1つ以上のサウンドサンプルを識別する、ユーザからの入力を受け取り、I/Oユニット310のスピーカを通して、識別されたサウンドサンプルを再生するようI/Oユニット310に指示し得る。

10

## 【0054】

[0064] いくつかの実施形態では、サウンドサンプル検証ユニット342は、サウンドセンサ300または記憶ユニット330から1つ以上の代替のまたは追加のサウンドサンプルを受け取り得る。例えば、1つ以上の前に受け取ったサウンドサンプルがサウンド検出モデルを生成する際に使用されない可能性があることをサウンドサンプル検証ユニット342が決定する場合、1つ以上の前に受け取ったサウンドサンプルを廃棄し、サウンドセンサ300を介して、廃棄されたサウンドサンプルに対する1つ以上の新規のサウンドサンプルを受け取り得る。追加的にまたは代替的に、サウンド検出モデルを生成する際に使用されると決定される、前に受け取ったサウンドサンプルがない場合、サウンドサンプル検証ユニット342は、前に受け取ったサウンドサンプルを廃棄し、サウンドセンサ300を介して1つ以上の新規のサウンドサンプルを受け取り得る。

20

## 【0055】

[0065] 1つの実施形態によると、サウンドサンプル検証ユニット342は、サウンドサンプルを次々に受け取り、次に、受け取ったサウンドサンプルを一括モードで処理し得る。例えば、5つのサウンドサンプルが受け取られるとき、サウンドサンプル検証ユニット342は、5つのサウンドサンプル全てがサウンド検出モデルを生成する際に使用され得るかどうかを決定し、5つのサウンドサンプル全ての決定結果の指示を出力するようI/Oユニット310に指示し得る。別の実施形態では、サウンドサンプル検証ユニット342は、サウンドサンプルを1つずつ受け取り、処理し得る。

30

## 【0056】

[0066] サウンド検出モデル生成ユニット344は、サウンド検出モデルを生成する際に使用されると決定されている1つ以上のサウンドサンプルをサウンドサンプル検証ユニット342から受け取るように構成され得る。サウンド検出モデル生成ユニット344は次に、受け取ったサウンドサンプルの少なくとも1つに基づいて、サウンド検出モデルを生成し得る。サウンド検出モデルは、スピーカ検証モデル、予め定められたキーワードまたはユーザによって定義されたキーワードについてのキーワード検出モデルなどを含み得る。サウンド検出モデル生成ユニット344は、サウンド検出モデルを生成する際に使用されると決定されている受け取ったサウンドサンプルのうちの少なくとも1つに基づいて、サウンド検出モデルについての閾値を決定し得る。1つの実施形態では、サウンド検出モデル生成ユニット344は、低いS/NRのために廃棄されているかまたは新規のサウンドサンプルと置き換えられている可能性がある少なくとも1つのサウンドサンプルを、生成されたサウンド検出モデルを調整する際に、ノイズが含まれたサウンドサンプルとしてそのサウンドサンプルを使用するために、サウンドサンプル検証ユニット342から受け取り得る。追加的にまたは代替的に、サウンド検出モデル生成ユニット344は、低いS/NR、低いサウンド強度レベルなどのために廃棄されている可能性のある少なくとも1つのサウンドサンプルをサウンドサンプル検証ユニット342から受け取り、サウンド検出モデルについての閾値を調整する際にそのサウンドサンプルを使用し得る。サウンド検出モデル生成ユニット344は、生成されたサウンド検出モデルを記憶ユニット330中に記憶し得る。

40

## 【0057】

50

[0067] サウンド検出ユニット346は、入力サウンドストリームをサウンドセンサ300または記憶ユニット330から受け取り、記憶ユニット330に記憶された少なくとも1つのサウンド検出モデルに基づいて、入力サウンドストリームのサウンド信号またはデータにおいて、キーワードおよび/または特定のユーザを検出するように構成され得る。例えば、キーワードを示す入力サウンドストリームを受け取ると、サウンド検出ユニット346は、入力サウンドストリームから複数の音響特徴を順に抽出し、少なくとも1つのキーワード検出モデルに関連付けられた少なくとも1つのキーワード（少なくとも1つのユーザによって定義されたキーワードおよび少なくとも1つの予め定められたキーワードを含む）に対するマッチングスコアを決定し得る。サウンド検出ユニット346は次に、最も高いマッチングスコアを有するキーワード検出モデルを識別し、かつそのマッチングスコアが入力サウンドストリームへのマッチングしたキーワード検出モデルとして閾値以上であるかを識別し得る。  
10

#### 【0058】

[0068] キーワードに関連付けられるマッチングしたキーワード検出モデルを識別すると、サウンド検出ユニット346は、キーワードに関連付けられた機能を実行するか、またはキーワードに関連付けられたアプリケーションをアクティブ化、制御、または動作し得る。追加的にまたは代替的に、サウンド検出ユニット346は、キーワードに関連付けられ得るボイスアシスタントユニット348をオンにするために、アクティブ化信号を生成および送信し得る。ボイスアシスタントユニット348は、サウンド検出ユニット346からのアクティブ化信号に応答してアクティブ化され得る。一旦アクティブ化されると、ボイスアシスタントユニット348は、ディスプレイスクリーン110上におよび/またはI/Oユニット310のスピーカを通じて「ご用件は何でしょうか？」のようなメッセージを出力することによって、ボイスアシスタント機能を実行し得る。これに応答して、ユーザは、電子デバイス100の様々な関連付けられた機能をアクティブ化するためのボイスコマンドを話し得る。例えば、インターネット検索のためのボイスコマンドが受け取られるとき、ボイスアシスタントユニット348は、検索コマンドとしてボイスコマンドを認識し、通信ユニット320を介してウェブ検索を実行し得る。  
20

#### 【0059】

[0069] 図4は、本開示の1つの実施形態に従った、1つ以上のサウンドサンプルが、サウンド検出モデルを生成する際に使用され得るかどうかを決定するように構成された、サウンドサンプル検証ユニット342のブロック図を図示する。サウンドサンプル検証ユニット342は、特徴抽出器400、特徴バッファ410、および類似性決定ユニット420を含み得る。特徴抽出器400は、1つ以上のサウンドサンプルをサウンドセンサ300または記憶ユニット330から受け取り、受け取ったサウンドサンプルの各々から1つ以上の音響特徴を抽出するように構成され得る。例えば、特徴抽出器400は、受け取ったサウンドサンプルの各々からスペクトル特徴、時間領域特徴などを抽出し得る。1つの実施形態では、特徴抽出器400は、各サウンドサンプルについてのサブワードのシーケンスを音響特徴として生成するために、受け取ったサウンドサンプルの各々でサブワード認識を実行し得る。別の実施形態では、特徴抽出器400は、受け取ったサウンドサンプルの各々から、サウンド強度レベル（例えば、平均的なサウンドインテンシブレベル（sound intensive level））またはS N Rを抽出し得る。  
30

#### 【0060】

[0070] いくつかの実施形態では、特徴抽出器400は、受け取ったサウンドサンプルの各々のR Tを計算し得る。この処理では、特徴抽出器400は、サウンドサンプルの各々を均等な時間期間の複数の連続的なフレームに分割し、1つ以上の音響特徴（例えば、スペクトル特徴）をそれらフレームから抽出し得る。フレーム中の音響特徴に基づいて、複数のフレーム間の1つ以上の音響特徴における類似性が相関値として決定され得る。相関値に基づいて、特徴抽出器400は、サウンドサンプルのR Tとして、閾値相関値以上の相関値を有するサウンドサンプル中の2つのフレーム間の最も長い時間期間を決定し得る。  
40

## 【0061】

[0071] サウンドサンプルの各々から 1 つ以上の音響特徴を抽出すると、特徴抽出器 400 は、抽出された音響特徴を特徴バッファ 410 中に記憶し得る。特徴バッファ 410 は、抽出された音響特徴を特徴抽出器 400 から受け取り、類似性決定ユニット 420 にそれらを送信するように構成され得る。抽出された音響特徴はまた、記憶ユニット 330 中に記憶され得る。

## 【0062】

[0072] 類似性決定ユニット 420 は、抽出された音響特徴を特徴バッファ 410 または記憶ユニット 330 から受け取り、サウンドサンプルの抽出された音響特徴が類似するかどうかを決定し得る。1 つの実施形態では、類似性決定ユニット 420 は、関連付けられた音響特徴を比較することによって、サウンドサンプルの各々のペアの音響特徴が類似するかどうかを決定し得る。少なくとも 2 つのサウンドサンプルの音響特徴が類似すると決定される場合、類似性決定ユニット 420 は、その少なくとも 2 つのサウンドサンプルが類似することを決定し得、サウンド検出モデルを生成する際に使用され得る。一方、類似性決定ユニット 420 が、サウンドサンプルの音響特徴が少なくとも 2 つの他のサウンドサンプルの音響特徴に類似しないと決定する場合、それは、サウンドサンプルがサウンド検出モデルを生成する際に使用されない可能性があることを決定し得る。

10

## 【0063】

[0073] 別の実施形態では、類似性決定ユニット 420 は、2 つ以上のサウンドサンプルの組み合わせられた音響特徴を決定し、サウンドサンプルの音響特徴が、その組み合わせられた音響特徴に類似するかどうかを決定し得る。組み合わせられた音響特徴は、特徴バッファ 410 または記憶ユニット 330 中に記憶され得る。例えば、第 1 および第 2 のサウンドサンプルの音響特徴が類似すると決定される場合、類似性決定ユニット 420 は、第 1 および第 2 のサウンドサンプルの音響特徴に基づいて、組み合わせられた音響特徴を決定し得る。その際、第 3 のサウンドサンプルの音響特徴が、組み合わせられた音響特徴に類似すると決定される場合、類似性決定ユニット 420 は、第 1 、第 2 、および第 3 のサウンドサンプルが類似することを決定し、サウンド検出モデルを生成する際に使用され得る。一方、類似性決定ユニット 420 が、第 3 のサウンドサンプルの音響特徴が組み合わせられた音響特徴に類似しないと決定する場合、それは、第 3 のサウンドサンプルがサウンド検出モデルを生成する際に使用されない可能性があることを決定し得る。

20

## 【0064】

[0074] 図 5 は、本開示の 1 つの実施形態に従った、サウンド検出モデルを生成する際に使用される 1 つ以上のサウンドサンプルを検証するために、電子デバイス 100 において実行される方法 500 のフローチャートである。始めに、電子デバイス 100 は、510 において、サウンド検出モデルを生成するために第 1 のサウンドサンプルを受け取り得る。電子デバイス 100 は、520 において、第 1 のサウンドサンプルから第 1 の音響特徴を抽出し得る。電子デバイス 100 は、530 において、サウンド検出モデルを生成するために第 2 のサウンドサンプルを受け取り得る。電子デバイス 100 は、540 において、第 2 のサウンドサンプルから第 2 の音響特徴を抽出し得る。電子デバイス 100 は、550 において、第 2 の音響特徴が第 1 の音響特徴に類似するかどうかを決定し得る。

30

## 【0065】

[0075] 図 6 は、本開示の 1 つの実施形態に従った、組み合わせられた音響特徴を決定するために、電子デバイス 100 において実行される方法 600 のフローチャートである。始めに、電子デバイス 100 は、610 において、第 1 のサウンドサンプルを受け取り、第 1 のサウンドサンプルから第 1 の音響特徴を抽出し得る。電子デバイス 100 は、620 において、第 2 のサウンドサンプルを受け取り、第 2 のサウンドサンプルから第 2 の音響特徴を抽出し得る。電子デバイス 100 は、630 において、第 2 の音響特徴が第 1 の音響特徴に類似するかどうかを決定し得る。第 1 および第 2 の音響特徴が類似すると決定される（すなわち、640 で YES である）場合、電子デバイス 100 は、650 において、第 1 および第 2 の音響特徴に基づいて、組み合わせられた音響特徴を決定し得る。

40

50

第1および第2の音響特徴が類似しないと決定される（すなわち、640でNOである）場合、方法600は、第1のサウンドサンプルを受け取るために610へと進み、第1のサウンドサンプルから第1の音響特徴を抽出し得る。

【0066】

[0076] 図7は、本開示の別の実施形態に従った、サウンド検出モデルを生成する際に使用される1つ以上のサウンドサンプルを検証するために、電子デバイス100で実行される方法700のフローチャートである。図6で図示されるように、電子デバイス100は、第1および第2の音響特徴に基づいて、組み合わせられた音響特徴を決定し得る。電子デバイス100は、710において、サウンド検出モデルを生成するために第3のサウンドサンプルを受け取り得る。電子デバイス100は、720において、第3のサウンドサンプルから第3の音響特徴を抽出し得る。電子デバイス100は、730において、第3の音響特徴が第1および第2の音響特徴の組み合わせられた音響特徴に類似するかどうかを決定し得る。第3の音響特徴が、組み合わせられた音響特徴に類似すると決定される（すなわち、740でYESである）場合、電子デバイス100は、750において、少なくとも1つの第1のサウンドサンプル、第2のサウンドサンプル、または第3のサウンドサンプルに基づいて、サウンド検出モデルを生成し得る。第3の音響特徴が、組み合わせられた音響特徴に類似しないと決定される（すなわち、740でNOである）場合、方法700は、第3のサウンドサンプルを受け取るために710へと進む。

10

【0067】

[0077] 図8は、本開示の1つの実施形態に従った、一括モードでサウンド検出モデルを生成する際に使用される1つ以上のサウンドサンプルを検証するために、電子デバイスにおいて実行される方法800のフローチャートである。電子デバイス100は、810において、第2のサウンド検出モデルを生成するために複数のサウンドサンプルを受け取り得る。電子デバイス100は、820において、受け取ったサウンドサンプルが類似するかどうかを決定し得る。サウンドサンプルが類似すると決定される（すなわち、830でYESである）場合、電子デバイス100は、850において、サウンドサンプルの少なくとも1つに基づいて、サウンド検出モデルを生成し得る。サウンドサンプルが類似しないと決定される（すなわち、830でNOである）場合、電子デバイス100は、840において、他のサウンドサンプルに類似しないと決定されている少なくとも1つのサウンドサンプルに対する代替のサウンドサンプルとして、少なくとも1つの新規のサウンドサンプルを受け取り得、方法800は、受け取ったサウンドサンプルが類似するかどうかを決定するために、820へと進み得る。

20

30

【0068】

[0078] 図9は、本開示の1つの実施形態に従った、1つ以上のサウンドサンプルの音響特徴が類似するかどうかを決定するために、電子デバイス100において実行される方法820のフローチャートである。電子デバイス100は、900において、サウンドサンプルの各々から1つ以上の音響特徴を抽出し得る。電子デバイス100は、910において、サウンドサンプルの各ペアの音響特徴が類似するかどうかを決定し得る。電子デバイス100は、920において、サウンドサンプルの各ペアの音響特徴が類似するかどうかを決定することに基づいて、サウンドサンプルが類似するかどうかの指示を出力し得る。

40

【0069】

[0079] 図10は、いくつかの実施形態に従った、サウンド検出モデルを生成する際に使用される1つ以上のサウンドサンプルを検証するために、本開示の方法および装置が実装され得るワイヤレス通信システム中の例示的な電子デバイス1000のブロック図を図示する。例示的な電子デバイス1000は、セルラフォン、スマートフォン、ウェアラブルコンピュータ、スマートウォッチ、スマートメガネ、タブレットパーソナルコンピュータ、端末、ハンドセット、パーソナルデジタルアシスタント（PDA）、ワイヤレスモデム、コードレスフォン、タブレットなどであり得る。ワイヤレス通信システムは、CDMAシステム、GSM（登録商標）システム、W-CDMA（登録商標）システム、LTE

50

(登録商標)システム、LTEアドバンストシステムなどであり得る。

【0070】

[0080] 例示的な電子デバイス1000は、受信バスおよび送信バスを介して双方向通信を提供することが可能であり得る。受信バスにおいて、基地局によって送信された信号は、アンテナ1012によって受信され得、受信機(RCVR)1014に提供され得る。受信機1014は、受信した信号を調整およびデジタル化し得、さらなる処理のために、調整されデジタル化されたデジタル信号をデジタルセクションに提供し得る。送信バスでは、送信機(TMTR)1016が、デジタルセクション1020から送信されるデータを受信し、このデータを処理および調整し、変調された信号を生成し得、それは、アンテナ1012を基地局に送信される。受信機1014および送信機1016は、CDMA、GSM、W-CDMA、LTE、LTEアドバンストなどをサポートし得るトランシーバーの一部であり得る。

10

【0071】

[0081] デジタルセクション1020は、例えば、モデムプロセッサ1022、縮小命令セットコンピュータ/デジタル信号プロセッサ(RISC/DSP)1024、コントローラ/プロセッサ1026、内部メモリ1028、汎用オーディオ/ビデオエンコーダ1032、汎用オーディオデコーダ1034、グラフィックス/ディスプレイプロセッサ1036、および外部バスインターフェース(EBI)1038のような様々な処理、インターフェース、およびメモリユニットを含み得る。モデムプロセッサ1022は、例えば、符号化、変調、復調、変調、および復号などの、データ送信および受信のための処理を実行し得る。RISC/DSP1024は、例示的な電子デバイス1000のための汎用および専用の処理を実行し得る。コントローラ/プロセッサ1026は、デジタルセクション1020中の様々な処理およびインターフェースユニットの動作を実行し得る。内部メモリ1028は、デジタルセクション1020中の様々なユニットについてのデータおよび/または命令を記憶し得る。

20

【0072】

[0082] 汎用オーディオ/ビデオエンコーダ1032は、オーディオ/ビデオソース1042、マイクロフォン1044、画像センサ1046などからの入力信号を符号化することを実行し得る。汎用オーディオデコーダ1034は、コード化されたオーディオデータに対する復号を実行し、スピーカ/ヘッドセット1048に出力信号を提供し得る。グラフィックス/ディスプレイプロセッサ1036は、グラフィックス、ビデオ、画像、およびテキストに対する処理を実行し得、それらは、ディスプレイユニット1050に表示され得る。EBI1038は、デジタルセクション1020とメインメモリ1052との間のデータの転送を容易にし得る。

30

【0073】

[0083] デジタルセクション1020は、1つ以上のプロセッサ、DSP、マイクロプロセッサ、RISCなどで実装され得る。デジタルセクション1020はまた、1つ以上の特定用途向け集積回路(ASSIC)および/またはいくつかの他のタイプの集積回路(ASIC)上で組み立てられ得る。

40

【0074】

[0084] 図11は、いくつかの実施形態に従って実装された、先に説明されたサーバのうちのいずれか1つであり得るサーバシステム1100を図示するブロック図である。サーバシステム1100は、1つ以上の処理ユニット(例えば、CPU)1102、1つ以上のネットワークまたは他の通信ネットワークインターフェース、メモリ1112、およびこれらの構成要素を相互接続するための1つ以上の通信バス1114を含み得る。サーバシステム1100はまた、ディスプレイデバイスおよびキーボードを有するユーザインターフェース(図示されない)を含み得る。

【0075】

[0085] メモリ1112は、高速ランダムアクセスメモリ(例えば、DRAM、SRAM、DDR RAM、または他のランダムアクセスソリッドステートメモリデバイス)の

50

のような任意の適切なメモリであり得る。メモリ 1112 は、不揮発性メモリ（例えば、1つ以上の磁気ディスク記憶デバイス、光学ディスク記憶デバイス、フラッシュメモリデバイス、または他の不揮発性シリッドステート記憶デバイス）を含み得るか、または代替的にそれら不揮発性メモリであり得る。いくつかの実施形態では、メモリ 1112 は、CPU 1102 から離れて位置付けられるおよび／または複数の場所 (sites) に離れて位置付けられる1つ以上の記憶デバイスを含み得る。

【0076】

[0086] メモリ 1112 によって表わされる上記メモリデバイスのうちの任意の1つは、前に説明された任意の処理、動作、および方法を実行するおよび／または行うための命令のセットに対応する任意の数のモジュールまたはプログラムを記憶し得る。例えば、メモリ 1112 は、様々な基本的システムサービスに対処するための、およびハードウェア依存タスクを実行するためのプロシージャを含む命令を記憶するように構成されたオペレーティングシステム 1116 を含み得る。メモリ 1112 のネットワーク通信モジュール 1118 は、インターネット、他のワイドエリアネットワーク、ローカルエリアネットワーク、大都市エリアネットワークなどのような、1つ以上の通信ネットワークおよび1つ以上の通信ネットワークインターフェース 1110（有線またはワイヤレス）を介して、サーバシステム 1100 を他のコンピュータに接続するために使用され得る。

10

【0077】

[0087] メモリ 1112 はまた、1つ以上のオブジェクト（例えば、テキストオブジェクトおよび非テキストオブジェクト）、分類データベース、文字情報データベース、辞書データベースなどを有する複数の画像の画像データベースを含むように構成されるデータベース 1120 を含み得る。オペレーティングシステム 1116 は、ネットワーク通信モジュール 1118 を通じて、受信および／またはキャプチャされ得る様々な画像を用いて画像データベースをアップデートし得る。オペレーティングシステム 1116 はまた、ネットワーク通信モジュール 1118 を介して、画像を複数の電子デバイスに提供し得る。加えて、分類データベース、文字情報データベース、辞書データベースは、画像の少なくとも1つのテキスト領域を検出する際に使用するために、複数の電子デバイスに提供される、および／または少なくとも1つのテキスト領域において1つ以上の文字ストリングを認識し得る。

20

【0078】

[0088] 一般に、本明細書で説明される任意のデバイスは、ワイヤレスフォン、セルラーフォン、ラップトップコンピュータ、ワイヤレスマルチメディアデバイス、ワイヤレス通信パーソナルコンピュータ（P C）カード、P D A、外部または内部モデム、無線チャネルを通じて通信するデバイスなどのような、様々なタイプのデバイスを表し得る。デバイスは、アクセス端末（A T）、アクセスユニット、加入者ユニット、移動局、モバイルデバイス、モバイルユニット、モバイルフォン、モバイル、遠隔局、遠隔端末、遠隔ユニット、ユーザデバイス、ユーザ機器、ハンドヘルドデバイスなどのような様々な名称を有し得る。本明細書で説明される任意のデバイスは、命令およびデータを記憶するためのメモリ、並びにハードウェア、ソフトウェア、ファームウェア、またはこれらの組み合わせを有し得る。

30

【0079】

[0089] 本明細書で説明される技法は、様々な手段によって実装され得る。例えば、これらの技術は、ハードウェア、ファームウェア、ソフトウェア、またはこれらの組み合わせで実装され得る。当業者はさらに、本明細書の開示に関連して説明される様々な例示のための論理ブロック、モジュール、回路、アルゴリズムステップが電子ハードウェア、コンピュータソフトウェア、または両方の組み合わせとして実装され得ることを理解するだろう。ハードウェアおよびソフトウェアのこの互換性を明確に例示するために、様々な例示のための構成要素、ブロック、モジュール、回路、およびステップがそれらの機能の観点から一般的に上記に説明されている。このような機能が、ハードウェアとして実装されるか、またはソフトウェアとして実装されるかは、特定のアプリケーションおよびシステ

40

50

ム全体に課せられた設計制約に依存する。当業者は、説明される機能を特定のアプリケーションごとに異なる方法で実装し得るが、このような実装の決定は、本開示の範囲からの逸脱を引き起こしていると解釈されるべきではない。

#### 【0080】

[0090] ハードウェアの実装について、技術を実行するために使用される処理ユニットは、1つ以上のASIC、DSP、デジタル信号処理デバイス(DSPD)、プログラマブル論理デバイス(PLD)、フィールドプログラマブルゲートアレイ(FPGA)、プロセッサ、コントローラ、マイクロコントローラ、マイクロプロセッサ、電子デバイス、本明細で説明される機能を実行するように設計された他の電子ユニット、コンピュータ、またはこれらの組み合わせ内で実装され得る。

10

#### 【0081】

[0091] よって、本明細書の開示に関連して説明される様々な例示のための論理ブロック、モジュール、および回路は、汎用プロセッサ、DSP、ASIC、FPGAまたは他のプログラマブル論理デバイス、ディスクリートゲートまたはトランジスタ論理、ディスクリートハードウェア構成要素、または本明細書で説明される機能を実行するように設計されるこれらの任意の組み合わせで実装または実行され得る。汎用プロセッサはマイクロプロセッサであり得るが、代替として、プロセッサは、任意のプロセッサ、コントローラ、マイクロプロセッサ、またはステートマシンであり得る。プロセッサはまた、例えば、DSPとマイクロプロセッサとの組み合わせ、複数のマイクロプロセッサ、DSPコアと結合された1つ以上のマイクロプロセッサ、または任意の他のこののような構成のようなコンピューティングデバイスの組み合わせとして実装され得る。

20

#### 【0082】

[0092] ソフトウェアで実装される場合、機能は、コンピュータ可読媒体に記憶され得る。コンピュータ可読媒体は、1つの場所から別の場所へのコンピュータプログラムの転送を容易にする任意の媒体を含む通信媒体とコンピュータ記憶媒体との両方を含む。記憶媒体は、コンピュータによってアクセスされ得る任意の利用可能な媒体であり得る。例として、限定はされないが、このようなコンピュータ可読媒体は、RAM、ROM、EEPROM、CD-ROM、または他の光ディスク記憶装置、磁気ディスク記憶装置または他の磁気記憶デバイス、あるいは、命令またはデータ構造の形式で所望されるプログラムコードを搬送または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備え得る。本明細書で使用されるようなディスク(disk)およびディスク(disc)は、コンパクトディスク(CD)、レーザーディスク(登録商標)、光ディスク、デジタル多目的ディスク(DVD)、フロッピー(登録商標)ディスク、およびBluray(登録商標)ディスクを含み、ここで、ディスク(disks)が通常磁気的にデータを再生する一方、ディスク(discs)は、レーザーを用いて光学的にデータを再生する。上記の組み合わせもまた、コンピュータ可読媒体の範囲内に含まれるべきである。例えば、コンピュータ可読記憶媒体は、プロセッサによって実行可能な命令を含む、非一時的なコンピュータ可読記憶デバイスであり得る。よって、コンピュータ可読記憶媒体は、信号ではない可能性がある。

30

#### 【0083】

[0093] 本開示の上記説明は、当業者が本開示を実施および使用することを可能にするために提供される。本開示への様々な修正は、当業者にとって容易に明らかであり、本明細書に定義される一般的な原理は、本開示の範囲から逸脱することなく、他のバリエーションに適用され得る。よって、本開示は、本明細書で説明される例に限定されることが意図されるものではなく、本明細書に開示される原理および新規の特徴と一致する最も広い範囲が付与されるべきものである。

40

#### 【0084】

[0094] 例示的な実装は、1つ以上の独立型(stand-alone)コンピュータシステムのコンテキストにおいて、本開示の主題の態様を利用するように述べられているが、主題は、限定はされないが、むしろ、ネットワークまたは分散型コンピューティング環境のよう

50

な、任意のコンピューティング環境に関連して実装され得る。さらに、本開示の主題の態様は、複数の処理チップ内でまたはデバイスを介して実装され得、記憶媒体も同様に、複数のデバイスを介して影響を受け得る。このようなデバイスは、P C、ネットワークサーバ、およびハンドヘルドデバイスを含み得る。

## 【0085】

[0095] 主題は、構造上の特徴および／または方法論的動作に特有の表現で説明されているが、添付の請求項に定義される主題は、上述される特定の特徴または動作に必ずしも限定されないことが理解されるべきである。むしろ、上述された特定の特徴および動作は、請求項を実現する形態の例として開示されている。

## 【0086】

[0096] 上述されるモジュールまたはプログラム（すなわち、命令のセット）は、別個のソフトウェアプログラム、プロシージャ、またはモジュールとして実装される必要がなく、そのため、様々な実施形態においてこれらのモジュールの様々なサブセットが組み合わせられ得るか、またはそうでなければ再配置され得ることが理解されるだろう。さらに、メモリ 1 1 1 2 は、上述されていない追加のモジュールおよびデータ構造を記憶し得る。

## 【0087】

<本開示の態様>

[0097] 以下に、本開示のいくつかの態様がさらに記載される。

## 【0088】

[0098] (例 1) 本開示の態様によると、サウンド検出モデルを生成する際に使用される少なくとも 1 つのサウンドサンプルを検証するための方法が提供され、方法は、第 1 のサウンドサンプルを受け取ることと、第 1 のサウンドサンプルから第 1 の音響特徴を抽出することと、第 2 のサウンドサンプルを受け取ることと、第 2 のサウンドサンプルから第 2 の音響特徴を抽出することと、第 2 の音響特徴が第 1 の音響特徴に類似するかどうかを決定することと、を備える。

## 【0089】

[0099] (例 2) 例 1 の方法は、第 2 の音響特徴が第 1 の音響特徴に類似することを決定することに応答して、第 1 のサウンドサンプルまたは第 2 のサウンドサンプルのうちの少なくとも 1 つに基づいて、サウンド検出モデルを生成することをさらに備える。

## 【0090】

[00100] (例 3) 例 1 または 2 の方法は、第 1 のサウンドサンプルの信号対雑音比 (S N R) を決定することと、第 2 のサウンドサンプルの S N R を決定することと、第 1 および第 2 のサウンドサンプルの S N R に基づいて、第 1 のサウンドサンプルまたは第 2 のサウンドサンプルのうちの少なくとも 1 つを選択することと、第 2 の音響特徴が第 1 の音響特徴に類似することを決定すると、選択された少なくとも 1 つのサウンドサンプルに基づいて、サウンド検出モデルを生成することと、選択されていないサウンドサンプルに基づいて、生成されたサウンド検出モデルを調整することと、をさらに備える。

## 【0091】

[00101] (例 4) 例 1 乃至 3 のうちのいずれか 1 つの方法は、第 2 の音響特徴が第 1 の音響特徴に類似することを決定することに応答して、第 1 の音響特徴および第 2 の音響特徴に基づいて、組み合わせられた音響特徴を決定することをさらに備える。

## 【0092】

[00102] (例 5) 例 1 乃至 4 のうちのいずれか 1 つの方法は、第 3 のサウンドサンプルを受け取ることと、第 3 のサウンドサンプルから第 3 の音響特徴を抽出することと、第 3 の音響特徴が組み合わせられた音響特徴に類似するかどうかを決定することと、をさらに備える。

## 【0093】

[00103] (例 6) 例 1 乃至 5 のうちのいずれか 1 つの方法において、第 1 の音響特徴および第 2 の音響特徴の各々は、スペクトル特徴または時間領域特徴のうちの少なくとも

10

20

30

40

50

1つを含む。

【0094】

[00104] (例7) 例1乃至6のうちのいずれか1つの方法において、第1の音響特徴および第2の音響特徴の各々は、サブワードのシーケンスを含む。

【0095】

[00105] (例8) 例1乃至7のうちのいずれか1つの方法において、サブワードは、音、音素、トライフォン、または音節のうちの少なくとも1つを含む。

【0096】

[00106] (例9) 例1乃至8のうちのいずれか1つの方法は、第2の音響特徴が第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルを受け取ることと、新規のサウンドサンプルから新規の音響特徴を抽出することと、新規の音響特徴が第1の音響特徴に類似するかどうかを決定することと、をさらに備える。

10

【0097】

[00107] (例10) 例1乃至9の例のうちのいずれか1つの方法は、新規の音響特徴が第1の音響特徴に類似することを決定することに応答して、第1のサウンドサンプルまたは新規のサウンドサンプルのうちの少なくとも1つに基づいて、サウンド検出モデルを生成することをさらに備える。

【0098】

[00108] (例11) 例1乃至10のうちのいずれか1つの方法において、サウンド検出モデルを生成することは、第1のサウンドサンプルまたは新規のサウンドサンプルのうちの少なくとも1つに基づいて、サウンド検出モデルの閾値を決定することと、第2のサウンドサンプルに基づいて、閾値を調整することと、を備える。

20

【0099】

[00109] (例12) 例1乃至11のいずれか1つの方法は、第2の音響特徴が第1の音響特徴に類似するかどうかを決定することに基づいて、第2のサウンドサンプルが第1のサウンドサンプルに類似するかどうかの指示を出力することと、サウンド検出モデルを生成する際に使用される第1のサウンドサンプルまたは第2のサウンドサンプルのうちの少なくとも1つを示す入力を受け取ることと、をさらに備える。

【0100】

[00110] (例13) 例1乃至12のうちのいずれか1つの方法において、第1のサウンドサンプルおよび第2のサウンドサンプルの各々は、電子デバイスをアクティブ化するためのコマンド、あるいは電子デバイスにおけるアプリケーションまたは機能を制御するためのコマンドのうちの少なくとも1つを示す。

30

【0101】

[00111] (例14) 例1乃至13のうちのいずれか1つの方法は、音声入力を受け取ることと、サウンド検出モデルに基づいて、音声入力からキーワードまたはユーザのうちの少なくとも1つを認識することと、をさらに備える。

【0102】

[00112] (例15) 本開示の別の態様によると、第1のサウンドサンプルおよび第2のサウンドサンプルを受け取るように構成されたサウンドセンサと、第1のサウンドサンプルから第1の音響特徴を抽出し、第2のサウンドサンプルから第2の音響特徴を抽出し、第2の音響特徴が第1の音響特徴に類似するかどうかを決定するように構成されたサウンドサンプル検証ユニットと、を備える、サウンド検出モデルを生成する際に使用される少なくとも1つのサウンドサンプルを検証するための電子デバイスが提供される。

40

【0103】

[00113] (例16) 例15の電子デバイスは、第2の音響特徴が第1の音響特徴に類似することを決定することに応答して、第1のサウンドサンプルまたは第2のサウンドサンプルのうちの少なくとも1つに基づいて、サウンド検出モデルを生成するように構成されたサウンド検出モデル生成ユニットをさらに備える。

【0104】

50

[00114] (例14) 例15または16の電子デバイスにおいて、サウンドサンプル検証ユニットは、第2の音響特徴が第1の音響特徴に類似することを決定することに応答して、第1の音響特徴および第2の音響特徴に基づいて、組み合わせられた音響特徴を決定するように構成される。

【0105】

[00115] (例18) 例15乃至17のうちのいずれか1つの電子デバイスにおいて、サウンドセンサは、第3のサウンドサンプルを受け取るように構成され、サウンドサンプル検証ユニットは、第3のサウンドサンプルから第3の音響特徴を抽出し、第3の音響特徴が組み合わせられた音響特徴に類似するかどうかを決定するように構成される。

【0106】

[00116] (例19) 例15乃至18のうちのいずれか1つの電子デバイスにおいて、サウンドセンサは、第2の音響特徴が第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルを受け取るように構成され、サウンドサンプル検証ユニットは、新規のサウンドサンプルから新規の音響特徴を抽出し、新規の音響特徴が第1の音響特徴に類似するかどうかを決定するように構成される。

10

【0107】

[00117] (例20) 本開示のさらに別の態様によると、電子デバイスの少なくとも1つのプロセッサに、第1のサウンドサンプルを受け取り、第1のサウンドサンプルから第1の音響特徴を抽出し、第2のサウンドサンプルを受け取り、第2のサウンドサンプルから第2の音響特徴を抽出し、第2の音響特徴が第1の音響特徴に類似するかどうかを決定する、動作を実行させる命令を備える非一時的コンピュータ可読記憶媒体が提供される。

20

【0108】

[00118] (例21) 例20に記載の非一時的コンピュータ可読記憶媒体は、電子デバイスの少なくとも1つのプロセッサに、第2の音響特徴が第1の音響特徴に類似すると決定することに応答して、第1のサウンドサンプルまたは第2のサウンドサンプルのうちの少なくとも1つに基づいて、サウンド検出モデルを生成する動作を実行させる命令をさらに備える。

【0109】

[00119] (例22) 例20または21に記載の非一時的コンピュータ可読記憶媒体は、電子デバイスの少なくとも1つのプロセッサに、第2の音響特徴が第1の音響特徴に類似することを決定することに応答して、第1の音響特徴および第2の音響特徴に基づいて、組み合わせられた音響特徴を決定する動作を実行させる命令をさらに備える。

30

【0110】

[00120] (例23) 例20乃至22のうちのいずれか1つの非一時的コンピュータ可読記憶媒体は、電子デバイスの少なくとも1つのプロセッサに、第3のサウンドサンプルを受け取り、第3のサウンドサンプルから第3の音響特徴を抽出し、第3の音響特徴が組み合わせられた音響特徴に類似するかどうかを決定する動作を実行させる命令をさらに備える。

【0111】

[00121] (例24) 例20乃至23のうちのいずれか1つの非一時的コンピュータ可読記憶媒体は、電子デバイスの少なくとも1つのプロセッサに、第2の音響特徴が第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルを受け取り、新規のサウンドサンプルから新規の音響特徴を抽出し、新規の音響特徴が第1の音響特徴に類似するかどうかを決定する動作を実行させる命令をさらに備える。

40

【0112】

[00122] (例25) 本開示のさらに別の態様によると、第1のサウンドサンプルを受け取るための手段と、第1のサウンドサンプルから第1の音響特徴を抽出するための手段と、第2のサウンドサンプルを受け取るための手段と、第2のサウンドサンプルから第2の音響特徴を抽出するための手段と、第2の音響特徴が第1の音響特徴に類似するかどうかを決定するための手段と、を備える、サウンド検出モデルを生成する際に使用される少

50

なくとも 1 つのサウンドサンプルを検証するための電子デバイスが提供される。

【 0 1 1 3 】

[00123] (例 2 6) 例 2 5 の電子デバイスは、第 2 の音響特徴が第 1 の音響特徴に類似することを決定することに応答して、第 1 のサウンドサンプルまたは第 2 のサウンドサンプルのうちの少なくとも 1 つに基づいて、サウンド検出モデルを生成するための手段をさらに備える。

【 0 1 1 4 】

[00124] (例 2 7) 例 2 5 または 2 6 の電子デバイスは、第 2 の音響特徴が第 1 の音響特徴に類似することを決定することに応答して、第 1 の音響特徴および第 2 の音響特徴に基づいて、組み合わせられた音響特徴を決定するための手段をさらに備える。

10

【 0 1 1 5 】

[00125] (例 2 8) 例 2 5 乃至 2 7 のうちのいずれか 1 の電子デバイスは、第 3 のサウンドサンプルを受け取るための手段と、第 3 のサウンドサンプルから第 3 の音響特徴を抽出するための手段と、第 3 の音響特徴が組み合わせられた音響特徴に類似するかどうかを決定するための手段と、をさらに備える。

【 0 1 1 6 】

[00126] (例 2 9) 例 2 5 乃至 2 8 のうちのいずれか 1 つの電子デバイスは、第 2 の音響特徴が第 1 の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルを受け取るための手段と、新規のサウンドサンプルから新規の音響特徴を抽出するための手段と、新規の音響特徴が第 1 の音響特徴に類似するかどうかを決定するための手段と、をさらに備える。

20

【 0 1 1 7 】

[00127] (例 3 0) 例 2 5 乃至 2 9 のうちのいずれか 1 つの電子デバイスは、第 2 の音響特徴が第 1 の音響特徴に類似するかどうかを決定することに基づいて、第 2 のサウンドサンプルが第 1 のサウンドサンプルに類似するかどうかの指示を出力するための手段と、サウンド検出モデルを生成する際に使用される第 1 のサウンドサンプルまたは第 2 のサウンドサンプルのうちの少なくとも 1 つを示す入力を受け取るための手段と、をさらに備える。

【図1】

【図2】

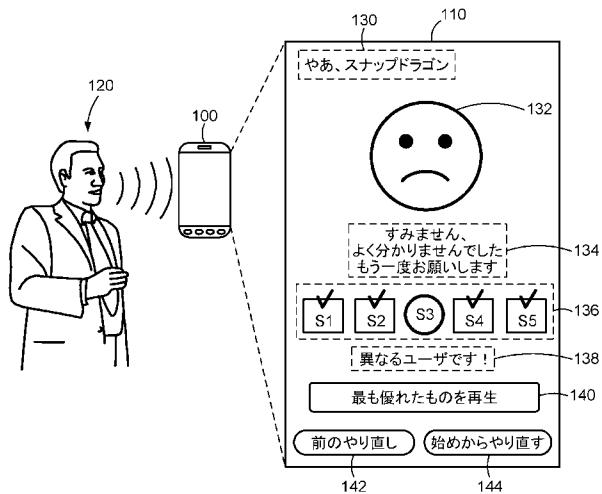


FIG. 1

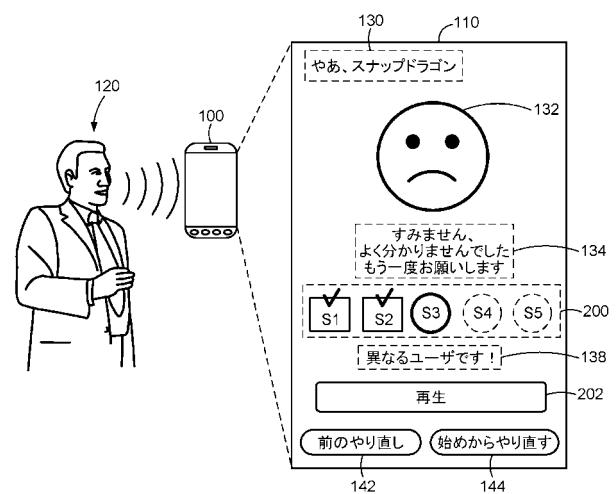


FIG. 2

【図3】

【図4】

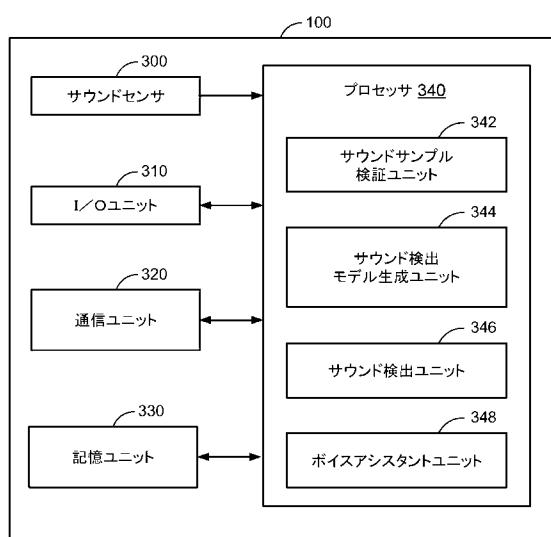


FIG. 3

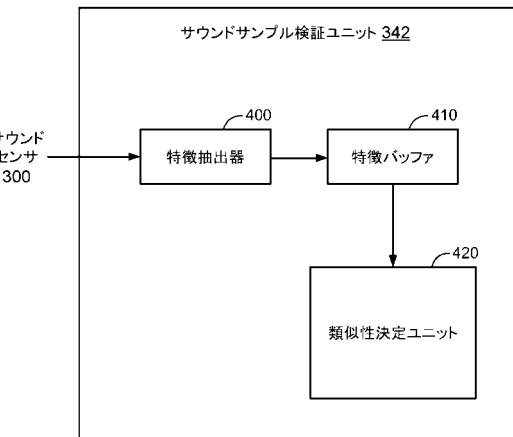


FIG. 4

【図5】

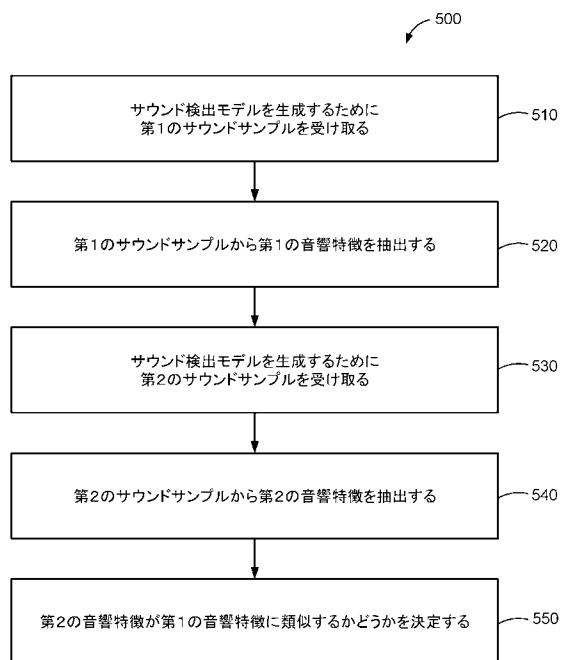


FIG. 5

【図6】

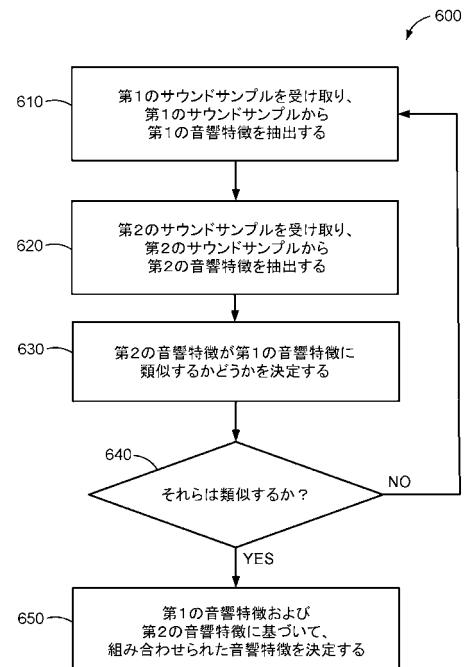


FIG. 6

【図7】

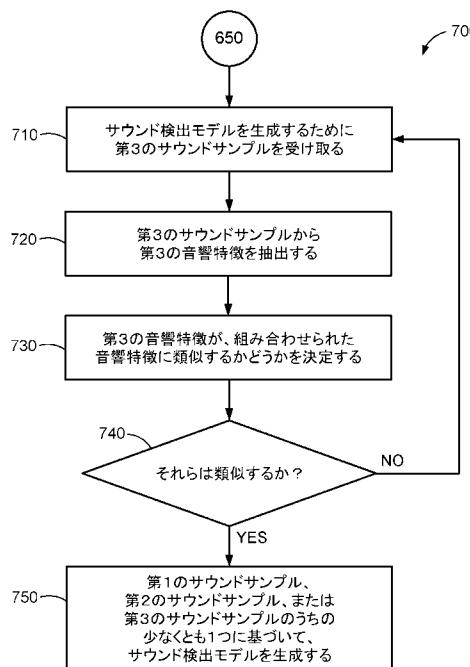


FIG. 7

【図8】

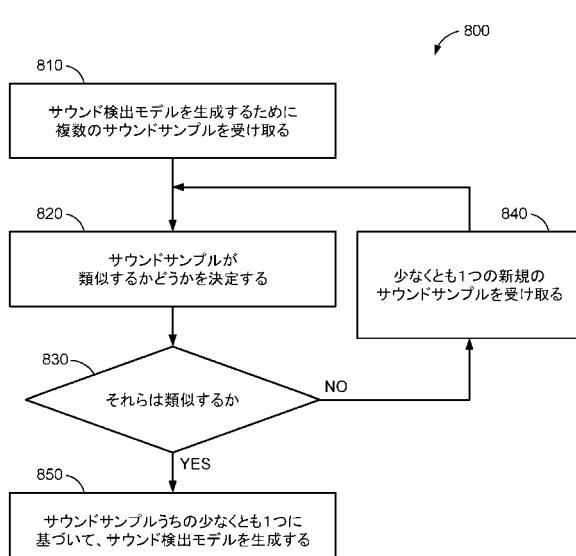


FIG. 8

【図 9】

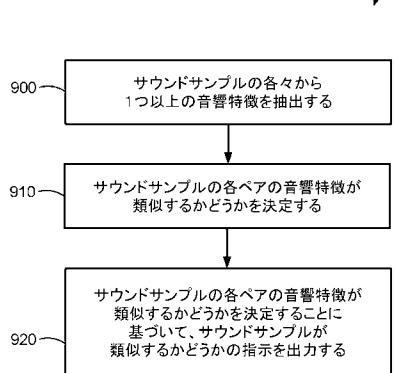


FIG. 9

【図 10】

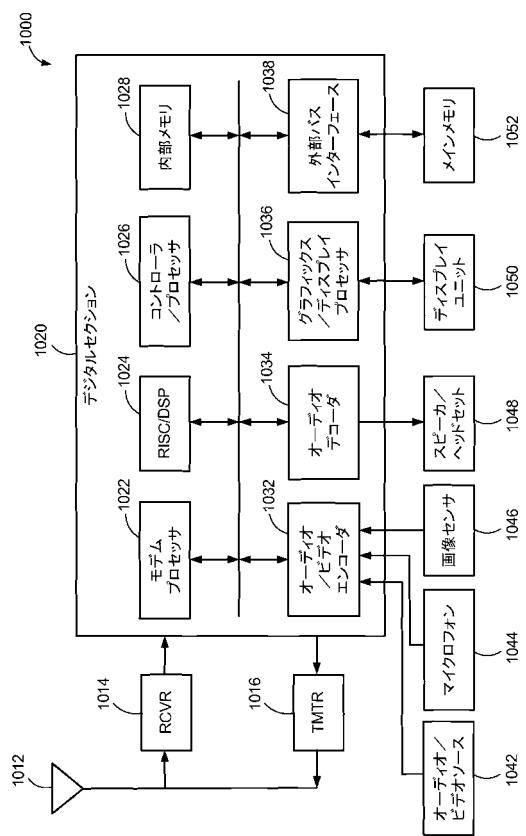


FIG. 10

【図 11】

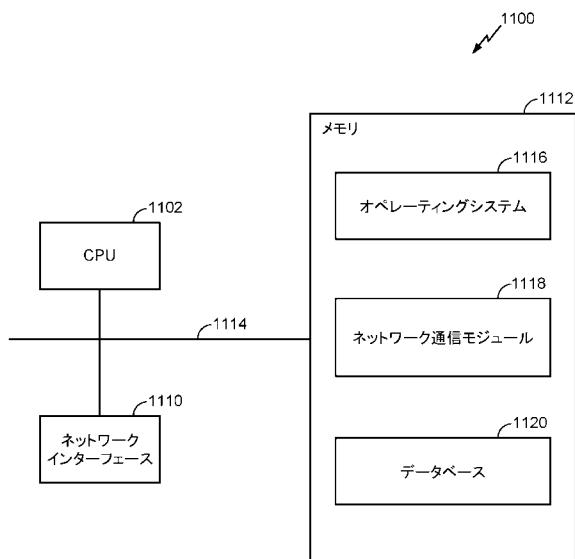


FIG. 11

**【手続補正書】**

【提出日】平成29年7月4日(2017.7.4)

**【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】**

サウンド検出モデルを生成する際に使用されるデータを決定するために、電子デバイスにおいて実行される方法であって、前記方法は、

第1のサウンドサンプルに関連付けられた第1のデータを受け取ることと、

前記第1のデータから第1の音響特徴を抽出することと、

第2のサウンドサンプルに関連付けられた第2のデータを受け取ることと、

前記第2のデータから第2の音響特徴を抽出することと、

前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定することと、

前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、決定結果を出力することと、

を備える、方法。

**【請求項2】**

前記決定結果は、サウンド検出モデルを生成する際に前記第2のサウンドサンプルが使用されない可能性があることを示す、請求項1に記載の方法。

**【請求項3】**

前記第1のデータに基づいて、前記第1のサウンドサンプルの信号対雑音比(SNR)を決定することと、

前記第2のデータに基づいて、前記第2のサウンドサンプルのSNRを決定することと、

前記第1のサウンドサンプルの前記SNRおよび前記第2のサウンドサンプルの前記SNRに基づいて、前記第1のデータまたは前記第2のデータのうちの少なくとも1つを選択することと、

前記第2の音響特徴が前記第1の音響特徴に類似すると決定すると、前記第1のデータまたは前記第2のデータのうちの前記選択された少なくとも1つに基づいて、前記サウンド検出モデルを生成することと、

前記第1のデータまたは前記第2のデータのうちの選択されていないものに基づいて、前記生成されたサウンド検出モデルを調整することと

をさらに備える、請求項1に記載の方法。

**【請求項4】**

代替のサウンドサンプルとして、第3のサウンドサンプルに関連付けられた第3のデータについての要求の指示を含む特定の決定結果を生成することと、

前記第3のデータを受け取ることと

をさらに備える、請求項1に記載の方法。

**【請求項5】**

前記第2の音響特徴が前記第1の音響特徴に類似すると決定することに応答して、前記第1の音響特徴および前記第2の音響特徴に基づいて、組み合わせられた音響特徴を決定することと、

第3のサウンドサンプルに関連付けられた第3のデータを受け取ることと、

前記第3のデータから第3の音響特徴を抽出することと、

前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定することと

をさらに備える、請求項1に記載の方法。

**【請求項 6】**

前記第1の音響特徴および前記第2の音響特徴の各々は、スペクトル特徴または時間領域特徴のうちの少なくとも1つを含み、前記決定することは、モバイル通信デバイスを備えるデバイスにおいて実行される、請求項1に記載の方法。

**【請求項 7】**

前記決定することは、マルチメディアデバイスを備えるデバイスにおいて実行される、請求項1に記載の方法。

**【請求項 8】**

前記第1の音響特徴および前記第2の音響特徴の各々は、サブワードのシーケンスを含み、前記サブワードは、音、音素、トライфон、または音節のうちの少なくとも1つを含む、請求項1に記載の方法。

**【請求項 9】**

前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルに関連付けられた追加のデータを受け取ることと、

前記追加のデータから新規の音響特徴を抽出することと、

前記新規の音響特徴が前記第1の音響特徴に類似するかどうかを決定することとをさらに備える、請求項1に記載の方法。

**【請求項 10】**

前記新規の音響特徴が前記第1の音響特徴に類似すると決定することに応答して、前記第1のデータまたは前記追加のデータのうちの少なくとも1つに基づいて、前記サウンド検出モデルを生成することをさらに備える、請求項9に記載の方法。

**【請求項 11】**

前記サウンド検出モデルを生成することは、

前記第1のデータまたは前記追加のデータのうちの少なくとも1つに基づいて、前記サウンド検出モデルの閾値を決定することと、

前記第2のデータに基づいて、前記閾値を調整することとを備える、請求項10に記載の方法。

**【請求項 12】**

前記第2の音響特徴が前記第1の音響特徴に類似すると決定することに応答して、前記第2のサウンドサンプルが前記第1のサウンドサンプルに類似するとの指示を出力することをさらに備える、請求項1に記載の方法。

**【請求項 13】**

前記第1のデータおよび前記第2のデータの各々は、前記電子デバイスをアクティブ化するためのコマンド、あるいは前記電子デバイスにおけるアプリケーションまたは機能を制御するためのコマンドのうちの少なくとも1つを示す、請求項1に記載の方法。

**【請求項 14】**

前記第2の音響特徴が前記第1の音響特徴に類似すると決定することに応答して、前記第1のデータまたは前記第2のデータのうちの少なくとも1つに基づいて、前記サウンド検出モデルを生成することと、

音声入力に関連付けられた入力データを受け取ることと、

前記サウンド検出モデルに基づいて、キーワードまたはユーザのうちの少なくとも1つを前記音声入力から認識することとをさらに備える、請求項1に記載の方法。

**【請求項 15】**

電子デバイスであって、

第1のサウンドサンプルに関連付けられた第1のデータを生成し、第2のサウンドサンプルに関連付けられた第2のデータを生成するように構成されたサウンドセンサと、

前記第1のデータから第1の音響特徴を抽出し、前記第2のデータから第2の音響特徴を抽出し、および前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定するように構成されたサウンドサンプル検証ユニットと、

前記第2の音響特徴が前記第1の音響特徴に類似しないとの前記サウンドサンプル検証ユニットによる決定に応答して、決定結果を出力するように構成された出力ユニットと、を備える、電子デバイス。

**【請求項16】**

アンテナと、  
前記アンテナに結合され、符号化されたオーディオ信号を受信するように構成された受信機と

をさらに備える、請求項15に記載の電子デバイス。

**【請求項17】**

前記受信機に結合された復調器をさらに備え、前記復調器は、前記符号化されたオーディオ信号を復調するように構成される、請求項16に記載の電子デバイス。

**【請求項18】**

前記復調器に結合されたプロセッサと、  
前記プロセッサに結合されたデコーダと、前記デコーダは、前記符号化されたオーディオ信号を復号するように構成され、  
ここにおいて、前記受信機、前記復調器、前記プロセッサ、前記デコーダ、前記サウンドセンサ、および前記サウンドサンプル検証ユニットは、モバイル通信デバイスに統合される、

をさらに備える、請求項17に記載の電子デバイス。

**【請求項19】**

前記サウンドセンサおよび前記サウンドサンプル検証ユニットは、アクセス端末に統合される、請求項15に記載の電子デバイス。

**【請求項20】**

前記第2の音響特徴が前記第1の音響特徴に類似するとの前記サウンドサンプル検証ユニットによる決定に応答して、前記第1のデータまたは前記第2のデータのうちの少なくとも1つに基づいて、サウンド検出モデルを生成するように構成されたサウンド検出モデル生成ユニットと、

前記第1のサウンドサンプルまたは前記第2のサウンドサンプルのうちの少なくとも1つを示す入力を受け取るよう構成された入力ユニットと

ここにおいて、前記サウンドサンプル検証ユニットは、前記第2の音響特徴が前記第1の音響特徴に類似するとの前記サウンドサンプル検証ユニットによる決定に応答して、前記第1の音響特徴および前記第2の音響特徴に基づいて、組み合わせられた音響特徴を決定し、前記サウンドセンサによって受け取られた第3のサウンドサンプルに関連付けられた第3のデータから第3の音響特徴を抽出し、および前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定するように構成され、

ここにおいて、前記出力ユニットは、前記第2の音響特徴が前記第1の音響特徴に類似するとの前記サウンドサンプル検証ユニットによる決定に応答して、前記第2のサウンドサンプルが前記第1のサウンドサンプルに類似するとの指示を出力するように構成される、

をさらに備える、請求項15に記載の電子デバイス。

**【請求項21】**

実行可能な命令を記憶するコンピュータ可読媒体であって、前記命令は、電子デバイスの少なくとも1つのプロセッサに、

第1のサウンドサンプルに関連付けられた第1のデータを受け取ることと  
前記第1のデータから第1の音響特徴を抽出することと  
第2のサウンドサンプルに関連付けられた第2のデータを受け取ることと  
前記第2のデータから第2の音響特徴を抽出することと  
前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定することと、  
前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、決定結果を出力すること、

を備える動作を実行させる、コンピュータ可読媒体。

【請求項 2 2】

前記動作は、前記第2の音響特徴が前記第1の音響特徴に類似すると決定することに応答して、前記第1のデータまたは前記第2のデータのうちの少なくとも1つに基づいて、サウンド検出モデルを生成することを備える、請求項2 1に記載のコンピュータ可読媒体。

【請求項 2 3】

前記動作は、前記第2の音響特徴が前記第1の音響特徴に類似すると決定することに応答して、前記第1の音響特徴および前記第2の音響特徴に基づいて、組み合わせられた音響特徴を決定することを備える、請求項2 1に記載のコンピュータ可読媒体。

【請求項 2 4】

前記動作は、

第3のサウンドサンプルに関連付けられた第3のデータを受け取ることと、

前記第3のデータから第3の音響特徴を抽出することと、

前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定することと

を備える、請求項2 3に記載のコンピュータ可読媒体。

【請求項 2 5】

前記動作は、

前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルに関連付けられた追加のデータを受け取ることと

前記追加のデータから新規の音響特徴を抽出することと、

前記新規の音響特徴が前記第1の音響特徴に類似するかどうかを決定することと

を備える、請求項2 1に記載のコンピュータ可読媒体。

【請求項 2 6】

サウンド検出モデルを生成する際に使用される少なくとも1つのサウンドサンプルを検証するための電子デバイスであって、

サウンドサンプルを受け取るための手段と、受け取るための前記手段は、第1のサウンドサンプルおよび第2のサウンドサンプルを受け取るように構成され、

音響特徴を抽出するための手段と、抽出するための前記手段は、前記第1のサウンドサンプルに関連付けられた第1のデータから第1の音響特徴を抽出し、前記第2のサウンドサンプルに関連付けられた第2のデータから第2の音響特徴を抽出するように構成され、

前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定するための手段と

前記第2の音響特徴が前記第1の音響特徴に類似しないと決定するための前記手段による決定に応答して、決定結果を出力するための手段と

を備える、電子デバイス。

【請求項 2 7】

受け取るための前記手段、抽出するための前記手段、決定するための前記手段、および出力するための前記手段は、モバイル通信デバイスに統合される、請求項2 6に記載の電子デバイス。

【請求項 2 8】

決定するための前記手段は、前記第2の音響特徴が前記第1の音響特徴に類似すると決定することに応答して、前記第1の音響特徴および前記第2の音響特徴に基づいて、組み合わせられた音響特徴を決定するように構成され、受け取るための前記手段は、第3のサウンドサンプルを受け取るように構成され、抽出するための前記手段は、前記第3のサウンドサンプルに関連付けられた第3のデータから第3の音響特徴を抽出するように構成され、決定するための前記手段は、前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定するように構成される、請求項2 6に記載の電子デバイス。

【請求項 2 9】

受け取るための前記手段、抽出するための前記手段、決定するための前記手段、および出力するための前記手段は、基地局に統合される、請求項 2 6 に記載の電子デバイス。

【請求項 3 0】

前記第 2 の音響特徴が前記第 1 の音響特徴に類似すると決定するための前記手段による決定に応答して、前記第 1 のデータまたは前記第 2 のデータのうちの少なくとも 1 つに基づいて、前記サウンド検出モデルを生成するための手段と、

前記第 2 の音響特徴が前記第 1 の音響特徴に類似しないと決定するための前記手段による決定に応答して、新規のサウンドサンプルを受け取るための手段と、

前記新規のサウンドサンプルに関連付けられた追加のデータから新規の音響特徴を抽出するための手段と、

前記新規の音響特徴が前記第 1 の音響特徴に類似するかどうかを決定するための手段とをさらに備える、請求項 2 6 に記載の電子デバイス。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0 1 1 7

【補正方法】変更

【補正の内容】

【0 1 1 7】

[00127] (例 3 0) 例 2 5 乃至 2 9 のうちのいずれか 1 つの電子デバイスは、第 2 の音響特徴が第 1 の音響特徴に類似するかどうかを決定することに基づいて、第 2 のサウンドサンプルが第 1 のサウンドサンプルに類似するかどうかの指示を出力するための手段と、サウンド検出モデルを生成する際に使用される第 1 のサウンドサンプルまたは第 2 のサウンドサンプルのうちの少なくとも 1 つを示す入力を受け取るための手段と、をさらに備える。

以下に本願の出願当初の特許請求の範囲に記載された発明を付記する。

[ C 1 ]

サウンド検出モデルを生成する際に使用される少なくとも 1 つのサウンドサンプルを検証するために、電子デバイスにおいて実行される方法であって、前記方法は、

第 1 のサウンドサンプルを受け取ることと、

前記第 1 のサウンドサンプルから第 1 の音響特徴を抽出することと、

第 2 のサウンドサンプルを受け取ることと、

前記第 2 のサウンドサンプルから第 2 の音響特徴を抽出することと、

前記第 2 の音響特徴が前記第 1 の音響特徴に類似するかどうかを決定することとを備える、方法。

[ C 2 ]

前記第 2 の音響特徴が前記第 1 の音響特徴に類似することを決定することに応答して、前記第 1 のサウンドサンプルまたは前記第 2 のサウンドサンプルのうちの少なくとも 1 つに基づいて、前記サウンド検出モデルを生成することをさらに備える、C 1 に記載の方法。

[ C 3 ]

前記第 1 のサウンドサンプルの信号対雑音比 ( S N R ) を決定することと、

前記第 2 のサウンドサンプルの S N R を決定することと、

前記第 1 および第 2 のサウンドサンプルの前記 S N R に基づいて、前記第 1 のサウンドサンプルまたは前記第 2 のサウンドサンプルのうちの少なくとも 1 つを選択することと、

前記第 2 の音響特徴が前記第 1 の音響特徴に類似することを決定すると、前記選択された少なくとも 1 つのサウンドサンプルに基づいて、前記サウンド検出モデルを生成することと、

選択されていないサウンドサンプルに基づいて、前記生成されたサウンド検出モデルを調整することと、

をさらに備える、C 1 に記載の方法。

[ C 4 ]

前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、前記第1の音響特徴および前記第2の音響特徴に基づいて、組み合わせられた音響特徴を決定することをさらに備える、C 1に記載の方法。

[ C 5 ]

第3のサウンドサンプルを受け取ることと、

前記第3のサウンドサンプルから第3の音響特徴を抽出することと、

前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定することと

をさらに備える、C 4に記載の方法。

[ C 6 ]

前記第1の音響特徴および前記第2の音響特徴の各々は、スペクトル特徴または時間領域特徴のうちの少なくとも1つを含む、C 1に記載の方法。

[ C 7 ]

前記第1の音響特徴および前記第2の音響特徴の各々は、サブワードのシーケンスを含む、C 1に記載の方法。

[ C 8 ]

前記サブワードは、音、音素、トライфон、または音節のうちの少なくとも1つを含む、C 7に記載の方法。

[ C 9 ]

前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルを受け取ることと、

前記新規のサウンドサンプルから新規の音響特徴を抽出することと、

前記新規の音響特徴が前記第1の音響特徴に類似するかどうかを決定することと

をさらに備える、C 1に記載の方法。

[ C 10 ]

前記新規の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、前記第1のサウンドサンプルまたは前記新規のサウンドサンプルのうちの少なくとも1つに基づいて、前記サウンド検出モデルを生成することをさらに備える、C 9に記載の方法。

[ C 11 ]

前記サウンド検出モデルを生成することは、

前記第1のサウンドサンプルまたは前記新規のサウンドサンプルのうちの少なくとも1つに基づいて、前記サウンド検出モデルの閾値を決定することと、

前記第2のサウンドサンプルに基づいて、前記閾値を調整することと

を備える、C 10に記載の方法。

[ C 12 ]

前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定することに基づいて、前記第2のサウンドサンプルが前記第1のサウンドサンプルに類似するかどうかの指示を出力することと、

前記サウンド検出モデルを生成する際に使用される前記第1のサウンドサンプルまたは前記第2のサウンドサンプルのうちの少なくとも1つを示す入力を受け取ることと

をさらに備える、C 1に記載の方法。

[ C 13 ]

前記第1のサウンドサンプルおよび前記第2のサウンドサンプルの各々は、前記電子デバイスをアクティブ化するためのコマンド、あるいは前記電子デバイスにおけるアプリケーションまたは機能を制御するためのコマンドのうちの少なくとも1つを示す、C 1に記載の方法。

[ C 14 ]

音声入力を受け取ることと、

前記サウンド検出モデルに基づいて、前記音声入力からキーワードまたはユーザのうちの少なくとも1つを認識することと  
をさらに備える、C 2に記載の方法。

[ C 1 5 ]

サウンド検出モデルを生成する際に使用される少なくとも1つのサウンドサンプルを検証するための電子デバイスであって、

第1のサウンドサンプルおよび第2のサウンドサンプルを受け取るように構成されたサウンドセンサと、

前記第1のサウンドサンプルから第1の音響特徴を抽出し、前記第2のサウンドサンプルから第2の音響特徴を抽出し、前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定するように構成されたサウンドサンプル検証ユニットと  
を備える、電子デバイス。

[ C 1 6 ]

前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、  
前記第1のサウンドサンプルまたは前記第2のサウンドサンプルのうちの少なくとも1つに基いて、前記サウンド検出モデルを生成するように構成されたサウンド検出モデル生成ユニットをさらに備える、C 1 5に記載の電子デバイス。

[ C 1 7 ]

前記サウンドサンプル検証ユニットは、前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、前記第1の音響特徴および前記第2の音響特徴に基いて、組み合わせられた音響特徴を決定するように構成される、C 1 5に記載の電子デバイス。

[ C 1 8 ]

前記サウンドセンサは、第3のサウンドサンプルを受け取るように構成され、

前記サウンドサンプル検証ユニットは、前記第3のサウンドサンプルから第3の音響特徴を抽出し、前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定するように構成される、

C 1 7に記載の電子デバイス。

[ C 1 9 ]

前記サウンドセンサは、前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルを受け取るように構成され、

前記サウンドサンプル検証ユニットは、前記新規のサウンドサンプルから新規の音響特徴を抽出し、前記新規の音響特徴が前記第1の音響特徴に類似するかどうかを決定するように構成される、

C 1 5に記載の電子デバイス。

[ C 2 0 ]

前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定することに基づいて、前記第2のサウンドサンプルが前記第1のサウンドサンプルに類似するかどうかの指示を出力するように構成された出力ユニットと、

前記サウンド検出モデルを生成する際に使用される前記第1のサウンドサンプルまたは前記第2のサウンドサンプルのうちの少なくとも1つを示す入力を受け取るための入力ユニットと

をさらに備える、C 1 5に記載の電子デバイス。

[ C 2 1 ]

命令を備える非一時的コンピュータ可読記憶媒体であって、前記命令は、電子デバイスの少なくとも1つのプロセッサに、

第1のサウンドサンプルを受け取り、

前記第1のサウンドサンプルから第1の音響特徴を抽出し、

第2のサウンドサンプルを受け取り、

前記第2のサウンドサンプルから第2の音響特徴を抽出し、

前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定する動作を実行させる、非一時的コンピュータ可読記憶媒体。

[ C 2 2 ]

前記電子デバイスの前記少なくとも1つのプロセッサに、前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、前記第1のサウンドサンプルまたは前記第2のサウンドサンプルのうちの少なくとも1つに基づいて、サウンド検出モデルを生成する動作を実行させる命令をさらに備える、C 2 1に記載の非一時的コンピュータ可読記憶媒体。

[ C 2 3 ]

前記電子デバイスの前記少なくとも1つのプロセッサに、前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、前記第1の音響特徴および前記第2の音響特徴に基づいて、組み合わせられた音響特徴を決定する動作を実行させる命令をさらに備える、C 2 1に記載の非一時的コンピュータ可読記憶媒体。

[ C 2 4 ]

前記電子デバイスの前記少なくとも1つのプロセッサに、  
第3のサウンドサンプルを受け取り、  
前記第3のサウンドサンプルから第3の音響特徴を抽出し、  
前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定する動作を実行させる命令をさらに備える、C 2 1に記載の非一時的コンピュータ可読記憶媒体。

[ C 2 5 ]

前記電子デバイスの前記少なくとも1つのプロセッサに、  
前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルを受け取り、  
前記新規のサウンドサンプルから新規の音響特徴を抽出し、  
前記新規の音響特徴が前記第1の音響特徴に類似するかどうかを決定する動作を実行させる命令をさらに備える、C 2 1に記載の非一時的コンピュータ可読記憶媒体。

[ C 2 6 ]

サウンド検出モデルを生成する際に使用される少なくとも1つのサウンドサンプルを検証するための電子デバイスであって、  
第1のサウンドサンプルを受け取るための手段と、  
前記第1のサウンドサンプルから第1の音響特徴を抽出するための手段と、  
第2のサウンドサンプルを受け取るための手段と、  
前記第2のサウンドサンプルから第2の音響特徴を抽出するための手段と、  
前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定するための手段とを備える、電子デバイス。

[ C 2 7 ]

前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、前記第1のサウンドサンプルまたは前記第2のサウンドサンプルのうちの少なくとも1つに基づいて、前記サウンド検出モデルを生成するための手段をさらに備える、C 2 1に記載の電子デバイス。

[ C 2 8 ]

前記第2の音響特徴が前記第1の音響特徴に類似することを決定することに応答して、前記第1の音響特徴および前記第2の音響特徴に基づいて、組み合わせられた音響特徴を決定するための手段をさらに備える、C 2 1に記載の電子デバイス。

[ C 2 9 ]

第3のサウンドサンプルを受け取るための手段と、  
前記第3のサウンドサンプルから第3の音響特徴を抽出するための手段と、  
前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定するた

めの手段と

をさらに備える、C 2 8 に記載の電子デバイス。

[ C 3 0 ]

前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルを受け取るための手段と、

前記新規のサウンドサンプルから新規の音響特徴を抽出するための手段と、

前記新規の音響特徴が前記第1の音響特徴に類似するかどうかを決定するための手段と、

をさらに備える、C 2 6 に記載の電子デバイス。

**【手続補正書】**

【提出日】平成29年7月5日(2017.7.5)

**【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】**

サウンド検出モデルを生成する際に使用されるデータを決定するために、電子デバイスにおいて実行される方法であって、前記方法は、

第1のサウンドサンプルに関連付けられた第1のデータを受け取ることと、

前記第1のデータから第1の音響特徴を抽出することと、

第2のサウンドサンプルに関連付けられた第2のデータを受け取ることと、

前記第2のデータから第2の音響特徴を抽出することと、

前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定することと、

前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、前記第2のサウンドサンプルが前記サウンド検出モデルを生成する際に使用されるかどうかをユーザに示す決定結果を出力すること、

を備える、方法。

**【請求項2】**

前記決定結果は、前記サウンド検出モデルを生成する際に前記第2のサウンドサンプルが使用されない可能性があることを前記ユーザに示す、請求項1に記載の方法。

**【請求項3】**

前記第1のデータに基づいて、前記第1のサウンドサンプルの信号対雑音比(SNR)を決定することと、

前記第2のデータに基づいて、前記第2のサウンドサンプルのSNRを決定することと、

前記第1のサウンドサンプルの前記SNRおよび前記第2のサウンドサンプルの前記SNRに基づいて、前記第1のデータまたは前記第2のデータのうちの少なくとも1つを選択することと、

前記第2の音響特徴が前記第1の音響特徴に類似すると決定すると、前記第1のデータまたは前記第2のデータのうちの前記選択された少なくとも1つに基づいて、前記サウンド検出モデルを生成することと、

前記第1のデータまたは前記第2のデータのうちの選択されていないものに基づいて、前記生成されたサウンド検出モデルを調整することと

をさらに備える、請求項1に記載の方法。

**【請求項4】**

前記決定結果は、

代替のサウンドサンプルについての前記ユーザへの要求

を含む、請求項1に記載の方法。

**【請求項5】**

前記第2の音響特徴が前記第1の音響特徴に類似すると決定することに応答して、前記第1の音響特徴および前記第2の音響特徴に基づいて、組み合わせられた音響特徴を決定することと、

第3のサウンドサンプルに関連付けられた第3のデータを受け取ることと、

前記第3のデータから第3の音響特徴を抽出することと、

前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定することと

をさらに備える、請求項1に記載の方法。

#### 【請求項6】

前記第1の音響特徴および前記第2の音響特徴の各々は、スペクトル特徴または時間領域特徴のうちの少なくとも1つを含み、前記決定することは、モバイル通信デバイスを備えるデバイスにおいて実行される、請求項1に記載の方法。

#### 【請求項7】

前記決定結果は、前記第1のサウンドサンプルと前記第2のサウンドサンプルとが類似しないことを前記ユーザに示し、

前記決定することは、マルチメディアデバイスを備えるデバイスにおいて実行される、請求項1に記載の方法。

#### 【請求項8】

前記第1の音響特徴および前記第2の音響特徴の各々は、サブワードのシーケンスを含み、前記サブワードは、音、音素、トライフォン、または音節のうちの少なくとも1つを含む、請求項1に記載の方法。

#### 【請求項9】

前記第2の音響特徴が前記第1の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルに関連付けられた追加のデータを受け取ることと、

前記追加のデータから新規の音響特徴を抽出することと、

前記新規の音響特徴が前記第1の音響特徴に類似するかどうかを決定することとをさらに備える、請求項1に記載の方法。

#### 【請求項10】

前記新規の音響特徴が前記第1の音響特徴に類似すると決定することに応答して、前記第1のデータまたは前記追加のデータのうちの少なくとも1つに基づいて、前記サウンド検出モデルを生成することをさらに備える、請求項9に記載の方法。

#### 【請求項11】

前記サウンド検出モデルを生成することは、

前記第1のデータまたは前記追加のデータのうちの少なくとも1つに基づいて、前記サウンド検出モデルの閾値を決定することと、

前記第2のデータに基づいて、前記閾値を調整することと

を備える、請求項10に記載の方法。

#### 【請求項12】

前記第2の音響特徴が前記第1の音響特徴に類似すると決定することに応答して、前記第2のサウンドサンプルが前記第1のサウンドサンプルに類似するとの指示を表示することをさらに備える、請求項1に記載の方法。

#### 【請求項13】

前記第1のデータおよび前記第2のデータの各々は、前記電子デバイスをアクティブ化するためのコマンド、あるいは前記電子デバイスにおけるアプリケーションまたは機能を制御するためのコマンドのうちの少なくとも1つを示す、請求項1に記載の方法。

#### 【請求項14】

前記第2の音響特徴が前記第1の音響特徴に類似すると決定することに応答して、前記第1のデータまたは前記第2のデータのうちの少なくとも1つに基づいて、前記サウンド検出モデルを生成することと、

音声入力に関連付けられた入力データを受け取ることと、

前記サウンド検出モデルに基づいて、キーワードまたは前記ユーザのうちの少なくとも1つを前記音声入力から認識することと  
をさらに備える、請求項1に記載の方法。

【請求項15】

電子デバイスであって、

第1のサウンドサンプルに関連付けられた第1のデータを生成し、第2のサウンドサンプルに関連付けられた第2のデータを生成するように構成されたサウンドセンサと、

前記第1のデータから第1の音響特徴を抽出し、前記第2のデータから第2の音響特徴を抽出し、および前記第2の音響特徴が前記第1の音響特徴に類似するかどうかを決定するように構成されたサウンドサンプル検証ユニットと、

前記第2の音響特徴が前記第1の音響特徴に類似しないとの前記サウンドサンプル検証ユニットによる決定に応答して、決定結果を出力するように構成された出力ユニットと、前記決定結果は、前記第2のサウンドサンプルがサウンド検出モデルを生成する際に使用されるかどうかをユーザに示す、

を備える、電子デバイス。

【請求項16】

アンテナと、

前記アンテナに結合され、符号化されたオーディオ信号を受信するように構成された受信機と

をさらに備える、請求項15に記載の電子デバイス。

【請求項17】

前記受信機に結合された復調器をさらに備え、前記復調器は、前記符号化されたオーディオ信号を復調するように構成される、請求項16に記載の電子デバイス。

【請求項18】

前記復調器に結合されたプロセッサと、

前記プロセッサに結合されたデコーダと、前記デコーダは、前記符号化されたオーディオ信号を復号するように構成され、

ここにおいて、前記受信機、前記復調器、前記プロセッサ、前記デコーダ、前記サウンドセンサ、および前記サウンドサンプル検証ユニットは、モバイル通信デバイスに統合される、

をさらに備える、請求項17に記載の電子デバイス。

【請求項19】

前記サウンドセンサおよび前記サウンドサンプル検証ユニットは、マルチメディアデバイスに統合される、請求項15に記載の電子デバイス。

【請求項20】

前記第2の音響特徴が前記第1の音響特徴に類似するとの前記サウンドサンプル検証ユニットによる決定に応答して、前記第1のデータまたは前記第2のデータのうちの少なくとも1つに基づいて、前記サウンド検出モデルを生成するように構成されたサウンド検出モデル生成ユニットと、

前記第1のサウンドサンプルまたは前記第2のサウンドサンプルのうちの少なくとも1つを示す入力を受け取るように構成された入力ユニットと

ここにおいて、前記サウンドサンプル検証ユニットは、前記第2の音響特徴が前記第1の音響特徴に類似するとの前記サウンドサンプル検証ユニットによる決定に応答して、前記第1の音響特徴および前記第2の音響特徴に基づいて、組み合わせられた音響特徴を決定し、前記サウンドセンサによって受け取られた第3のサウンドサンプルに関連付けられた第3のデータから第3の音響特徴を抽出し、および前記第3の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定するように構成され、

ここにおいて、前記出力ユニットは、前記第2の音響特徴が前記第1の音響特徴に類似するとの前記サウンドサンプル検証ユニットによる決定に応答して、前記第2のサウンドサンプルが前記第1のサウンドサンプルに類似するとのユーザへの指示を出力するように

構成される、

をさらに備える、請求項 1 5 に記載の電子デバイス。

【請求項 2 1】

実行可能な命令を記憶するコンピュータ可読媒体であって、前記命令は、電子デバイスの少なくとも 1 つのプロセッサに、

第 1 のサウンドサンプルに関連付けられた第 1 のデータを受け取ることと

前記第 1 のデータから第 1 の音響特徴を抽出することと

第 2 のサウンドサンプルに関連付けられた第 2 のデータを受け取ることと

前記第 2 のデータから第 2 の音響特徴を抽出することと

前記第 2 の音響特徴が前記第 1 の音響特徴に類似するかどうかを決定することと、

前記第 2 の音響特徴が前記第 1 の音響特徴に類似しないと決定することに応答して、前記第 2 のサウンドサンプルがサウンド検出モデルを生成する際に使用されるかどうかをユーザに示す決定結果を出力することと、

を備える動作を実行させる、コンピュータ可読媒体。

【請求項 2 2】

前記動作は、前記第 2 の音響特徴が前記第 1 の音響特徴に類似すると決定することに応答して、前記第 1 のデータまたは前記第 2 のデータのうちの少なくとも 1 つに基づいて、前記サウンド検出モデルを生成することを備える、請求項 2 1 に記載のコンピュータ可読媒体。

【請求項 2 3】

前記動作は、前記第 2 の音響特徴が前記第 1 の音響特徴に類似すると決定することに応答して、前記第 1 の音響特徴および前記第 2 の音響特徴に基づいて、組み合わせられた音響特徴を決定することを備える、請求項 2 1 に記載のコンピュータ可読媒体。

【請求項 2 4】

前記動作は、

前記第 2 のサウンドサンプルが前記第 1 のサウンドサンプルに類似しないとの指示を表示した後、第 3 のサウンドサンプルに関連付けられた第 3 のデータを受け取ることと、

前記第 3 のデータから第 3 の音響特徴を抽出することと、

前記第 3 の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定することと

を備える、請求項 2 3 に記載のコンピュータ可読媒体。

【請求項 2 5】

前記動作は、

前記第 2 の音響特徴が前記第 1 の音響特徴に類似しないと決定することに応答して、新規のサウンドサンプルに関連付けられた追加のデータを受け取ることと

前記追加のデータから新規の音響特徴を抽出することと、

前記新規の音響特徴が前記第 1 の音響特徴に類似するかどうかを決定することと

を備える、請求項 2 1 に記載のコンピュータ可読媒体。

【請求項 2 6】

サウンド検出モデルを生成する際に使用される少なくとも 1 つのサウンドサンプルを検証するための電子デバイスであって、

サウンドサンプルを受け取るための手段と、受け取るための前記手段は、第 1 のサウンドサンプルおよび第 2 のサウンドサンプルを受け取るように構成され、

音響特徴を抽出するための手段と、抽出するための前記手段は、前記第 1 のサウンドサンプルに関連付けられた第 1 のデータから第 1 の音響特徴を抽出し、前記第 2 のサウンドサンプルに関連付けられた第 2 のデータから第 2 の音響特徴を抽出するように構成され、

前記第 2 の音響特徴が前記第 1 の音響特徴に類似するかどうかを決定するための手段と、

前記第 2 の音響特徴が前記第 1 の音響特徴に類似しないと決定するための前記手段による決定に応答して、決定結果を出力するための手段と、前記決定結果は、前記第 2 のサウ

ンドサンプルが前記サウンド検出モデルを生成する際に使用されるかどうかをユーザに示す

を備える、電子デバイス。

【請求項 27】

受け取るための前記手段、抽出するための前記手段、決定するための前記手段、および出力するための前記手段は、モバイル通信デバイスに統合される、請求項 26 に記載の電子デバイス。

【請求項 28】

決定するための前記手段は、前記第 2 の音響特徴が前記第 1 の音響特徴に類似すると決定することに応答して、前記第 1 の音響特徴および前記第 2 の音響特徴に基づいて、組み合わせられた音響特徴を決定するように構成され、受け取るための前記手段は、第 3 のサウンドサンプルを受け取るように構成され、抽出するための前記手段は、前記第 3 のサウンドサンプルに関連付けられた第 3 のデータから第 3 の音響特徴を抽出するように構成され、決定するための前記手段は、前記第 3 の音響特徴が前記組み合わせられた音響特徴に類似するかどうかを決定するように構成される、請求項 26 に記載の電子デバイス。

【請求項 29】

受け取るための前記手段、抽出するための前記手段、決定するための前記手段、および出力するための前記手段は、マルチメディアデバイスに統合される、請求項 26 に記載の電子デバイス。

【請求項 30】

前記第 2 の音響特徴が前記第 1 の音響特徴に類似すると決定するための前記手段による決定に応答して、前記第 1 のデータまたは前記第 2 のデータのうちの少なくとも 1 つに基づいて、前記サウンド検出モデルを生成するための手段と、

前記第 2 の音響特徴が前記第 1 の音響特徴に類似しないと決定するための前記手段による決定に応答して、新規のサウンドサンプルを受け取るための手段と、

前記新規のサウンドサンプルに関連付けられた追加のデータから新規の音響特徴を抽出するための手段と、

前記新規の音響特徴が前記第 1 の音響特徴に類似するかどうかを決定するための手段とをさらに備える、請求項 26 に記載の電子デバイス。

## 【国際調査報告】

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2015/053665

A. CLASSIFICATION OF SUBJECT MATTER  
INV. G10L15/06  
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 02/21509 A1 (SNAP ON TECH INC [US]) 14 March 2002 (2002-03-14) figures 1, 2 paragraph bridging pages 11 and 12 -----	1-30
X	US 6 134 527 A (MEUNIER JEFFREY ARTHUR [US] ET AL) 17 October 2000 (2000-10-17) figure 1 column 3, line 54 - column 4, line 47 -----	1-30
A	US 4 994 983 A (LANDELL BLAKELY P [US] ET AL) 19 February 1991 (1991-02-19) abstract figure 2 -----	1-30



Further documents are listed in the continuation of Box C.



See patent family annex.

## \* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

16 December 2015

04/01/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

Müller, Achim

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2015/053665

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0221509	A1 14-03-2002	AU 7917201 A CN 1451152 A EP 1314151 A1 JP 2004508593 A TW 571290 B US 6556971 B1 WO 0221509 A1	22-03-2002 22-10-2003 28-05-2003 18-03-2004 11-01-2004 29-04-2003 14-03-2002
US 6134527	A 17-10-2000	NONE	
US 4994983	A 19-02-1991	NONE	

---

フロントページの続き

(81)指定国 AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,R0,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,JP,KE,KG,KN,KP,KR,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US

(72)発明者 ムン、ソンクク

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

(72)発明者 ジン、ミンホ

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

(72)発明者 シャ、ハイイン

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

(72)発明者 フアン、ヘス

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

(72)発明者 デール、ウォーレン・フレデリック

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775