

(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2015-508205

(P2015-508205A)

(43) 公表日 平成27年3月16日 (2015.3.16)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 17/30 (2006.01)	G06F 17/30	3 1 0 Z 5 K 1 2 7
H04M 1/00 (2006.01)	H04M 1/00	R 5 K 2 0 1
H04M 11/00 (2006.01)	H04M 11/00	3 0 2

審査請求 未請求 予備審査請求 有 (全 49 頁)

(21) 出願番号	特願2014-556569 (P2014-556569)	(71) 出願人	595020643
(86) (22) 出願日	平成25年1月29日 (2013.1.29)		クアルコム・インコーポレイテッド
(85) 翻訳文提出日	平成26年9月18日 (2014.9.18)		QUALCOMM INCORPORATED
(86) 国際出願番号	PCT/US2013/023607		ED
(87) 国際公開番号	W02013/119418		アメリカ合衆国、カリフォルニア州 92
(87) 国際公開日	平成25年8月15日 (2013.8.15)		121-1714、サン・ディエゴ、モア
(31) 優先権主張番号	61/596,576		ハウス・ドライブ 5775
(32) 優先日	平成24年2月8日 (2012.2.8)	(74) 代理人	100108855
(33) 優先権主張国	米国 (US)		弁理士 蔵田 昌俊
(31) 優先権主張番号	13/752,080	(74) 代理人	100109830
(32) 優先日	平成25年1月28日 (2013.1.28)		弁理士 福原 淑弘
(33) 優先権主張国	米国 (US)	(74) 代理人	100103034
			弁理士 野河 信久
		(74) 代理人	100075672
			弁理士 峰 隆司

最終頁に続く

(54) 【発明の名称】 音識別に基づくモバイルデバイスの制御

(57) 【要約】

モバイルデバイスにおいて機能を実行するための方法および装置が開示される。モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャし、キャプチャされたメディア音から音特徴を抽出する。抽出された音特徴に基づいて基準音特徴のセット中の少なくとも1つの基準音特徴を識別することによって、モバイルデバイスにおいて実行されるべき機能を決定し、基準音特徴のセット中の各基準音特徴は、複数のメディア音のうちの少なくとも1つと、複数の機能のうちの少なくとも1つとに関連する。さらに、モバイルデバイスにおいて、決定された機能を実行する。

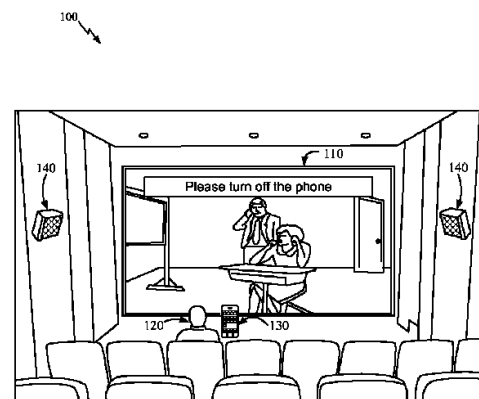


FIG. 1

【特許請求の範囲】**【請求項 1】**

モバイルデバイスにおいて機能を実行するための方法であって、
前記モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャすることと、
前記キャプチャされたメディア音から音特徴を抽出することと、
前記抽出された音特徴に基づいて、基準音特徴のセット中の少なくとも 1 つの基準音特徴を識別することによって、前記モバイルデバイスにおいて実行されるべき前記機能を決定することと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも 1 つと、複数の機能のうちの少なくとも 1 つとに関連し、
前記モバイルデバイスにおいて前記決定された機能を実行することと、
を備える、方法。

10

【請求項 2】

前記少なくとも 1 つの基準音特徴は広告に関連し、前記決定された機能を実行することは、前記モバイルデバイス上に前記広告を表示することを含む、請求項 1 に記載の方法。

【請求項 3】

前記少なくとも 1 つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記決定された機能を実行することは、
前記キーワードを前記検索エンジンに送信することと、
前記検索エンジンから検索結果を受信することと、
前記モバイルデバイス上に前記検索結果を出力することと、
を含む、請求項 1 に記載の方法。

20

【請求項 4】

前記決定された機能を実行することは、前記モバイルデバイスの構成を調整することを含む、請求項 1 に記載の方法。

【請求項 5】

前記決定された機能を実行することは、
実行されるべき前記機能の通知を出力することと、
前記機能を実行するための入力を受信することと、
前記受信された入力に応答して前記機能を実行することと、
を含む、請求項 1 に記載の方法。

30

【請求項 6】

前記キャプチャされたメディア音の音レベルが所定の音レベルよりも大きい場合、前記キャプチャされたメディア音から前記音特徴を抽出することと、前記モバイルデバイスによって実行されるべき前記機能を決定することとが実行される、請求項 1 に記載の方法。

【請求項 7】

前記キャプチャされたメディア音から前記音特徴を抽出することは、
前記キャプチャされたメディア音から中間音特徴を抽出することと、
前記音特徴を生成するために、前記中間音特徴に少なくとも 1 つの基礎フィルタを適用することと、
を含む、請求項 1 に記載の方法。

40

【請求項 8】

前記少なくとも 1 つの基礎フィルタは、複数の基礎フィルタから、前記複数の基礎フィルタの各々によって 1 ペアのサンプル音から抽出された音特徴間の類似度の程度に基づいて、選択される、請求項 7 に記載の方法。

【請求項 9】

前記少なくとも 1 つの基準音特徴は、前記モバイルデバイス上で実行されるべき前記機能とコンテキスト情報とを含むメタデータに関連する、請求項 1 に記載の方法。

【請求項 10】

モバイルデバイスにおいて機能を実行するための方法であって、

50

複数の音センサーによって、少なくとも１つのメディア音を含む合成音をキャプチャすることと、

前記合成音から前記少なくとも１つのメディア音を分離することと、

前記少なくとも１つの分離されたメディア音から少なくとも１つの音特徴を抽出することと、

前記少なくとも１つの抽出された音特徴に基づいて、基準音特徴のセット中の少なくとも１つの基準音特徴を識別することによって、前記モバイルデバイスにおいて実行されるべき前記機能を決定することと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも１つと、複数の機能のうちの少なくとも１つとに関連し、

前記モバイルデバイスにおいて前記決定された機能を実行することと、

を備える、方法。

10

【請求項 1 1】

前記少なくとも１つの基準音特徴は広告に関連し、前記決定された機能を実行することは、前記モバイルデバイス上に前記広告を表示することを含む、請求項 1 0 に記載の方法。

【請求項 1 2】

前記少なくとも１つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記決定された機能を実行することは、

前記キーワードを前記検索エンジンに送信することと、

前記検索エンジンから検索結果を受信することと、

前記モバイルデバイス上に前記検索結果を出力することと、

を含む、請求項 1 0 に記載の方法。

20

【請求項 1 3】

前記決定された機能を実行することは、前記モバイルデバイスの構成を調整することを含む、請求項 1 0 に記載の方法。

【請求項 1 4】

前記決定された機能を実行することは、

実行されるべき前記機能の通知を出力することと、

前記機能を実行するための入力を受信することと、

前記受信された入力に応答して前記機能を実行することと、

を含む、請求項 1 0 に記載の方法。

30

【請求項 1 5】

前記少なくとも１つの分離されたメディア音から前記少なくとも１つの音特徴を抽出することは、

前記少なくとも１つの分離されたメディア音から１つのメディア音を選択することと、

前記選択されたメディア音から音特徴を抽出することと、

を含む、請求項 1 0 に記載の方法。

【請求項 1 6】

前記少なくとも１つの分離されたメディア音から前記少なくとも１つの音特徴を抽出することは、すべての分離されたメディア音から音特徴を抽出することを含む、請求項 1 0 に記載の方法。

40

【請求項 1 7】

モバイルデバイスによって実行されるべき機能を決定するためのサーバであって、

音特徴を受信するように構成された受信機と、前記音特徴は、前記モバイルデバイスによってキャプチャされたメディア音から抽出され、

前記音特徴に基づいて、基準音特徴のセット中の少なくとも１つの基準音特徴を識別するように構成された識別ユニットと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも１つと、複数の機能のうちの少なくとも１つとに関連し、

前記少なくとも１つの識別された基準音特徴に基づいて前記モバイルデバイスによって

50

実行されるべき前記機能を決定するように構成された機能決定ユニットと、
を備える、サーバ。

【請求項 18】

前記少なくとも 1 つの基準音特徴は広告に関連し、前記機能決定ユニットは、前記モバイルデバイス上に前記広告を表示するような前記機能を決定するように構成された、請求項 17 に記載のサーバ。

【請求項 19】

前記少なくとも 1 つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記機能決定ユニットは、

前記キーワードを前記検索エンジンに送信し、

前記検索エンジンから検索結果を受信し、

前記モバイルデバイス上に前記検索結果を出力するような

前記機能を決定するように構成された、請求項 17 に記載のサーバ。

【請求項 20】

前記機能決定ユニットは、前記モバイルデバイスの構成を調整するような前記機能を決定するように構成された、請求項 17 に記載のサーバ。

【請求項 21】

前記機能決定ユニットは、

実行されるべき前記機能の通知を出力し、

前記機能を実行するための入力を受信し、

前記受信された入力に応答して前記機能を実行するような

前記機能を決定するように構成された、請求項 17 に記載のサーバ。

【請求項 22】

前記音特徴は、前記キャプチャされたメディア音の音レベルが所定の音レベルよりも大きい場合、前記キャプチャされたメディア音から抽出される、請求項 17 に記載のサーバ。

【請求項 23】

前記音特徴は、前記キャプチャされたメディア音から中間音特徴を抽出することと、前記音特徴を生成するために前記中間音特徴に少なくとも 1 つの基礎フィルタを適用することによって、前記キャプチャされたメディア音から抽出される、請求項 17 に記載のサーバ。

【請求項 24】

複数の基礎フィルタから、前記複数の基礎フィルタの各々によって 1 ペアのサンプル音から抽出された音特徴間の類似度の程度に基づいて、前記少なくとも 1 つの基礎フィルタを選択するように構成された基礎フィルタ生成器をさらに備える、請求項 23 に記載のサーバ。

【請求項 25】

前記少なくとも 1 つの基準音特徴は、前記モバイルデバイスによって実行されるべき前記機能とコンテキスト情報とを含むメタデータに関連する、請求項 17 に記載のサーバ。

【請求項 26】

モバイルデバイスであって、

前記モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャするように構成された音センサーと、

前記キャプチャされたメディア音から音特徴を抽出するように構成された音特徴抽出器と、

前記音特徴に基づいて基準音特徴のセット中の少なくとも 1 つの基準音特徴を識別するように構成された識別ユニットと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも 1 つと、複数の機能のうちの少なくとも 1 つとに関連し、

前記少なくとも 1 つの識別された基準音特徴に基づいて前記モバイルデバイスによって実行されるべき機能を決定するように構成された機能決定ユニットと、

10

20

30

40

50

前記モバイルデバイスにおいて前記決定された機能を実行するように構成された制御ユニットと、

を備える、モバイルデバイス。

【請求項 27】

前記少なくとも 1 つの基準音特徴は広告に関連し、前記制御ユニットは、前記モバイルデバイス上に前記広告を表示するように構成された、請求項 26 に記載のモバイルデバイス。

【請求項 28】

前記少なくとも 1 つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記制御ユニットは、

前記キーワードを前記検索エンジンに送信することと、

前記検索エンジンから検索結果を受信することと、

前記モバイルデバイス上に前記検索結果を出力することと、

を行うように構成された、請求項 26 に記載のモバイルデバイス。

【請求項 29】

前記制御ユニットは、前記モバイルデバイスの構成を調整するように構成された、請求項 26 に記載のモバイルデバイス。

【請求項 30】

前記制御ユニットは、

実行されるべき前記機能の通知を出力することと、

前記機能を実行するための入力を受信することと、

前記受信された入力に応答して前記機能を実行することと、

を行うように構成された、請求項 26 に記載のモバイルデバイス。

【請求項 31】

前記音センサーは、所定の時間間隔で環境音を受信することと、前記受信された環境音が前記メディア音であるかどうかを決定することによって、前記メディア音をキャプチャするようにさらに構成された、請求項 26 に記載のモバイルデバイス。

【請求項 32】

モバイルデバイスにおいて機能を実行するための命令を記憶す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、

前記モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャすることと、

前記キャプチャされたメディア音から音特徴を抽出することと、

前記抽出された音特徴に基づいて基準音特徴のセット中の少なくとも 1 つの基準音特徴を識別することによって、前記モバイルデバイスにおいて実行されるべき前記機能を決定することと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも 1 つと、複数の機能のうちの少なくとも 1 つとに関連し、

前記モバイルデバイスにおいて前記決定された機能を実行することと、

の動作をプロセッサに実行させる、非一時的コンピュータ可読記憶媒体。

【請求項 33】

前記少なくとも 1 つの基準音特徴は広告に関連し、前記決定された機能を実行することは、前記モバイルデバイス上に前記広告を表示することを含む、請求項 32 に記載の媒体。

【請求項 34】

前記少なくとも 1 つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記決定された機能を実行することは、

前記キーワードを前記検索エンジンに送信することと、

前記検索エンジンから検索結果を受信することと、

前記モバイルデバイス上に前記検索結果を出力することと、

を含む、請求項 32 に記載の媒体。

10

20

30

40

50

【請求項 3 5】

前記決定された機能を実行することは、前記モバイルデバイスの構成を調整することを含む、請求項 3 2 に記載の媒体。

【請求項 3 6】

前記決定された機能を実行することは、
実行されるべき前記機能の通知を出力することと、
前記機能を実行するための入力を受信することと、
前記受信された入力に応答して前記機能を実行することと、
を含む、請求項 3 2 に記載の媒体。

【請求項 3 7】

前記キャプチャされたメディア音の音レベルが所定の音レベルよりも大きい場合、前記キャプチャされたメディア音から前記音特徴を抽出することと、前記モバイルデバイスによって実行されるべき前記機能を決定することとが実行される、請求項 3 2 に記載の媒体。

10

【請求項 3 8】

前記キャプチャされたメディア音から前記音特徴を抽出することは、
前記キャプチャされたメディア音から中間音特徴を抽出することと、
前記音特徴を生成するために前記中間音特徴に少なくとも 1 つの基礎フィルタを適用することと、
を含む、請求項 3 2 に記載の媒体。

20

【請求項 3 9】

前記少なくとも 1 つの基礎フィルタは、複数の基礎フィルタから、前記複数の基礎フィルタの各々によって 1 ペアのサンプル音から抽出された音特徴間の類似度の程度に基づいて、選択される、請求項 3 8 に記載の媒体。

【請求項 4 0】

前記少なくとも 1 つの基準音特徴は、前記モバイルデバイスにおいて実行されるべき前記機能とコンテキスト情報とを含むメタデータに関連する、請求項 3 2 に記載の媒体。

【請求項 4 1】

モバイルデバイスによって実行されるべき機能を決定するためのサーバであって、
音特徴を受信する手段と、前記音特徴は、前記モバイルデバイスによってキャプチャされたメディア音から抽出され、
前記音特徴に基づいて、基準音特徴のセット中の少なくとも 1 つの基準音特徴を識別する手段と、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも 1 つと、複数の機能のうちの少なくとも 1 つとに関連し、
前記少なくとも 1 つの識別された基準音特徴に基づいて前記モバイルデバイスによって実行されるべき前記機能を決定する手段と、
を備える、サーバ。

30

【請求項 4 2】

前記少なくとも 1 つの基準音特徴は広告に関連し、前記決定する手段は、前記モバイルデバイス上に前記広告を表示するような前記機能を決定するように構成された、請求項 4 1 に記載のサーバ。

40

【請求項 4 3】

前記少なくとも 1 つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記決定する手段は、
前記キーワードを前記検索エンジンに送信し、
前記検索エンジンから検索結果を受信し、
前記モバイルデバイス上に前記検索結果を出力するような
前記機能を決定するように構成された、請求項 4 1 に記載のサーバ。

【請求項 4 4】

前記決定する手段は、前記モバイルデバイスの構成を調整するような前記機能を決定す

50

るように構成された、請求項 4 1 に記載のサーバ。

【請求項 4 5】

前記決定する手段は、
実行されるべき前記機能の通知を出力し、
前記機能を実行するための入力を受信し、
前記受信された入力に応答して前記機能を実行するような
前記機能を決定するように構成された、請求項 4 1 に記載のサーバ。

【請求項 4 6】

前記音特徴は、前記キャプチャされたメディア音の音レベルが所定の音レベルよりも大きい場合、前記キャプチャされたメディア音から抽出される、請求項 4 1 に記載のサーバ。 10

【請求項 4 7】

前記音特徴は、前記キャプチャされたメディア音から中間音特徴を抽出することと、前記音特徴を生成するために前記中間音特徴に少なくとも 1 つの基礎フィルタを適用することによって、前記キャプチャされたメディア音から抽出される、請求項 4 1 に記載のサーバ。

【請求項 4 8】

複数の基礎フィルタから、前記複数の基礎フィルタの各々によって 1 ペアのサンプル音から抽出された音特徴間の類似度の程度に基づいて、前記少なくとも 1 つの基礎フィルタを選択する手段をさらに備える、請求項 4 7 に記載のサーバ。 20

【請求項 4 9】

前記少なくとも 1 つの基準音特徴は、前記モバイルデバイスによって実行されるべき前記機能とコンテキスト情報とを含むメタデータに関連する、請求項 4 1 に記載のサーバ。

【請求項 5 0】

モバイルデバイスであって、
前記モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャする手段と、
前記キャプチャされたメディア音から音特徴を抽出する手段と、
前記音特徴に基づいて、基準音特徴のセット中の少なくとも 1 つの基準音特徴を識別する手段と、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも 1 つと、複数の機能のうちの少なくとも 1 つとに関連し、
前記少なくとも 1 つの識別された基準音特徴に基づいて前記モバイルデバイスによって実行されるべき機能を決定する手段と、
前記モバイルデバイスにおいて前記決定された機能を実行する手段と、
を備える、モバイルデバイス。 30

【請求項 5 1】

前記少なくとも 1 つの基準音特徴は広告に関連し、前記実行する手段は、前記モバイルデバイス上に前記広告を表示するように構成された、請求項 5 0 に記載のモバイルデバイス。

【請求項 5 2】

前記少なくとも 1 つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記実行する手段は、
前記キーワードを前記検索エンジンに送信することと、
前記検索エンジンから検索結果を受信することと、
前記モバイルデバイス上に前記検索結果を出力することと
を行うように構成された、請求項 5 0 に記載のモバイルデバイス。 40

【請求項 5 3】

前記実行する手段は、前記モバイルデバイスの構成を調整するように構成された、請求項 5 0 に記載のモバイルデバイス。

【請求項 5 4】

前記実行する手段は、
実行されるべき前記機能の通知を出力することと、
前記機能を実行するための入力を受信することと、
前記受信された入力に応答して前記機能を実行することと、
を行うように構成された、請求項 50 に記載のモバイルデバイス。

【発明の詳細な説明】

【優先権の主張】

【0001】

本特許出願は、各々が本出願の譲受人（出願人）に譲渡され、参照により本明細書に明確に組み込まれる、2012年2月8日に提出された米国仮出願第61/596,576号および2013年1月28日に提出された米国非仮出願第13/752,080号の優先権を主張する。

10

【技術分野】

【0002】

本開示は、一般に、ポータブルコンピューティングデバイス上の機能を制御することに関する。より詳細には、本開示は、音識別に基づいてモバイルデバイスを制御することに関する。

【背景技術】

【0003】

[0003]近年、スマートフォンおよびタブレットコンピュータなど、モバイルデバイスの使用が普及してきた。これらのデバイスは、一般に、ワイヤレスネットワークを介して音声およびデータ通信機能を与える。さらに、そのようなモバイルデバイスは、ユーザが、インターネットをブラウズすること、写真またはビデオを撮ること、電話呼またはビデオ呼を行うことなど、様々な機能を実行することを可能にする。したがって、たいていのユーザは、便利な機能のために、そのようなモバイルデバイスを常に携帯している。

20

【0004】

[0004]しかしながら、モバイルデバイスは、しばしば、望ましくない状況で動作させられる。たとえば、モバイルフォンのユーザは、しばしば、映画館などのいくつかの状況ではモバイルフォンの電源を切ることを期待される。最善の意図にもかかわらず、何人かのユーザは、電話の電源を切るのを忘れることがあり、近くの観客を迷惑がられる。

30

【0005】

[0005]さらに、モバイルデバイスユーザは、日常生活において様々な情報を受ける。ユーザが、興味の対象に関するより多くの情報を見つけ出すことを望むとき、ユーザは、一般に、自分のモバイルデバイス上でクエリを手動で入力することによって情報の検索を実行する。運転中のユーザにとって、これはあまり便利な作業でないことがある。

【0006】

[0006]したがって、ユーザによる手入力なしにモバイルデバイスにおいて機能を自動的に実行することができる方法および装置が必要である。

【発明の概要】

【0007】

[0007]本開示は、音識別に基づいてモバイルデバイス上で機能を制御するための方法、システムおよびデバイスを提供する。

40

【0008】

[0008]本開示の一態様によれば、モバイルデバイスを制御するための方法が開示される。本方法では、モバイルデバイスの外部の音出力デバイスからのメディア音がキャプチャされ、キャプチャされたメディア音から音特徴が抽出される。抽出された音特徴に基づいて基準音特徴のセット中の少なくとも1つの基準音特徴を識別することによって、モバイルデバイスによって実行されるべき機能が決定され、基準音特徴のセット中の各基準音特徴は、複数のメディア音のうちの少なくとも1つと、複数の機能のうちの少なくとも1つとに関連する。当該決定された機能は、モバイルデバイスにおいて実行される。さらに、

50

本方法は、音特徴を生成するために少なくとも１つの基礎フィルタ（basis filter）を適用することを含み得る。本開示は、本方法に係るデバイス、システム、サーバ、手段の組合せ、およびコンピュータ可読媒体についても説明する。

【０００９】

【0009】本開示の別の態様によれば、モバイルデバイスにおいて機能を実行するための方法が開示される。本方法では、少なくとも１つのメディア音を含む合成音（composite sound）は、複数の音センサーによってキャプチャされる。少なくとも１つのメディア音は、合成音から分離される。また、少なくとも１つの抽出された音特徴に基づいて基準音特徴のセット中の少なくとも１つの基準音特徴を識別することによって、モバイルデバイスにおいて実行されるべき機能が決定される。基準音特徴のセット中の各基準音特徴は、複数のメディア音のうちの少なくとも１つと、複数の機能のうちの少なくとも１つとに関連する。決定された機能はモバイルデバイスにおいて実行される。本開示では、本方法に係るデバイス、システム、サーバ、手段の組合せ、およびコンピュータ可読媒体についても説明する。

10

【００１０】

【0010】本開示のさらに別の態様によれば、サーバは、受信機と、識別ユニットと、機能決定ユニットとを含む。受信機は音特徴を受信するように構成され、音特徴は、モバイルデバイスによってキャプチャされたメディア音から抽出される。識別ユニットは、音特徴に基づいて基準音特徴のセット中の少なくとも１つの基準音特徴を識別するように構成され、基準音特徴のセット中の各基準音特徴は、複数のメディア音のうちの少なくとも１つと、複数の機能のうちの少なくとも１つとに関連する。機能決定ユニットは、少なくとも１つの識別された基準音特徴に基づいてモバイルデバイスによって実行されるべき機能を決定するように構成される。

20

【００１１】

【0011】本開示のまた別の態様によれば、モバイルデバイスは、音センサーと、音特徴抽出器と、識別ユニットと、機能決定ユニットと、制御ユニットとを含む。音センサーは、モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャするように構成される。音特徴抽出器は、キャプチャされたメディア音から音特徴を抽出するように構成される。識別ユニットは、音特徴に基づいて基準音特徴のセット中の少なくとも１つの基準音特徴を識別するように構成され、基準音特徴のセット中の各基準音特徴は、複数のメディア音のうちの少なくとも１つと、複数の機能のうちの少なくとも１つとに関連する。機能決定ユニットは、少なくとも１つの識別された基準音特徴に基づいてモバイルデバイスにおいて実行されるべき機能を決定するように構成される。制御ユニットは、モバイルデバイスにおいて、決定された機能を実行するように構成される。

30

【００１２】

【0012】本開示の発明的態様の実施形態は、以下の発明を実施するための形態を参照し、添付の図面とともに読めば理解されよう。

【図面の簡単な説明】

【００１３】

【図１】いくつかの実施形態に従って、モバイルデバイスが、機能を実行するためにメディア音をキャプチャするように構成された、映画館を示す図。

40

【図２】いくつかの実施形態に従って、テレビジョンからキャプチャされたメディア音に基づいて機能を実行するように構成されたモバイルデバイスを装備した視聴者を示す図。

【図３】いくつかの実施形態に従って、それぞれ、テレビジョンからのメディア音に基づいてチャット機能を起動するように構成されたモバイルデバイスを装備した視聴者を示す図。

【図４】いくつかの実施形態に従って、車中のカーオーディオからのメディア音にตอบสนองして機能を実行するように構成され得るモバイルデバイスを装備した運転者を示す図。

【図５】いくつかの実施形態に従って、バス中のスピーカーからのメディア音にตอบสนองして機能を実行するように構成され得るモバイルデバイスを装備した、バスの乗客を示す図。

50

【図 6 A】いくつかの実施形態に従って、メディア音から抽出された音特徴に基づいてモバイルデバイスにおいて実行されるべき機能を決定するように構成されたシステムを示すブロック図。

【図 6 B】いくつかの実施形態に従って、サーバが、モバイルデバイスから受信したメディア音に基づいてモバイルデバイスにおいて実行されるべき機能を決定するように構成された、システムを示すブロック図。

【図 7】いくつかの実施形態に従って、実行されるべき機能を決定するための音特徴データベースを含むモバイルデバイスのブロック図。

【図 8】一実施形態に従って、音特徴データベースのデータ構造を示すブロック図。

【図 9】いくつかの実施形態に従って、サーバにおいて実行されるべき機能を決定することによって、メディア音から抽出された音特徴に基づいてモバイルデバイスによって機能を実行するための方法のフローチャート。

【図 10】いくつかの実施形態に従って、メディア音の音特徴に基づいてモバイルデバイスにおいて実行されるべき機能を決定するための、サーバによって実行される方法のフローチャート。

【図 11】いくつかの実施形態に従って、メディア音に基づいて機能を決定し、実行するための、モバイルデバイスによって実行される方法のフローチャート。

【図 12】いくつかの実施形態に従って、基礎フィルタに基づいてメディア音から音特徴を抽出するための、モバイルデバイスによって実行される方法のフローチャート。

【図 13】いくつかの実施形態に従って、音特徴を抽出する際に使用する複数の基礎フィルタから少なくとも 1 つの基礎フィルタを選択するための方法のフローチャート。

【図 14】いくつかの実施形態に従って、複数のメディア音のうちの少なくとも 1 つをキャプチャすることに応答して少なくとも 1 つの機能を実行するように構成された、複数の音センサーを有するモバイルデバイスを示す図。

【図 15】いくつかの実施形態に従って、複数の音から選択されたメディア音に応答して機能を実行するためのモバイルデバイスにおける方法のフローチャート。

【図 16】いくつかの他の実施形態に従って、複数のメディア音に応答して少なくとも 1 つの機能を実行するためのモバイルデバイスにおける方法のフローチャート。

【図 17】いくつかの実施形態に従って、音識別に基づいて機能を制御するための方法および装置が実装され得る例示的なモバイルデバイスのブロック図。

【図 18】いくつかの実施形態に従って、音識別に基づいて機能を制御するための方法および装置が実装され得るサーバシステムのブロック図。

【発明を実施するための形態】

【0014】

[0032]次に、添付の図面にその例が示されている、様々な実施形態を詳細に参照する。以下の発明を実施するための形態では、本主題の完全な理解を与えるために多数の具体的な詳細を記載する。ただし、本主題はこれらの具体的な詳細なしに実施され得ることが当業者には明らかであろう。他の事例では、様々な実施形態の態様を不必要に不明瞭にしないように、よく知られている方法、プロシージャ、システム、および構成要素については詳細に説明していない。

【0015】

[0033]図 1 に、いくつかの実施形態による、モバイルデバイス 130 が、機能を実行するためにメディア音をキャプチャするように構成された、映画館 100 を示す。映画館 100 にいる視聴者 120（たとえば、映画ファン）はモバイルデバイス 130 を有し得る。モバイルデバイス 130 は、たとえば、マイクロフォンなどの音キャプチャ機能、ならびに / あるいはデータおよび / または通信ネットワークを通じた通信機能を装備した、セルラーフォン、スマートフォン、ラップトップコンピュータまたはタブレットパーソナルコンピュータなど、任意の好適なデバイスであり得る。視聴者 120 は、映画館 100 においてモバイルデバイス 130 の電源を切るのを忘れることがある。

【0016】

10

20

30

40

50

[0034] 視聴者 120 のモバイルデバイス 130 は環境音をキャプチャするように動作可能である。映画が始まる前に、映画館 100 は、1 つまたは複数のスピーカー 140 を通して「電話の電源を切ってください」などのメディア音を再生し得、映画スクリーン 110 上に、対応する視覚メッセージを表示することもある。モバイルデバイス 130 は、メディア音をキャプチャし、キャプチャされたメディア音から音特徴を抽出する。

【0017】

[0035] スピーカー 140 を通して再生されるメディア音は、再生のためにキャプチャされ、録音された音である。本明細書で使用するメディア音（たとえば、ソース音またはデータベース中に登録された音）は、再生のために記憶媒体（たとえば、光媒体、磁気媒体、またはデジタル媒体）上に記録され得る、音声、音楽、または任意のタイプの音であり得る。メディア音は、音を一意に記述するかまたは表す音特徴によって特徴づけられ、オーディオフィンガープリント、音シグネチャ（sound signature）、または任意の他の音識別子を含み得る。メディア音を特徴づける音特徴は、メディア音に応答してモバイルデバイスの電源を切ることなど、モバイルデバイスによって実行されるべき機能に関連し得る。「機能」（または代替的に「アクション」と呼ばれる）という用語は、モバイルデバイスにおいて実行または履行され得る、任意の動作またはその動作のための命令を指す。

【0018】

[0036] キャプチャされたメディア音から音特徴を抽出した後、モバイルデバイス 130 は、抽出された音特徴に関連する機能を決定し得る。図示の例では、「電話の電源を切ってください」というアナウンスは、映画館 100 における再生のために声優によってあらかじめ録音され得る。機能が決定されると、モバイルデバイス 130 は機能を実行する。たとえば、モバイルデバイス 130 は、視覚、オーディオ、または触覚通知を与えることによって、視聴者 120 にモバイルデバイス 130 の電源を切るように促す機能を実行し得る。いくつかの実施形態では、モバイルデバイス 130 は、視聴者 120 による介入を必要とすることなしに、モバイルデバイス 130 の電源を自動的に切るように構成され得る。

【0019】

[0037] 図 2 に、いくつかの実施形態による、テレビジョン 210 からキャプチャされたメディア音に基づいて機能を実行するように構成されたモバイルデバイス 130 を装備した視聴者 120 を示す。図示の実施形態では、テレビジョン 210 からブロードキャストされるいくつかの音はメディア音であり得、メディア音の各々は、モバイルデバイス機能に関連する音特徴によって特徴づけられる。テレビジョン 210 がメディア音をブロードキャストしたとき、モバイルデバイス 130 は、メディア音をキャプチャし、メディア音から音特徴を抽出する。

【0020】

[0038] 一実施形態では、テレビジョン 210 は、スクリーン上にコマーシャルを表示し、同時に、テレビジョン 210 の 1 つまたは複数のスピーカー 240 を通してそのコマーシャルに関連するメディア音を再生し得る。たとえば、テレビジョン 210 は、人気がある缶飲料の広告を再生し得、あらかじめ録音されている、缶飲料に関する音クリップ（たとえば、缶飲料を飲んでいる人）がモバイルデバイス 130 によってキャプチャされる。キャプチャされると、メディア音は、キャプチャされた音クリップのための一意の音特徴を抽出するためにデジタル的に処理される。音特徴は、モバイルデバイス 130 が、情報の検索、たとえば、インターネット検索を自動的に実行し得るように、モバイルデバイス 130 によって実行されるべき機能に関連し得る。缶飲料コマーシャルの例では、モバイルデバイス 130 は、製品詳細、クーポン、または缶飲料に関する検索結果を表示するために、たとえば、検索エンジンを介して検索を自動的に実行し得る。別の例では、モバイルデバイス 130 によって実行されるべき機能は、視聴者 120 がテレビジョン 210 でフットボールの試合を見ている間、フットボールプレーヤーに関する最新の統計の検索を実行することであり得る。また別の例では、モバイルデバイス 130 は、視聴者 120 が見ているテレビ番組、たとえばニュース番組または朝番組のウェブサイトの検索を実

10

20

30

40

50

行し得る。

【0021】

[0039]図3に、いくつかの実施形態による、それぞれテレビジョン312Aおよび312Bからのメディア音に基づいてチャット機能を起動するように構成された、それぞれモバイルデバイス130Aおよび130Bを装備した視聴者120Aおよび120Bを示す。図示の実施形態では、視聴者120Aは第1のロケーション310Aにおいてテレビジョン312Aを見ているが、視聴者120Bは第2のロケーション310Bにおいてテレビジョン312Bを見ている。視聴者120Aおよび120Bが同じテレビ番組（たとえば、ショー、コマーシャル、スポーツイベントなど）を見ているとき、モバイルデバイス130Aおよび130Bは、それぞれテレビジョン312Aおよび312Bからブロードキャストされたメディア音をキャプチャし、キャプチャされたメディア音から音特徴を抽出するように構成される。たとえば、テレビ番組中の俳優330の音声などのメディア音がテレビジョン312Aのスピーカー320Aを通して再生されたとき、モバイルデバイス130Aは俳優330の音声のメディア音をキャプチャし得る。同様に、モバイルデバイス130Bは、視聴者120Aおよび120Bが同じテレビ番組を見ているので、スピーカー320Bを通して俳優330の音声の同じメディア音をキャプチャし得る。

10

【0022】

[0040]モバイルデバイス130Aおよび130Bが俳優330の音声のメディア音をキャプチャすると、モバイルデバイス130Aおよび130Bは、キャプチャされたメディア音から音特徴を抽出し、ネットワーク350を通して、抽出された音特徴をサーバ340に送信し得る。ネットワーク350は、任意のタイプのネットワーク、ならびに、限定はしないが、セルラーネットワーク、回線交換ネットワーク、公衆交換ネットワーク、ローカルエリアネットワーク、ワイドエリアネットワーク、ワイヤレスネットワーク、および/またはパケット交換ネットワークなどを含む、ネットワークの任意の組合せであり得る。サーバ340は、俳優330の音声のメディア音から抽出された基準音特徴を記憶し得る。サーバ340は、モバイルデバイス130Aおよび130Bからの受信した音特徴は、両方の音特徴が基準音特徴と同じであると決定された場合に、同じメディア音を表すと決定する。

20

【0023】

[0041]サーバ340は、次いで、モバイルデバイス130Aおよび130Bによって実行されるべきである、基準音特徴に関連する機能を決定する。たとえば、機能は、音声またはビデオチャットセッション、ソーシャルネットワーキングサービス、テキストおよび/またはピクチャ交換サービスなどを開始することなど、モバイルデバイス130Aとモバイルデバイス130Bの両方の上で同様の動作をアクティブにする機能であり得る。サーバ340は、モバイルデバイスによって実行されるべき機能を識別する情報（たとえば、データまたは命令）をモバイルデバイス130Aおよび130Bに送信する。モバイルデバイス130Aおよび130Bは、実行されるべき機能を識別する情報を受信し、その機能を実行し得る。一実施形態では、モバイルデバイス130Aおよび130Bは、それぞれモバイルデバイス130Aの視聴者120Aとモバイルデバイス130Bの視聴者120Bとの間で通信セッション（たとえば、音声またはビデオチャットセッション）を開始するために、識別された機能を実行する。

30

40

【0024】

[0042]一実施形態では、サーバ340は、モバイルデバイス130Aおよび130Bによるキャプチャされた音が、サーバ340においてデータベース中に登録されたメディア音のうちの1つであるかどうかを決定することによって、音出力デバイスから出力される音（すなわち、メディア音）と音出力デバイス以外のソースから出力されるライブ音（すなわち、メディア音でない音）とを区別することが可能であり得る。たとえば、俳優330が、第2のロケーション310Bに物理的に存在し、テレビジョン312Bがオンになることなしに視聴者120Bに向かって話している場合、モバイルデバイス130Bは、俳優330のライブ音声（すなわち、メディア音でない音）をキャプチャし、キャプチャ

50

された音声から音特徴を抽出し得る。一方、モバイルデバイス130Aは、テレビジョン312Aからの俳優330の音（すなわち、メディア音）をキャプチャして、音特徴を抽出する。モバイルデバイス130Aおよび130Bは、次いで、抽出された音特徴を、モバイルデバイス130Aおよび130Bからの抽出された音特徴が同じ基準音特徴に関連するかどうかを決定するように構成されたサーバ340に送信する。第2のロケーション310Bにおいてキャプチャされた俳優のライブ音声のための音特徴の場合、サーバ340におけるデータベースは、一致する基準音特徴を有さず、したがって、モバイルデバイス130Aによってキャプチャされた音が、モバイルデバイス130Bによってキャプチャされた俳優のライブ音声とは異なると決定する。その結果、サーバ340は、モバイルデバイス130Aおよび130Bによって実行されるべき機能を識別する情報を与えない。

10

【0025】

[0043]図4に、いくつかの実施形態による、車410中のカーオーディオ430からのメディア音に応答して機能を実行するように構成され得るモバイルデバイス130を装備した運転者420を示す。図示の実施形態では、カーオーディオ430は、あらかじめ録音されている広告、音楽、ニュース、または天気予報など、ある音をブロードキャストし得る。たとえば、製品に関係する広告の音クリップが、メディア音としてモバイルデバイス130によってキャプチャされ得る。キャプチャされたメディア音から、モバイルデバイス130は、音特徴を抽出し、抽出された音特徴を図3のサーバ340などのサーバに送信し得る。いくつかの実施形態では、モバイルデバイス130はまた、全地球測位システム（GPS）など、モバイルデバイス130のロケーション追跡機能を使用することによって、抽出された音特徴とともにモバイルデバイス130のロケーション情報をサーバ340に送信し得る。

20

【0026】

[0044]モバイルデバイス130から音特徴を受信した後、サーバ340は、音特徴が、サーバ340に記憶された基準音特徴と同じであるかどうかを決定し得る。モバイルデバイス130から受信した音特徴が基準音特徴と同じであるとサーバ340が決定した場合、サーバ340は、モバイルデバイス130によって実行されるべき機能を決定し、機能を識別する情報をモバイルデバイス130に送信し得る。たとえば、機能は、メディア音を通して広告されている製品が購入され得るロケーションのリストを与えることであり得る。モバイルデバイス130は、サーバ340から、機能を識別する情報を受信し、マップ中にロケーションのリストを表示すること、または会社の名前、ロケーション情報および電話番号など、詳細な情報を与えることによって機能を実行する。

30

【0027】

[0045]別の実施形態では、サーバ340は、モバイルデバイス130の広告された製品および/またはロケーションに基づいて販売促進情報（たとえば、割引券、無料サービスクーポンなど）を与えるために機能を決定し得る。他の機能は、カーオーディオ430上でブロードキャストされ得る任意のメディア音に基づいて、モバイルデバイス130上で可能であり、広告に限定される必要がないことを諒解されよう。

【0028】

[0046]図5に、いくつかの実施形態による、バス510中のスピーカー540からのメディア音に応答して機能を実行するように構成され得るモバイルデバイス130を装備した、バス510の乗客520を示す。乗客520がバス510中にいる間、乗客に次の停留所について通知するために、バス510中のスピーカー540を通してアナウンスなどのメディア音が生じ得る。乗客520のモバイルデバイス130は、そのようなアナウンスをキャプチャし、前に説明したように、モバイルデバイス130上で機能を実行するために処理され得る音特徴を抽出するために、キャプチャされた音を処理するように構成され得る。

40

【0029】

[0047]次の停留所のアナウンスが行われたとき、乗客520が、イヤフォンを用いてモ

50

バイルデバイス 130 で音楽を聞いている場合、乗客 520 はアナウンスが聞こえないことがある。したがって、モバイルデバイス 130 は、アナウンスをキャプチャし、アナウンスの抽出された音特徴に基づいて、オーディオおよび / またはビジュアルメッセージ、触覚通知、またはそれらの任意の組合せを出力することなど、乗客 520 に次の停留所について通知する機能を実行し得る。たとえば、モバイルデバイス 130 は、乗客 520 に対して再生されている音楽を停止し、および / または振動通知を生成するように構成され得る。さらに、モバイルデバイス 130 は、モバイルデバイス 130 のスクリーン上にポップアップメッセージ（たとえば、「次の停留所が目的地です」）を表示し得る。いくつかの実施形態では、乗客 520 は、目的地の停留所がアナウンスされたときに通知を受信するために、目的地情報を用いてモバイルデバイス 130 を事前構成し得る。

10

【0030】

[0048] モバイルデバイス 130 は、メディア音に基づいて他の機能を実行するように構成され得る。たとえば、モバイルデバイス 130 は、目的地の停留所のあらかじめ録音されたアナウンスに基づく地図または公共交通ルートなど、ロケーション情報を表示するように構成され得る。別の例では、モバイルデバイス 130 は、バス 510 のスピーカー 540 を介してアナウンスされるあらかじめ録音された広告をキャプチャし、乗客 520 に、広告された製品に関連する製品情報および / またはロケーション情報を通知し得る。

【0031】

[0049] いくつかの実施形態では、モバイルデバイス 130 は、メディア音（たとえば、ラジオまたはテレビ番組、コマーシャル音クリップ、スピーカーアナウンスなどから）を、再生中いつでもキャプチャし得ることを諒解されよう。いくつかの実施形態では、メディア音は、一定の時間間隔においてモバイルデバイス 130 によってキャプチャされ得る。たとえば、図 3 に示したようにテレビ番組を見る場合、視聴者 120B が視聴者 120A よりも後で同じテレビ番組に同調した場合、モバイルデバイス 130A および 130B は、次に、テレビジョン 312B 上で、関連する音特徴を有するメディア音が再生されたときにチャットサービスを開始する。

20

【0032】

[0050] 図 6A は、いくつかの実施形態による、メディア音から抽出された音特徴に基づいてモバイルデバイス 130A において実行されるべき機能を決定するように構成されたシステム 600A を示すブロック図である。システム 600A は、ネットワーク 350 を介して接続された、モバイルデバイス 130A とサーバ 340A とを含み得る。モバイルデバイス 130A は、音センサー 612 と、音特徴抽出器 614 と、送信機 616 と、受信機 618 と、制御ユニット 620 と、出力ユニット 622 とを含み得る。音センサー 612 は、電話の電源を切ること、通知を送ること、またはインターネット上で検索を行い、検索結果を与えることなど、メディア音に関連する機能を識別するために使用され得る少なくとも 1 つの一意の音特徴を有するメディア音をキャプチャするように構成される。たとえば、音センサー 612 は、1 つまたは複数のマイクロフォン、または周囲の周囲音を検出、感知、および / またはキャプチャするために使用される他のタイプの音センサーを含み得る。音センサー 612 は、通話のためにまたはビデオを録音するためにモバイルデバイス 130A に組み込まれたマイクロフォン、または周囲音を受信するための専用のマイクロフォンであり得る。音特徴抽出器 614 は、音声およびオーディオコーディング、圧縮、認識ならびに分析方法を含む、好適な信号処理方法を使用することによって、音センサー 612 からメディア音を受信し、メディア音から一意の音特徴（たとえば、オーディオフィンガープリントまたは音シグネチャ）を抽出する。たとえば、そのような方法には、オーディオフィンガープリンティング、音声認識または音声およびオーディオコーディングのための周知の方法である、スペクトル分析、MFCC（メル周波数ケプストラム係数（Mel-frequency cepstral coefficient））、LPC（線形予測コーディング（linear predictive coding））、および / または LSP（線スペクトル対（line spectral pair））技法がある。また、以下でより詳細に説明するように、一意の音特徴は、キャプチャされたメディア音に基礎フィルタを適用することによって抽出され得る。

30

40

50

【 0 0 3 3 】

[0051] モバイルデバイス 1 3 0 A の送信機 6 1 6 は、ネットワーク 3 5 0 を介して、抽出された音特徴をサーバ 3 4 0 A に送信するように構成される。サーバ 3 4 0 A は、以下でより詳細に説明するように、モバイルデバイス 1 3 0 A によって実行されるべき機能を決定するためにモバイルデバイス 1 3 0 A から受信した音特徴を処理し、ネットワーク 3 5 0 を介してモバイルデバイス 1 3 0 A に機能を識別する情報を与える。モバイルデバイス 1 3 0 A の受信機 6 1 8 は、音特徴に基づいて決定された機能を識別する情報を受信し、機能を識別する情報を、モバイルデバイス 1 3 0 A において機能を実行するように構成された制御ユニット 6 2 0 に送信するように構成される。たとえば、制御ユニット 6 2 0 は、モバイルデバイス 1 3 0 A の電源を切ることまたは出力ユニット 6 2 2 を通して通知メッセージを表示することによって機能を実行し得る。

10

【 0 0 3 4 】

[0052] サーバ 3 4 0 A は、モバイルデバイス 1 3 0 A から受信した音特徴に基づいてモバイルデバイス 1 3 0 A 上で実行されるべき機能を決定するように構成される。サーバ 3 4 0 A は、受信機 6 4 2 と、機能決定ユニット 6 4 4 と、音特徴データベース 6 4 6 と、基礎フィルタ生成器 6 5 0 と、送信機 6 4 8 と、サンプル音データベース 6 5 2 とを含み得る。受信機 6 4 2 は、モバイルデバイス 1 3 0 A から音特徴を受信するように構成され、モバイルデバイス 1 3 0 A によって実行されるべき機能を決定するために音特徴を機能決定ユニット 6 4 4 に与える。

【 0 0 3 5 】

20

[0053] 機能決定ユニット 6 4 4 は、受信した音特徴との最も大きい類似度を有する基準音特徴を識別するために、音特徴データベース 6 4 6 にアクセスするように構成される。音特徴データベース 6 4 6 は複数の基準音特徴を含み、基準音特徴の各々は、メディア音から抽出され、モバイルデバイスによって実行されるべき機能に関連する。一実施形態では、抽出された音特徴または基準音特徴は、メディア音を一意に記述するかまたは表す値の任意の好適なセットによって表され得る。たとえば、値のセットは、周波数領域スペクトル、電力領域スペクトルなどに対応し得る。メディア音は、音を一意に記述するかまたは表す音特徴によって特徴づけられ、オーディオフィンガープリント、音シグネチャ、または任意の他の音識別子を含み得る。したがって、音特徴データベース 6 4 6 は、複数の抽出された音特徴または基準音特徴から構成され得、それらの音特徴の各々は、関連する機能を有する。本明細書で説明するように、音特徴がメディア音から抽出され、関連する機能とともに音特徴データベース 6 4 6 中に含まれているとき、メディア音は、音特徴データベース 6 4 6 中に登録されていると見なされ得る。したがって、入力メディア音が音特徴データベース 6 4 6 中に登録されているとき、音特徴データベース 6 4 6 から、モバイルデバイスによって実行されるべき対応する機能が識別され得る。

30

【 0 0 3 6 】

[0054] 以下の表 1 に示すように、音特徴データベース 6 4 6 は複数の基準音特徴 1 ~ N を含み得、基準音特徴の各々は、実行されるべき機能に関連する。たとえば、基準音特徴 1 は、「電話の電源を切る」という機能に関連するが、基準音特徴 2 は、インターネット上で製品検索を行い、検索結果を与える機能に関連し得る。

40

【表 1】

表 1

基準音特徴	実行されるべき機能
基準音特徴 1	電話の電源を切る
基準音特徴 2	インターネット上で「製品」検索を行い、検索結果を与える
⋮	⋮
基準音特徴 N	機能 N

10

20

【0037】

[0055]サーバ340A中の基礎フィルタ生成器650は、サンプル音データベース652のサンプル音を使用することによって、抽出された音特徴に適用され得る1つまたは複数の基礎フィルタを生成するように構成される。さらに、基礎フィルタは、音特徴データベース646の精度を改善するために、音特徴データベース646を更新するために使用され得る。基礎フィルタ生成器650の具体的な動作を以下でより詳細に与える。

【0038】

[0056]受信した音特徴に基づいて、機能決定ユニット644は、音特徴データベース646にアクセスし、受信した音特徴との最も大きい類似度を有する基準音特徴を決定する。基準音特徴が音特徴データベース646から識別されると、機能決定ユニット644は、音特徴データベース646からの識別された基準音特徴に関連する機能を決定する。サーバ340Aは、送信機648を介してネットワーク350を通して、機能を識別する情報をモバイルデバイス130Aに送信する。モバイルデバイス130Aの制御ユニット620は、受信機618を介してサーバ340Aから、機能を識別する情報を受信し、たとえば、モバイルデバイス130Aの電源を切ることまたは出力ユニット622を通して通知メッセージを表示することによって、モバイルデバイス130Aにおいて機能を実行する。

30

40

【0039】

[0057]一実施形態では、モバイルデバイス130Aまたは130Bは、ユーザ介入を必要とすることなしに、音処理機能を自動的に実行するように構成され得る。言い換えれば、ユーザは、モバイルデバイス130Aまたは130B上での機能の実行を手動で開始する必要がない。この場合、モバイルデバイス130Aまたは130Bの音センサー612は、モバイルデバイス130Aまたは130B上で実行されるべき機能を決定するために、たとえば、連続的に環境音をキャプチャして、登録されたメディア音を認識し得る。しかしながら、モバイルデバイス130Aまたは130Bが連続的に環境音をキャプチャし、処理するように動作している間、電力は、概して、一般に容量が制限されるバッテリーから消費される。したがって、バッテリーを節約するために、以下のいくつかの実施形態で説

50

明するように、低電力モードで動作することによって、モバイルデバイス 130A または 130B の電力消費が低減され得る。

【0040】

[0058]いくつかの実施形態では、モバイルデバイス 130A は、電力消費を低減するために、所定の時間間隔において環境音をキャプチャし得る。この場合、環境音をキャプチャまたは処理するために使用されるモバイルデバイス 130A 中の構成要素が、所定の時間間隔（たとえば、100ms の間隔において 10ms の持続時間をもつ音）においてアクティブにされ得る。たとえば、そのような構成要素は、10ms では「アクティブ状態」に（すなわち、高電力状態に）あり、次の 90ms では「アイドル状態」に（すなわち、低電力状態またはオフ状態に）あるように構成され得る。アクティブ状態の間に、モバイルデバイス 130A は、環境音をキャプチャし、キャプチャされた音が所定のしきい値音レベル（たとえば、電力レベル）を超えるかどうかを決定する。キャプチャされた音の音レベルがしきい値音レベルを超えた場合、モバイルデバイス 130A は、キャプチャされた音を処理して、モバイルデバイス 130A 上で実行されるべき機能を決定し得る。アイドル状態では、環境音をキャプチャまたは処理するために使用される構成要素は、デューティサイクルの 90% について非アクティブである。したがって、そのような場合、モバイルデバイス 130A は、連続的に環境音をキャプチャし、処理する場合よりも、音処理における電力の 90% をも節約し得る。

【0041】

[0059]モバイルデバイス 130A はまた、キャプチャされた音が、音声または音楽など、メディア音であるかどうかを決定するように構成され得る。一実施形態では、音センサー 612 は、雑音などの非メディア音と区別するために、キャプチャされた音がメディア音であるかどうかを決定するためのデジタル信号プロセッサ（「DSP」）を含み得る。たとえば、キャプチャされた音の音レベルがしきい値音レベルを超えると決定されたとき、DSP は、MFCC（メル周波数ケプストラム係数）をもつガウス混合モデル（Gaussian Mixture Model）など、好適な信号処理方法を使用することによって、キャプチャされた音がメディア音であるか否かを決定するためにアクティブにされ得る。キャプチャされた音が、音声または音楽など、メディア音であると DSP が決定した場合、キャプチャされた音は、上記で説明したように、さらなる処理のために音特徴抽出器 614 に与えられる。一方、キャプチャされた音がメディア音でないと DSP が決定した場合、キャプチャされた音はそれ以上処理されない。音処理のフロントエンドにおいて非メディア音を識別し、破棄することによって、モバイルデバイス 130A は、すべてのキャプチャされた音を処理する場合よりもかなりの電力を節約し得る。

【0042】

[0060]図 6B は、別の実施形態による、サーバ 340B が、モバイルデバイス 130B から受信したメディア音に基づいてモバイルデバイス 130B において実行されるべき機能を決定するように構成された、システム 600B を示すブロック図ある。図示のように、システム 600B は、ネットワーク 350 を介して接続された、モバイルデバイス 130B とサーバ 340B とを含む。この実施形態では、音特徴抽出器 643 は、図 6A のシステム 600A とは対照的に、モバイルデバイス 130B ではなくサーバ 340B 中に含まれる。モバイルデバイス 130B の音センサー 612 がメディア音をキャプチャした後、モバイルデバイス 130B は、送信機 616 を通してメディア音をサーバ 340B に直接送信する。

【0043】

[0061]サーバ 340B では、受信機 642 は、モバイルデバイス 130B から送信されたメディア音を受信し、受信したメディア音を音特徴抽出器 643 に与える。音特徴抽出器 643 は、メディア音から一意の音特徴を抽出するように構成される。サーバ 340B は、次いで、モバイルデバイス 130B によって実行されるべき機能を決定し、ネットワーク 350 を介してモバイルデバイス 130B に機能を識別する情報を与えるために、図 6A に関して上記で説明したのと同じ方法で、抽出された音特徴を処理し得る。

【 0 0 4 4 】

[0062] 図 7 に、いくつかの実施形態による、実行されるべき機能を決定するための音特徴データベース 740 を含むモバイルデバイス 700 のブロック図を示す。モバイルデバイス 700 は、少なくとも 1 つの一意の音特徴を有するメディア音をキャプチャするように構成された音センサー 710 と、音センサー 710 からメディア音を受信し、キャプチャされたメディア音から音特徴（たとえば、オーディオフィンガープリント）を抽出するように構成された音特徴抽出器 720 とを含む。モバイルデバイス 700 は、抽出された音特徴を音特徴データベース 740 に記憶された複数の基準音特徴と比較するように構成された機能決定ユニット 730 をさらに含む。

【 0 0 4 5 】

[0063] 機能決定ユニット 730 は、受信した音特徴に対し最大類似度を有する基準音特徴を識別するために、音特徴データベース 740 にアクセスするように構成される。上記で説明したように、音特徴データベース 740 は複数の基準音特徴を含み、基準音特徴の各々は、メディア音から抽出され、モバイルデバイスによって実行されるべき機能に関連する。音特徴データベース 740 の複数の基準音特徴は、モバイルデバイス 700 においてローカルで更新されるか、またはサーバと通信してリモートで更新され得る。たとえば、メディア音をキャプチャし、関連する機能をモバイルデバイス 700 に入力するために、ユーザがモバイルデバイス 700 を操作し得る。モバイルデバイス 700 は、次いで、キャプチャされたメディア音から基準音特徴を抽出し、当該関連する機能とともに当該抽出された基準音特徴を音特徴データベース 740 に記憶し得る。代替として、音特徴データベース 740 は、リモートサーバから、関連する機能をもつ基準音特徴をダウンロードし得る。音特徴データベース 740 から、機能決定ユニット 730 は、識別された基準音特徴に関連する機能を決定する。機能決定ユニット 730 からの決定された機能に基づいて、制御ユニット 760 はモバイルデバイス 700 において機能を実行する。

【 0 0 4 6 】

[0064] 図 8 は、一実施形態による、音特徴データベース 646 のデータ構造 800 を示すブロック図である。図 6 A および図 6 B の音特徴データベース 646 はデータ構造 800 によって記述され得るが、図 7 の音特徴データベース 740 もデータ構造 800 を使用し得る。データ構造 800 は、複数のメタデータ 810 A ~ 810 N を含み、それらは、それぞれ、基準音特徴 812 A ~ 812 N、機能 814 A ~ 814 N、およびコンテキスト情報 816 A ~ 816 N を含む。コンテキスト情報 816 A ~ 816 N は、基準音特徴 812 A ~ 812 N に追加の情報またはパラメータをそれぞれ与え得る。したがって、メタデータ 810 A ~ 810 N の各々は、対応する基準音特徴の特性を記述する。たとえば、「電話の電源を切ってください」というメディア音は、メディア音から抽出された基準音特徴 812 A と、モバイルデバイス上で実行されるべき機能 814 A（たとえば、モバイルデバイスの電源を切ることまたはモバイルデバイスの構成をサイレントモードに調整すること）と、メディア音がキャプチャされ得る環境コンテキスト（たとえば、映画館）を識別するコンテキスト情報 816 A とを含む、メタデータ 810 A によって表され得る。

【 0 0 4 7 】

[0065] メタデータ 810 A ~ 810 N は、サーバ 340 または別のリモートサーバにおいてローカルに事前生成され、音特徴データベース 646 に記憶され得る。たとえば、メタデータ 810 A は、ターゲットメディア音から基準音特徴 812 A を抽出し、当該基準音特徴 812 A に関連する機能 814 A として望ましい機能を与えることによって生成され得る。基準音特徴 812 A を記述するかまたは特徴づけ得る追加の情報またはパラメータは、コンテキスト情報 816 A として与えられ得る。上記の方法で生成されたメタデータ 810 A ~ 810 N は、サーバ 340 にアップロードされ、音特徴データベース 646 に記憶され得る。

【 0 0 4 8 】

[0066] 図 9 は、いくつかの実施形態による、サーバ 340 において実行されるべき機能

10

20

30

40

50

を決定することによって、メディア音から抽出された音特徴に基づいてモバイルデバイス 130 によって機能を実行するための方法 900 のフローチャートである。初めに、モバイルデバイス 130 は、910 において、音出力デバイス（たとえば、図 1 に示したように映画館 100 におけるスピーカー 140）を介して再生されるメディア音をキャプチャする。モバイルデバイス 130 は、920 において、キャプチャされた音进行处理して、キャプチャされたメディア音から音特徴を抽出する。930 において、モバイルデバイス 130 は、抽出された音特徴をサーバ 340 に送信し、サーバ 340 は、モバイルデバイス 130 によって実行されるべき機能を決定し、ネットワークを介して、機能を識別する情報をモバイルデバイス 130 に送信する。モバイルデバイス 130 は、940 において、サーバ 340 から、機能を識別する情報を受信し、950 において機能を実行する。

10

【0049】

[0067] 図 10 は、いくつかの実施形態による、メディア音の音特徴に基づいてモバイルデバイス 130 において実行されるべき機能を決定するための、サーバ 340 において実行される方法 1000 のフローチャートである。1010 において、サーバ 340 は、モバイルデバイス 130 からの抽出された音特徴を受信する。次いで、1020 において、サーバ 340 は、モバイルデバイス 130 からの抽出された音特徴と音特徴データベース 646 からの基準音特徴の各々との間の類似度値を決定する。類似度値を決定した後、サーバ 340 は、1030 において、最も大きい類似度値を決定し、最も大きい類似度値に関連する音特徴データベース 646 における基準音特徴を識別する。一実施形態では、最も大きい類似度値が所定のしきい値よりも小さい場合、基準音特徴は識別されず、したがって、サーバ 340 は、送信された音特徴について識別された機能がないことを示す応答をモバイルデバイス 130 に送信し得る。最も大きい類似度値を有する基準音特徴が決定されると、サーバ 340 は、1040 において、音特徴データベース 646 から、識別された基準音特徴に関連する機能を決定する。次いで、1050 において、サーバ 340 は、ネットワークを通して、決定された機能を識別する情報をモバイルデバイス 130 に送信する。

20

【0050】

[0068] 図 11 は、いくつかの実施形態による、メディア音に基づいて機能を決定し、実行するための、モバイルデバイス 130 によって実行される方法 1100 のフローチャートである。モバイルデバイス 130 は、初めに、1110 においてメディア音をキャプチャする。モバイルデバイス 130 は、1120 において、キャプチャされた音进行处理して、キャプチャされたメディア音から音特徴を抽出する。次いで、1130 において、モバイルデバイス 130 は、抽出された音特徴とモバイルデバイス 130 中の音特徴データベースからの基準音特徴の各々との間の類似度値を決定する。類似度値を決定した後、モバイルデバイス 130 は、1140 において、最も大きい類似度値を決定し、音特徴データベース中の最も大きい類似度値に関連する基準音特徴を識別する。最も大きい類似度値を有する基準音特徴が決定されると、モバイルデバイス 130 は、1150 において、音特徴データベースから、識別された基準音特徴に関連する機能を決定する。次いで、1160 において、モバイルデバイス 130 は、決定された機能を実行する。

30

【0051】

[0069] 図 12 に、いくつかの実施形態に従って、基礎フィルタに基づいてメディア音から音特徴を抽出するための、モバイルデバイス 130 によって実行される方法 1200 のフローチャートを示す。モバイルデバイス 130 の音センサー 612 は、初めに、1210 においてメディア音をキャプチャする。1220 において、モバイルデバイス 130 の音特徴抽出器 614 は、音センサー 612 からのキャプチャされたメディア音を受信し、メディア音を所定の時間期間を有する複数のフレームにセグメント化する（「ウィンドウ処理」と呼ばれることもある）。一実施形態では、隣接するフレームは、そのフレームの一部がそれらのフレームに共通であるようにオーバーラップするようにセグメント化される。そのような場合、オーバーラップする部分の時間期間があらかじめ決定され得る。フレームおよびオーバーラップする部分の時間期間は、音一意性（sound uniqueness）に

40

50

おける所望のレベルの精度を達成するように調整され得る。たとえば、音一意性を高めることによって、フレームおよび／またはオーバーラップする部分のためのより小さい時間期間が、より高い精度をもたらし得る。

【 0 0 5 2 】

[0070]メディア音を複数のフレームにセグメント化した後、たとえば、高速フーリエ変換（FFT：Fast Fourier Transform）を使用することによって、1230において、各フレームを周波数領域スペクトルに変換する。代替的に、いくつかの他の変換プロセス（たとえば、限定はしないが、離散コサイン変換、ハール変換（Haar Transform）またはウォルシュアダマール変換（Walsh-Hadamard Transform）を含む）が、フレームを変換するために使用され得る。一実施形態では、周波数領域スペクトルにおける変換されたフレームは、抽出された音特徴、または抽出された音特徴を生成するための中間音特徴として使用され得る。

10

【 0 0 5 3 】

[0071]フレームに対応する各周波数領域スペクトルについて、音特徴抽出器614は、1240において電力スペクトルを計算する。このプロセスでは、各フレームについての周波数領域スペクトルは周波数帯域にセグメント化され得、各帯域におけるエネルギー値（すなわち、電力値）がそれぞれの周波数帯域の電力レベルとして計算される。したがって、周波数帯域の計算されたエネルギー値は、フレームを一意に識別する中間音特徴として使用され得るフレームの電力スペクトルを表す。

【 0 0 5 4 】

20

[0072]各フレーム（すなわち、各フレームの中間音特徴）についてスペクトルが取得されると、音特徴抽出器614は、1250において、各電力スペクトルに基礎フィルタを適用することによって音特徴（すなわち、抽出された音特徴）を生成する。本明細書で使用する「基礎フィルタ」は、中間音特徴よりも雑音およびひずみに対してよりロバストである抽出された音特徴を生成するために中間音特徴（たとえば、メディア音から抽出された電力スペクトル）に適用され得る、値のセット、ベクトル、または行列を指す。基礎フィルタを適用することによって中間音特徴から抽出された音特徴は、中間特徴よりも高いレベルの特徴を表し得る。たとえば、抽出された音特徴は、キャプチャされた音の周波数成分の変化または組合せを表すが、中間特徴は、キャプチャされた音の周波数成分を表すにすぎない。別の例では、抽出された音特徴は、キャプチャされた音のオンセットまたはオフセットを表し得る。

30

【 0 0 5 5 】

[0073]上記で生成された、抽出された音特徴は、キャプチャされた音を表す一連の値からなり得る。次いで、1260において、任意の好適な2値化アルゴリズムを採用することによって、抽出された音特徴の値を2進形式（たとえば、「0」または「1」）に変換する。一実施形態では、抽出された音特徴の各値は、抽出された音特徴の2値化された形態を生成するためにしきい値と比較され得る。たとえば、抽出された音特徴の値がしきい値に等しいかまたはそれよりも大きい場合、ビット「1」が値に割り当てられるが、値がしきい値よりも小さい場合、ビット「0」が値に割り当てられる。別の実施形態では、一連の値のうちの所定の数の値が選択され、「1」の値を割り当てられ得るが、残りの値は「0」の値を割り当てられる。たとえば、一連の値のうちのN個の最大値が選択され、「1」の値を割り当てられ得るが、残りの値は「0」の値を割り当てられる。2値化の後、1270において、抽出された音特徴として、2値化された音特徴を出力する。

40

【 0 0 5 6 】

[0074]いくつかの実施形態では、基準音と基準音のひずみ音との間の差を最小限に抑えるために値または行列を計算することによって、基礎フィルタが生成され得る。本明細書で使用する、「基準音セット」と総称され得る、「基準音」（または「サンプル音」）および基準音のひずみ音は、雑音のない元の音（すなわち、無ひずみ基準音）と、雑音のある元の音（すなわち、ひずみ基準音）とのセットを指す。たとえば、基準音セットは、音楽またはアナウンスなど、雑音のない元の音（すなわち、無ひずみ基準音）と、雑音のあ

50

る同じ元の音（すなわち、ひずみ基準音）とから抽出され得る。

【 0 0 5 7 】

[0075]一実施形態によれば、無ひずみ基準音とその対応するひずみ基準音との間の誤差 E を最小限に抑えるために、基礎フィルタ M が生成され得、

【数 1】

$$E = x_i - M \cdot x_i^k$$

【 0 0 5 8 】

のように行列形式で表され得る。

【 0 0 5 9 】

10

上式で、 x_i は、 i 番目の無ひずみ基準音の中間音特徴を示すベクトルであり、 x_i^k は、 i 番目の無ひずみ基準音の k 番目のひずみ音の中間音特徴を示すベクトルである。上式では、誤差 E は、基礎フィルタ M のための係数を調整または最適化することによって減少され得る。いくつかの実施形態では、基礎フィルタ M は、限定はしないが、PCA（主成分分析（Principal Components Analysis））、OPCA（指向性主成分分析（Oriented Principal Components Analysis））、ICA（独立成分分析（Independent Component Analysis））、スパースコーディング、および他の同様のアルゴリズムを含む、任意の好適な技法を採用することによって生成され得る。

【 0 0 6 0 】

20

[0076]上記の基礎フィルタ M に基づいて、基礎フィルタ M をメディア音から抽出された中間音特徴に適用することにより、中間音特徴よりも雑音およびひずみに対してよりロバストである音特徴を生成する。基礎フィルタ M は、サーバ 340 の基礎フィルタ生成器 650 によって生成され、音特徴を抽出する際に使用するために、送信機 648 を介してモバイルデバイス 130 に与えられ得る。代替として、モバイルデバイス 130 はサーバ 340 に中間音特徴を送信し得、サーバ 340 は、音特徴を抽出する際に基礎フィルタ M を使用し得る。別の実施形態では、モバイルデバイス 130 はサーバ 340 にメディア音を送信し得、サーバ 340 は、メディア音から中間音特徴を抽出し、音特徴を抽出する際に基礎フィルタ M を使用し得る。上記のように生成された基礎フィルタ M は、音特徴データベース 646 において基準音特徴を生成し、更新するために使用され得る。

【 0 0 6 1 】

30

[0077]いくつかの実施形態では、基礎フィルタが OPCA 法によって生成されたとき、基礎フィルタは、基礎フィルタの 4 つの行を有する行列によって特徴づけられ得、以下の式に従って音特徴 C が生成され得る。

【数 2】

$$C = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \\ c_1 & c_2 & \dots & c_n \\ d_1 & d_2 & \dots & d_n \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \\ \cdot \\ \cdot \\ f_n \end{bmatrix}$$

40

【 0 0 6 2 】

上式で、行 $[a_1 \ a_2 \ \dots \ a_n]$ 、 $[b_1 \ b_2 \ \dots \ b_n]$ 、 $[c_1 \ c_2 \ \dots \ c_n]$ 、および $[d_1 \ d_2 \ \dots \ d_n]$ は、それぞれ第 1 の基礎フィルタ、第 2 の基礎フィルタ、第 3 の基礎フィルタ、および第 4 の基礎フィルタを示すが、列 $[f_1 \ f_2 \ \dots \ f_n]$ は、メディア音から抽出された中間音特徴（たとえば、電力スペクトル）を示す。基礎フィルタ行列は、4 つの基礎フィルタを使用して示されるが、基礎フ

50

フィルタ行列は、中間音特徴の対応するエントリをもつ任意の好適な数の基礎フィルタを含み得る。

【 0 0 6 3 】

[0078] 上記で説明したように、基礎フィルタ（たとえば、基礎フィルタ行列）は、無ひずみ基準音とひずみ基準音とを比較することによって生成され、メディア音から雑音ロバスタな音特徴を抽出する際に使用され得る。複数の基礎フィルタの場合、いくつかの基礎フィルタは、他の基礎フィルタよりも雑音を低減する際に良好な結果を与え得、したがって、音特徴を抽出する際に使用するために選択され得る。上記の基礎フィルタ行列では、たとえば、類似する音特徴が、無ひずみ基準音とひずみ基準音とに基づいて第 1 の基礎フィルタおよび第 3 の基礎フィルタから生成され得る。一方、無ひずみ基準音とひずみ基準音とに基づいて、第 1 および第 3 の基礎フィルタと比べて、より類似しない音特徴が第 2 の基礎フィルタおよび第 4 の基礎フィルタから生成され得る。したがって、第 1 の基礎フィルタおよび第 3 の基礎フィルタは、以下で説明するように、メディア音から音特徴 C を生成するように選択され得る（基礎フィルタを「更新する」と呼ばれることもある）。

【 数 3 】

$$C = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \\ c_1 & c_2 & \dots & c_n \\ d_1 & d_2 & \dots & d_n \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \\ \cdot \\ \cdot \\ f_n \end{bmatrix} \rightarrow C = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ c_1 & c_2 & \dots & c_n \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \\ \cdot \\ \cdot \\ f_n \end{bmatrix}$$

【 0 0 6 4 】

[0079] 図 1 3 に、いくつかの実施形態による、音特徴を抽出する際に使用する複数の基礎フィルタから少なくとも 1 つの基礎フィルタを選択するための方法 1 3 0 0 のフローチャートを示す。基礎フィルタを選択するための方法 1 3 0 0 は、図 6 A および図 6 B のサーバ 3 4 0 中の基礎フィルタ生成器 6 5 0 またはモバイルデバイス 1 3 0 中の基礎フィルタ生成器によって実行され得る。さらに、基礎フィルタを選択するための方法 1 3 0 0 は、モバイルデバイス 1 3 0 においてメディア音を受信する前に実行され得、選択された基礎フィルタは、メディア音を受信した後にメディア音から音特徴を抽出するために使用され得る。方法 1 3 0 0 について、説明の目的で第 1 ~ 第 3 のサンプル音を参照しながら説明する。

【 0 0 6 5 】

[0080] 図示の方法 1 3 0 0 では、基礎フィルタ生成器 6 5 0 は、1 3 0 2、1 3 0 4、および 1 3 0 6 において、サンプル音データベース 6 5 2 から、それぞれ第 1 のサンプル音、第 2 のサンプル音、および第 3 のサンプル音を受信する。第 1 のサンプル音は、雑音のない無ひずみ音である。第 2 のサンプル音は、雑音を加えた第 1 のサンプル音のひずみ音であり、したがって、音コンテンツにおいて第 1 のサンプル音に類似する。第 3 のサンプル音は、雑音のない別の無ひずみ音であり、したがって、音コンテンツにおいて第 1 のサンプル音と異なる。たとえば、スピーカー（たとえば、最小雑音およびひずみをもつ高忠実度（high fidelity）スピーカー）を通して再生された音楽からキャプチャされた音は、コンテンツにおいて別のスピーカー（たとえば、低忠実度（low fidelity）スピーカー）を通して再生された同じ音楽に類似すると見なされ得る。一方、高忠実度スピーカーを通して再生されたアナウンス（たとえば、「電話の電源を切ってください」）は、コンテンツにおいて同じ高忠実度スピーカーを通して再生された音楽に類似しないと見なされ得る。

【 0 0 6 6 】

[0081] サンプル音データベース 6 5 2 からサンプル音を受信した後、基礎フィルタ生成

器 6 5 0 は、初めに、1 3 0 8 において複数の基礎フィルタから 1 つの基礎フィルタを選択する。第 1 ~ 第 3 のサンプル音の各々から、基礎フィルタ生成器 6 5 0 は、1 3 1 0 において、上記で説明した電力スペクトルなど、中間音特徴を抽出する。中間音特徴の各々について、基礎フィルタ生成器 6 5 0 は、1 3 1 2 において、中間音特徴よりも雑音およびひずみに対してよりロバストである音特徴を生成するために、選択された基礎フィルタを各中間音特徴に適用する。

【 0 0 6 7 】

[0082] サンプル音の音特徴が生成されると、選択された基礎フィルタを適用することによってサンプル音のペアから抽出された音特徴間の類似度の程度を示す 1 つまたは複数の信頼性値が、ペアワイズブースティング (P B : pairwise boosting) アルゴリズムなどの任意の好適なアルゴリズムを採用することによって計算され得る。選択された基礎フィルタがサンプル音のペアの中間音特徴に適用されたとき、信頼性値は雑音低減の程度を示し得る。一実施形態では、基礎フィルタ生成器 6 5 0 は、類似するサンプル音の抽出された音特徴間の差の信頼性値と、類似しないサンプル音の抽出された音特徴間の差の別の信頼性値とを計算し得る。

10

【 0 0 6 8 】

具体的には、1 3 1 4 において、第 1 のサンプル音の抽出された音特徴と第 2 のサンプル音の抽出された音特徴との間の差を計算することによって、基礎フィルタの第 1 の信頼性値を決定する。たとえば、第 1 および第 2 のサンプル音が、類似する音であると仮定すれば、選択された基礎フィルタを第 1 および第 2 のサンプル音の中間音特徴に適用することによって抽出された音特徴間の差が低い場合、基礎フィルタの信頼性値は高い。さらに、1 3 1 6 において、第 1 のサンプル音の抽出された音特徴と第 3 のサンプル音の抽出された音特徴との間の差を計算することによって、基礎フィルタの第 2 の信頼性値を決定する。この場合、第 1 および第 3 のサンプル音が、類似しない音であると仮定すれば、選択された基礎フィルタを第 1 および第 3 のサンプル音に適用することによって抽出された音特徴間の差が高い場合、基礎フィルタの信頼性値は高い。

20

【 0 0 6 9 】

[0083] 1 3 1 8 において、基礎フィルタ生成器 6 5 0 は、第 1 および第 2 の信頼性値に基づいて、選択された基礎フィルタのための総合信頼性値を決定する。一実施形態では、総合信頼性値は、第 1 および第 2 の信頼性値を加算することによって計算され得る。代替として、総合信頼性値は、重み付けされた第 1 および第 2 の信頼性値を加算することによって計算され得る。

30

【 0 0 7 0 】

[0084] 選択された基礎フィルタの総合信頼性値を生成した後、基礎フィルタ生成器 6 5 0 は、1 3 2 0 において、選択され得る別の基礎フィルタがあるかどうかを決定する。別の基礎フィルタがある場合、方法 1 3 0 0 は、1 3 0 8 に戻り、総合信頼性値を決定するために次の基礎フィルタを選択する。総合信頼性値がすべての基礎フィルタについて計算されると、基礎フィルタ生成器 6 5 0 は、1 3 2 2 において、総合信頼性値に基づいて複数の基礎フィルタから少なくとも 1 つの基礎フィルタを選択する。一実施形態では、最も高い総合信頼性値を有する基礎フィルタが選択され得る。代替として、所定のしきい値を超える総合信頼性値を有するすべての基礎フィルタが選択され得る。

40

【 0 0 7 1 】

[0085] 図 1 4 に、いくつかの実施形態による、複数のメディア音のうちの少なくとも 1 つをキャプチャすることに応答して少なくとも 1 つの機能を実行するように構成された、複数の音センサー 1 4 4 0 および 1 4 5 0 を有するモバイルデバイス 1 4 3 0 を示す。モバイルデバイス 1 4 3 0 において、2 つの音センサー 1 4 4 0 および 1 4 5 0 は、たとえば、2 つのマイクロフォンであり得る。各音センサー 1 4 4 0 および 1 4 5 0 は、周辺環境から、複数の音を含む合成音をキャプチャ (たとえば、TV 1 4 1 0 からのメディア音と時計 1 4 2 0 からのアラーム音とを同時にキャプチャ) し得る。一実施形態では、音センサー 1 4 4 0 および 1 4 5 0 の各々からキャプチャされた合成音は、たとえば、モバイ

50

ルデバイス 1 4 3 0 の異なるロケーションにおいて音をキャプチャする音センサー 1 4 4 0 および 1 4 5 0 により、互いに区別され得る。音センサー 1 4 4 0 および音センサー 1 4 5 0 から受信した異なる合成音に基づいて、モバイルデバイス 1 4 3 0 は合成音を音の各々に分離し得る。図示の実施形態では、モバイルデバイス 1 4 3 0 は、TV 1 4 1 0 からのメディア音および時計 1 4 2 0 からのアラーム音と混合された合成音をキャプチャし得る。モバイルデバイス 1 4 3 0 は、合成音を、それぞれ TV 1 4 1 0 からのメディア音と時計 1 4 2 0 の音とに分離し得る。

【 0 0 7 2 】

[0086] ソース（たとえば、TV またはアラーム）に従って合成音を分離した後、モバイルデバイス 1 4 3 0 は、分離された音の各々から音特徴を抽出し、抽出された音特徴のうちのいずれか 1 つが、モバイルデバイス 1 4 3 0 上で実行されるべき機能に関連するかどうかを決定し得る。モバイルデバイス 1 4 3 0 は、分離された音のうちの 1 つから、抽出された音特徴に関連する機能を実行し得る。音センサー 1 4 4 0 および 1 4 5 0 によってキャプチャされた複数のメディア音から複数の機能が決定され得ることを諒解されよう。

【 0 0 7 3 】

[0087] 図 1 5 に、いくつかの実施形態による、複数の音から選択された 1 つのメディア音に応答して機能を実行するためのモバイルデバイス 1 4 3 0 における方法 1 5 0 0 のフローチャートを示す。この方法 1 5 0 0 では、モバイルデバイス 1 4 3 0 は、1 5 1 0 において、音センサー（たとえば、マイクロフォン）を通して、少なくとも 1 つのメディア音を含む音（すなわち、合成音）をキャプチャする。次いで、モバイルデバイス 1 4 3 0 は、1 5 2 0 において、キャプチャされた合成音から少なくとも 1 つのメディア音を分離する。この実施形態では、キャプチャされた音は、独立成分分析（ICA : independent component analysis）またはブラインド音源分離（BSS : blind source separation）を採用することによって分離され得るが、それに限定されない。それらの音を分離した後、モバイルデバイス 1 4 3 0 は、それらの音のうちの少なくとも 1 つがメディア音であるか、またはおそらくメディア音であるという識別に基づいて、1 5 3 0 において、複数の分離された音のうち 1 つのメディア音を選択する。一実施形態では、モバイルデバイス 1 4 3 0 が、映画館に位置し、シアター音響システムから「電話の電源を切ってください」というアナウンスなどの音を受信したとき、他の音（たとえば、モバイルデバイスの呼出し音）は雑音と見なされ得る。この場合、モバイルデバイス 1 4 3 0 は、キャプチャされた音を分離した後、シアター音響システムからの音を選択し、モバイルフォンからの他の音を無視し得る。1 5 4 0 において、モバイルデバイス 1 4 3 0 は、選択されたメディア音から音特徴を抽出する。1 5 5 0 において、抽出された特徴をサーバに送信する。モバイルデバイス 1 4 3 0 からの音特徴に基づいて、サーバは、1 5 6 0 において、モバイルデバイス 1 4 3 0 において実行されるべき機能を決定し、機能を識別する情報をモバイルデバイス 1 4 3 0 に送る。サーバから、機能を識別する情報を受信した後、モバイルデバイス 1 4 3 0 は、1 5 7 0 において機能を実行する。

【 0 0 7 4 】

[0088] 図 1 6 に、いくつかの他の実施形態による、複数のメディア音に応答して少なくとも 1 つの機能を実行するためのモバイルデバイス 1 4 3 0 における方法 1 6 0 0 のフローチャートを示す。この方法 1 6 0 0 では、モバイルデバイス 1 4 3 0 は、1 6 1 0 において、音センサーを通して、複数のメディア音を含む合成音をキャプチャする。次いで、モバイルデバイス 1 4 3 0 は、1 6 2 0 において、キャプチャされた合成音から複数のメディア音を分離する。合成音を分離した後、モバイルデバイス 1 4 3 0 は、1 6 3 0 において、特にいずれか 1 つの音を選択することなしに、分離されたメディア音の各々から音特徴を抽出する。モバイルデバイス 1 4 3 0 は、1 6 4 0 において、抽出された音特徴をサーバに送信する。映画館の例では、モバイルデバイス 1 4 3 0 は、シアター音響システムの音から音特徴を抽出し、モバイルデバイスの呼出し音から音特徴を抽出し、両方の音特徴をサーバに送り得る。次いで、サーバは、受信した音特徴を音特徴データベースに記憶された基準音特徴と比較し、抽出された音特徴に関連する少なくとも 1 つの基準音特徴

10

20

30

40

50

を識別する。基準音特徴との一致を識別した後、サーバは、1650において、識別された少なくとも1つの基準音特徴に関連する少なくとも1つの機能を識別する情報をモバイルデバイス1430に送る。1660において、モバイルデバイス1430は当該少なくとも1つの機能を実行する。

【0075】

[0089]図17に、いくつかの実施形態による、音識別に基づいて機能を制御するための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におけるモバイルデバイス1700のブロック図を示す。モバイルデバイス1700は、セルラーフォン、端末、ハンドセット、携帯情報端末(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、符号分割多元接続(CDMA)システム、モバイル通信用グローバルシステム(GSM(登録商標): Global System for Mobile Communications)システム、広帯域CDMA(W-CDMA)(登録商標)システム、ロングタームエボリューション(LTE: Long Tern Evolution)システム、LTE Advancedシステムなどであり得る。

10

【0076】

[0090]モバイルデバイス1700は、受信経路および送信経路を介して双方向通信を行うことが可能であり得る。受信経路上では、基地局によって送信された信号は、アンテナ1712によって受信され、受信機(RCVR)1714に与えられる。受信機1714は、受信信号を調整し、デジタル化し、さらなる処理のために調整およびデジタル化された信号をデジタルセクション1720に与える。送信経路上では、送信機(TMTX)は、デジタルセクション1720から送信されるべきデータを受信し、データを処理し、調整し、被変調信号を生成し、被変調信号はアンテナ1712を介して基地局に送信される。受信機1714および送信機1716は、CDMA、GSM、W-CDMA、LTE、LTE Advancedなどをサポートするトランシーバの一部である。

20

【0077】

[0091]デジタルセクション1720は、たとえば、モデムプロセッサ1722、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)1724、コントローラ/プロセッサ1726、内部メモリ1728、一般化オーディオエンコーダ1732、一般化オーディオデコーダ1734、グラフィックス/ディスプレイプロセッサ1736、および/または外部バスインターフェース(EBI)1738など、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む。モデムプロセッサ1722は、データ送信および受信のための処理、たとえば、符号化、変調、復調、および復号を実行する。RISC/DSP1724は、モバイルデバイス1700のための一般的専用処理を実行する。コントローラ/プロセッサ1726は、デジタルセクション1720内の様々な処理およびインターフェースユニットの演算を制御する。内部メモリ1728は、デジタルセクション1720内の様々なユニットのためのデータおよび/または命令を記憶する。

30

【0078】

[0092]一般化オーディオエンコーダ1732は、オーディオソース1742、マイクロフォン1743などからの入力信号に対して符号化を実行する。一般化オーディオデコーダ1734は、コード化オーディオデータに対して復号を実行し、出力信号をスピーカ/ヘッドセット1744に与える。一般化オーディオエンコーダ1732および一般化オーディオデコーダ1734は、必ずしも、オーディオソース、マイクロフォン1743およびスピーカ/ヘッドセット1744とのインターフェースのために必要とされるとは限らず、したがって、モバイルデバイス1700に示されていないことに留意されたい。グラフィックス/ディスプレイプロセッサ1736は、ディスプレイユニット1746に提示されるグラフィックス、ビデオ、画像、およびテキストのための処理を実行する。外部バスインターフェース(EBI)1738は、デジタルセクション1720とメインメモリ1748との間のデータの転送を可能にする。

40

【0079】

[0093]デジタルセクション1720は、1つまたは複数のプロセッサ、DSP、マイク

50

ロプロセッサ、RISCなどを用いて実装される。デジタルセクション1720はまた、1つまたは複数の特定用途向け集積回路(ASIC)および/または何らかの他のタイプの集積回路(IC)上に作製される。

【0080】

[0094]概して、本明細書で説明したいかなるデバイスも、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(PC)カード、PDA、外部または内部モデム、ワイヤレスチャネルを介して通信するデバイスなど、様々なタイプのデバイスを示す。デバイスは、アクセス端末(AT)、アクセスユニット、加入者ユニット、移動局、クライアントデバイス、モバイルユニット、モバイルフォン、モバイル、リモート局、リモート端末、リモートユニット、ユーザデバイス、ユーザ機器、ハンドヘルドデバイスなど、様々な名前を有し得る。本明細書で説明したいかなるデバイスも、命令およびデータ、ならびにハードウェア、ソフトウェア、ファームウェア、またはそれらの組合せを記憶するためのメモリを有し得る。

【0081】

[0095]図18は、いくつかの実施形態に従って実装される音識別に基づいて機能を制御するための、前に説明し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1800を示すブロック図である。サーバシステム1800は、1つまたは複数の処理ユニット(CPU)1802と、1つまたは複数のネットワークまたは他の通信インターフェース1810と、メモリ1812と、これらの構成要素を相互接続するための1つまたは複数の通信バス1814とを含み得る。サーバシステム1800はまた、ディスプレイデバイスとキーボードとを有するユーザインターフェース(図示せず)を含み得る。

【0082】

[0096]メモリ1812は、高速ランダムアクセスメモリ(たとえば、DRAM、SRAM、DDR RAMまたは他のランダムアクセス固体メモリデバイス)など、任意の好適なメモリであり得る。メモリ1812は、不揮発性メモリ(たとえば、1つまたは複数の磁気ディスクストレージデバイス、光ディスクストレージデバイス、フラッシュメモリデバイス、または他の不揮発性固体ストレージデバイス)を含み得るか、または代替的に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1812は、(1つまたは複数の)CPU1802から遠隔に位置する、および/または複数のサイト中に遠隔に位置する1つまたは複数のストレージデバイスを含み得る。

【0083】

[0097]メモリ1812によって表される上記のメモリデバイスのうちのいずれか1つが、前に説明したプロセス、動作、および方法のいずれかを実行および/または履行する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。たとえば、メモリ1812は、様々な基本システムサービスを処理するためのプロシージャと、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と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1816を含み得る。メモリ1812のネットワーク通信モジュール1818は、1つまたは複数の通信ネットワークインターフェース1810(ワイヤードまたはワイヤレス)と、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、メトロポリタンエリアネットワークなど、1つまたは複数の通信ネットワークとを介して、サーバシステム1800を他のコンピュータに接続するために使用され得る。メモリ1812はまた、機能が関連するメディア音の基準音特徴から機能を決定するように構成された機能決定エンジン1844と、音特徴を抽出するための基礎フィルタを生成し、更新するように構成された基礎フィルタ生成器1850とを含み得る。機能決定エンジン1844および基礎フィルタ生成器1850は、基準音特徴データ、基礎フィルタデータ、および/またはサンプル音データを記憶するように構成されたデータベース1848にアクセスし得る。

【0084】

[0098]上記の識別されたモジュールまたはプログラム(すなわち、命令のセット)は、

10

20

30

40

50

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、したがって、これらのモジュールの様々なサブセットは、様々な実施形態において組み合わせられるか、あるいは並べ替えられ得ることを諒解されよう。さらに、メモリ 1812 は、上記で説明し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。

【0085】

[0099] 図 17 および図 18 は、本明細書で説明する実施形態の構造概略図というよりむしろ、クライアントシステムおよびサーバシステムの様々な特徴の機能説明として意図されている。実際には、当業者によって認識されるように、別々に示されたアイテムが組み合わせられ得、いくつかのアイテムが分離され得る。たとえば、図 18 で別々に示されたいくつかのアイテムは、単一のサーバ上で実装され得、単一のアイテムは、1 つまたは複数のサーバによって実装され得る。さらに、データベース 1848 は、サーバの異なるセット上で実装されるか、またはサーバシステム 1800 の他の構成要素において実装され得る。サーバシステム 1800 を実装するために使用されるサーバの実際の数、およびそれらの間の特徴の割振りは、実装ごとに異なり得る。

【0086】

[00100] 本明細書で説明した技法は様々な方法によって実装される。たとえば、これらの技法は、ハードウェア、ファームウェア、ソフトウェア、またはそれらの組合せで実装される。さらに、本明細書の開示に関連して説明した様々な例示的な論理ブロック、モジュール、回路、およびアルゴリズムステップは、電子ハードウェア、コンピュータソフトウェア、または両方の組合せが可能であるように実装され得ることを諒解されよう。ハードウェアとソフトウェアのこの互換性を明確に示すために、様々な例示的な構成要素、ブロック、モジュール、回路、およびステップについて、上記では概してそれらの機能に関して説明した。そのような機能をハードウェアとして実装するか、ソフトウェアとして実装するかは、特定の適用例および全体的なシステムに課された設計制約に依存する。説明した機能は、特定の適用例ごとに様々な方法で実装され得るが、そのような実装の決定は、本開示の範囲からの逸脱を生じるものと解釈すべきではない。

【0087】

[00101] ハードウェア実装の場合、本技法を実行するために使用される処理ユニットは、1 つまたは複数の 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、本明細書で説明した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、またはそれらの組合せ内で実装され得る。

【0088】

[00102] ソフトウェアで実装した場合、前に説明した方法、動作、およびプロセスは、1 つまたは複数の命令またはコードとしてコンピュータ可読媒体上に記憶され得るか、あるいはコンピュータ可読媒体を介して送信され得る。コンピュータ可読媒体は、ある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可能にする任意の媒体を含む、コンピュータ記憶媒体とコンピュータ通信媒体の両方を含む。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。限定ではなく例として、そのようなコンピュータ可読媒体は、RAM、ROM、EEPROM (登録商標)、CD-ROM、あるいは他の光ディスクストレージ、磁気ディスクストレージまたは他の磁気記憶デバイス、もしくは命令またはデータ構造の形態の所望のプログラムコードを搬送または記憶するために使用でき、コンピュータによってアクセスできる任意の他の媒体を備えることができる。また、いかなる接続もコンピュータ可読媒体と適切に呼ばれる。たとえば、ソフトウェアが、同軸ケーブル、光ファイバケーブル、ツイストペア、デジタル加入者回線 (DSL)、または赤外線、無線、およびマイクロ波などのワイヤレス技術を使用して、ウェブサイト、サーバ、または他のリモートソースから送信される場合、同軸ケーブル、光ファイバケーブル、ツイストペア、DSL、または赤外線、無線、およびマイクロ波などのワイヤレス技術は、媒体の定義に含まれる。本明細書で使用するディスク (disk

）およびディスク（disc）は、コンパクトディスク（disc）（CD）、レーザーディスク（登録商標）（disc）、光ディスク（disc）、デジタル多用途ディスク（disc）（DVD）、フロッピー（登録商標）ディスク（disk）およびblu-ray（登録商標）ディスク（disc）を含み、ディスク（disk）は、通常、データを磁氣的に再生し、ディスク（disc）は、データをレーザーで光学的に再生する。上記の組合せもコンピュータ可読媒体の範囲内に含まれるべきである。

【0089】

[00103]本開示についての以上の説明は、当業者が本開示を作成または使用することができるように与えたものである。本開示への様々な修正は当業者には容易に明らかとなり、本明細書で定義した一般原理は、本開示の趣旨または範囲から逸脱することなく他の変形形態に適用される。したがって、本開示は、本明細書で説明した例に限定されるものではなく、本明細書で開示した原理および新規の特徴に合致する最も広い範囲を与えられるべきである。

【0090】

[00104]主題について構造的機能および／または方法論的行為に特有の言語で説明したが、添付の特許請求の範囲において定義される主題は、必ずしも上記で説明した特定の特徴または行為に限定されとは限らないことを理解されたい。むしろ、上記の特定の特徴および行為は、特許請求の範囲を実施することの例示的な形態として開示される。

【図1】

図1

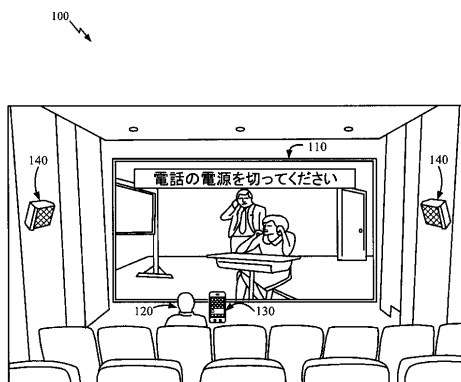


FIG. 1

【図2】

図2

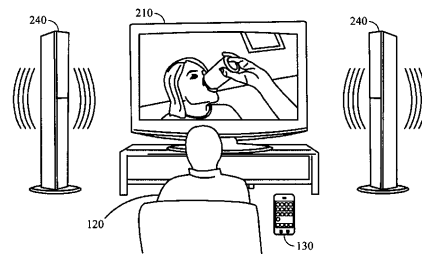


FIG. 2

【図 3】

図 3

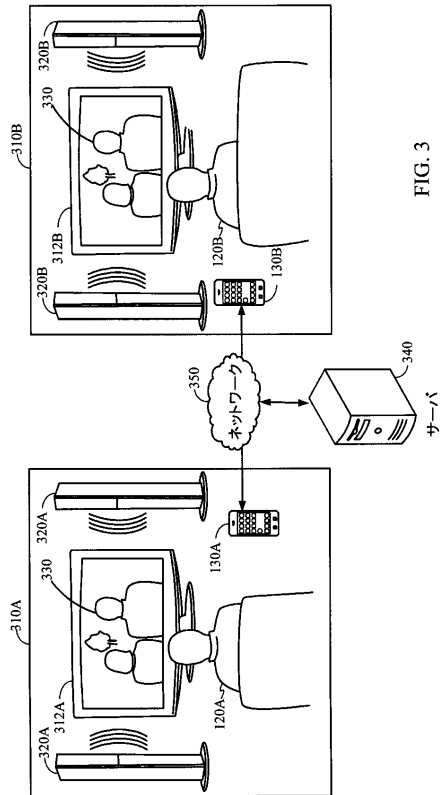


FIG. 3

【図 4】

図 4

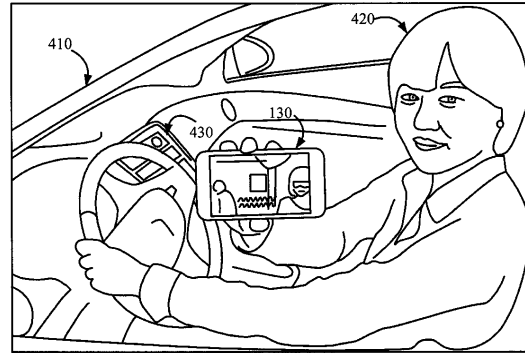


FIG. 4

【図 5】

図 5

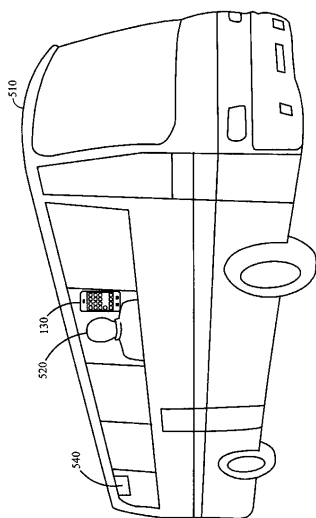


FIG. 5

【図 6 A】

図 6A

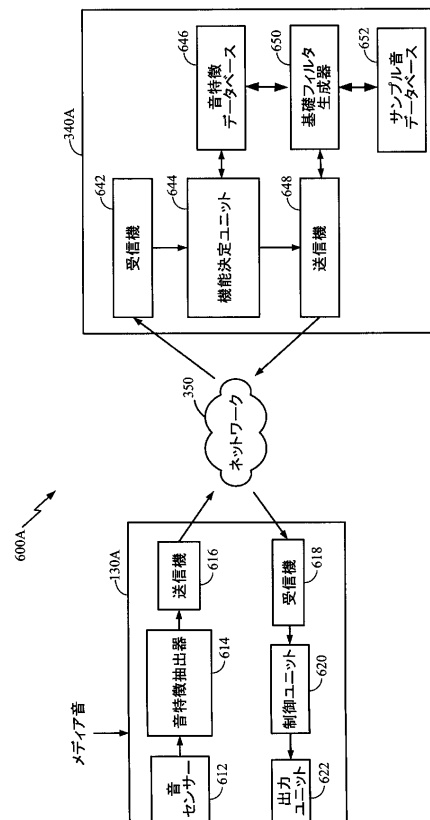


FIG. 6A

【図 6 B】

図 6B

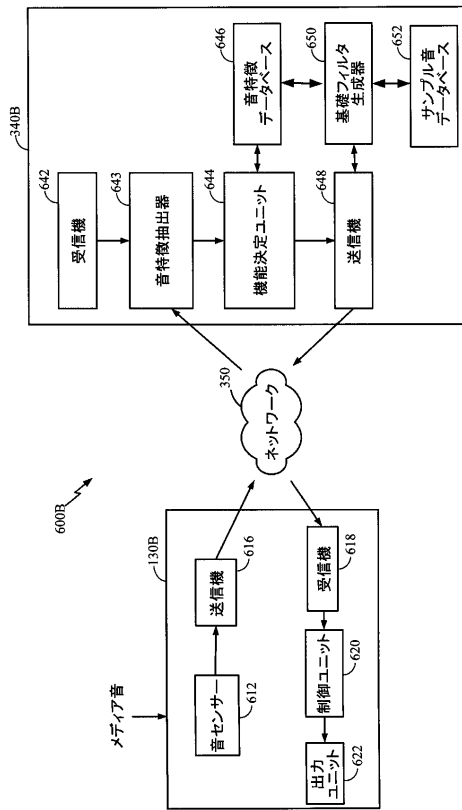


FIG. 6B

【図 7】

図 7

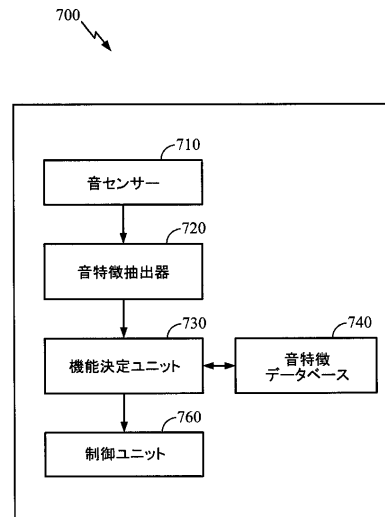


FIG. 7

【図 8】

図 8

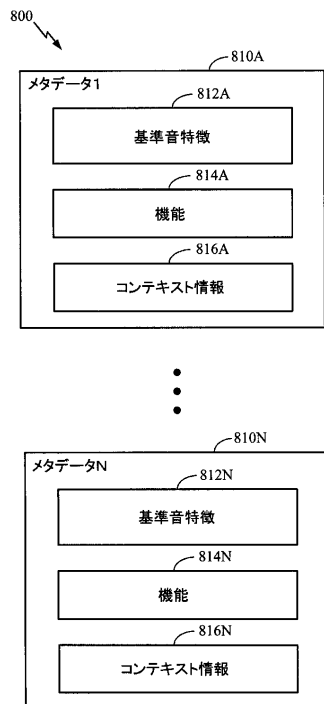


FIG. 8

【図 9】

図 9

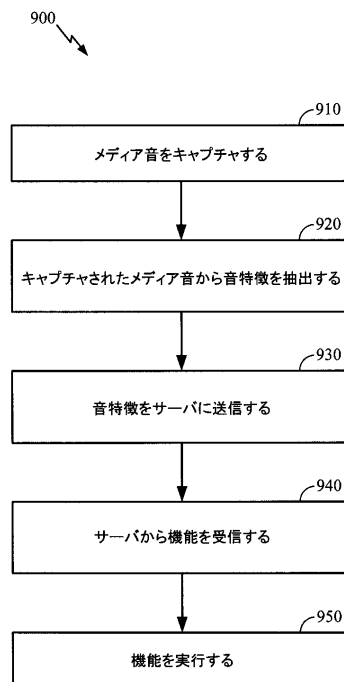


FIG. 9

【図 10】

図 10

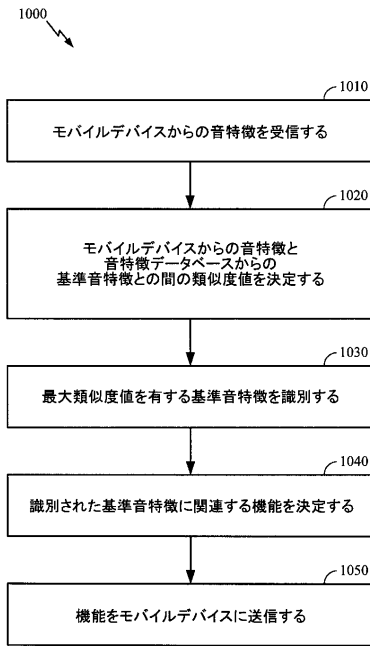


FIG. 10

【図 11】

図 11

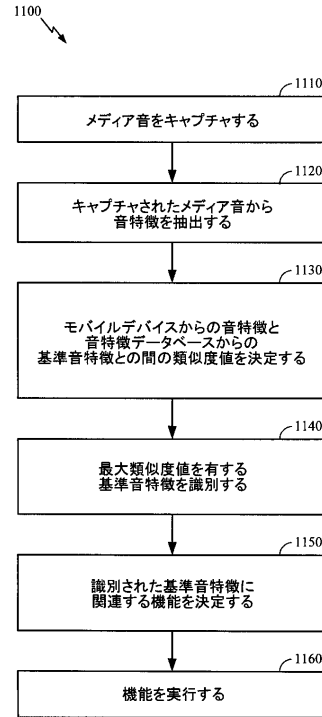


FIG. 11

【図 12】

図 12

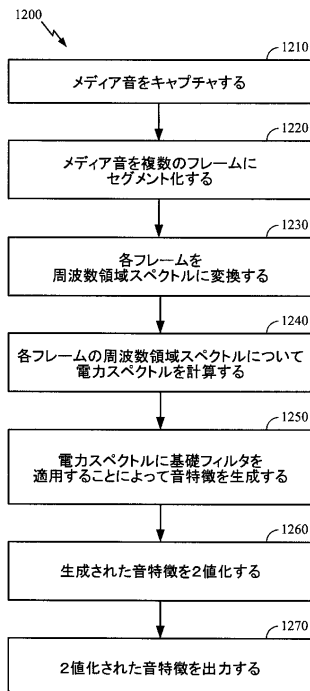


FIG. 12

【図 13】

図 13

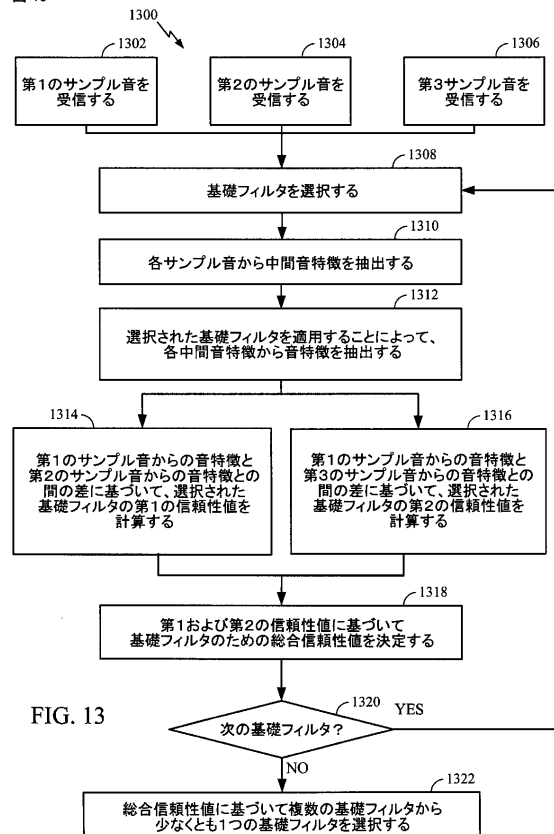


FIG. 13

【図 14】

図 14

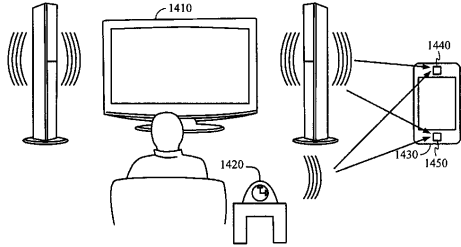


FIG. 14

【図 15】

図 15

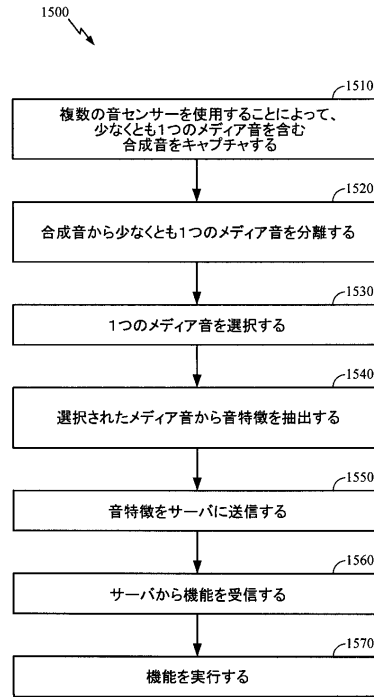


FIG. 15

【図 16】

図 16

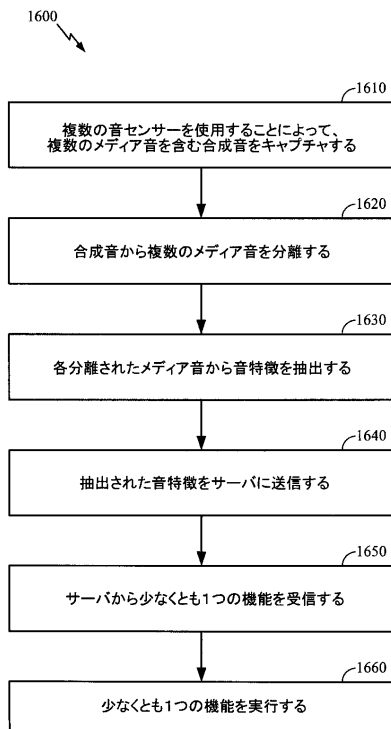


FIG. 16

【図 17】

図 17

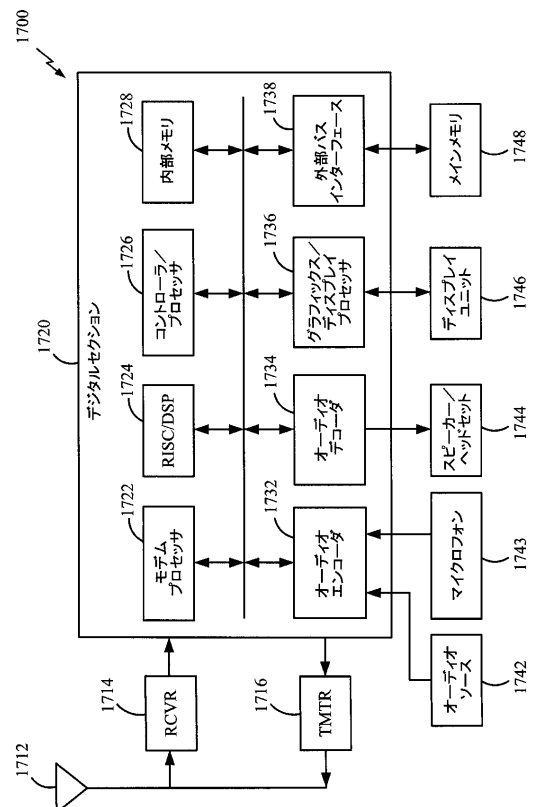


FIG. 17

【図 18】

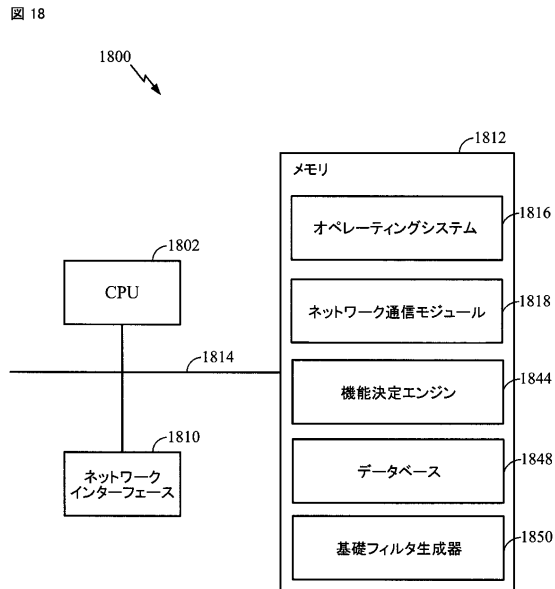


FIG. 18

【手続補正書】

【提出日】平成26年9月24日(2014.9.24)

【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】

モバイルデバイスにおいて機能を実行するための方法であって、

前記モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャすることと、

前記キャプチャされたメディア音から音特徴を抽出することと、

前記抽出された音特徴に基づいて、基準音特徴のセット中の少なくとも1つの基準音特徴を識別することによって、前記モバイルデバイスにおいて実行されるべき前記機能を決定することと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも1つと、複数の機能のうちの少なくとも1つとに関連し、

前記モバイルデバイスにおいて前記決定された機能を実行することと、前記少なくとも1つの基準音特徴は広告に関連し、前記決定された機能を実行することは、前記モバイルデバイスにおいて前記広告に基づいて販売促進情報を表示することを含む、

を備える、方法。

【請求項 2】

前記メディア音が環境音に対応する、請求項 1 に記載の方法。

【請求項 3】

前記少なくとも1つの基準音特徴は、検索エンジンによって検索されるべきキーワード

に関連し、前記決定された機能を実行することは、
前記キーワードを前記検索エンジンに送信することと、
前記検索エンジンから検索結果を受信することと、
前記モバイルデバイス上に前記検索結果を出力することと、
を含む、請求項 1 に記載の方法。

【請求項 4】

前記決定された機能を実行することは、前記モバイルデバイスの構成を調整することを含む、請求項 1 に記載の方法。

【請求項 5】

前記決定された機能を実行することは、
実行されるべき前記機能の通知を出力することと、
前記機能を実行するための入力を受信することと、
前記受信された入力に応答して前記機能を実行することと、
を含む、請求項 1 に記載の方法。

【請求項 6】

前記キャプチャされたメディア音の音レベルが所定の音レベルよりも大きい場合、前記キャプチャされたメディア音から前記音特徴を抽出することと、前記モバイルデバイスによって実行されるべき前記機能を決定することとが実行される、請求項 1 に記載の方法。

【請求項 7】

前記キャプチャされたメディア音から前記音特徴を抽出することは、
前記キャプチャされたメディア音から中間音特徴を抽出することと、
前記音特徴を生成するために、前記中間音特徴に少なくとも 1 つの基礎フィルタを適用することと、
を含む、請求項 1 に記載の方法。

【請求項 8】

前記少なくとも 1 つの基礎フィルタは、複数の基礎フィルタから、前記複数の基礎フィルタの各々によって 1 ペアのサンプル音から抽出された音特徴間の類似度の程度に基づいて、選択される、請求項 7 に記載の方法。

【請求項 9】

前記少なくとも 1 つの基準音特徴は、前記モバイルデバイス上で実行されるべき前記機能とコンテキスト情報とを含むメタデータに関連する、請求項 1 に記載の方法。

【請求項 10】

モバイルデバイスにおいて機能を実行するための方法であって、
複数の音センサーによって、少なくとも 1 つのメディア音を含む合成音をキャプチャすることと、

前記合成音から前記少なくとも 1 つのメディア音を分離することと、

前記少なくとも 1 つの分離されたメディア音から少なくとも 1 つの音特徴を抽出することと、

前記少なくとも 1 つの抽出された音特徴に基づいて、基準音特徴のセット中の少なくとも 1 つの基準音特徴を識別することによって、前記モバイルデバイスにおいて実行されるべき前記機能を決定することと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも 1 つと、複数の機能のうちの少なくとも 1 つとに関連し、

前記モバイルデバイスにおいて前記決定された機能を実行することと、前記少なくとも 1 つの基準音特徴は広告に関連し、前記決定された機能を実行することは、前記モバイルデバイスにおいて前記広告に基づいて販売促進情報を表示することを含む、

を備える、方法。

【請求項 11】

前記少なくとも 1 つのメディア音が環境音に対応する、請求項 10 に記載の方法。

【請求項 12】

前記少なくとも 1 つの基準音特徴は、検索エンジンによって検索されるべきキーワード

に関連し、前記決定された機能を実行することは、
前記キーワードを前記検索エンジンに送信することと、
前記検索エンジンから検索結果を受信することと、
前記モバイルデバイス上に前記検索結果を出力することと、
を含む、請求項 10 に記載の方法。

【請求項 13】

前記決定された機能を実行することは、前記モバイルデバイスの構成を調整することを含む、請求項 10 に記載の方法。

【請求項 14】

前記決定された機能を実行することは、
実行されるべき前記機能の通知を出力することと、
前記機能を実行するための入力を受信することと、
前記受信された入力に応答して前記機能を実行することと、
を含む、請求項 10 に記載の方法。

【請求項 15】

前記少なくとも 1 つの分離されたメディア音から前記少なくとも 1 つの音特徴を抽出することは、
前記少なくとも 1 つの分離されたメディア音から 1 つのメディア音を選択することと、
前記選択されたメディア音から音特徴を抽出することと、
を含む、請求項 10 に記載の方法。

【請求項 16】

前記合成音から前記少なくとも 1 つのメディア音を分離することは、
前記モバイルデバイスのロケーションを決定することと、
特定の音が一般的に前記ロケーションに関連するかどうかを決定することと、
前記特定の音が一般的に前記ロケーションに関連すると決定したことに応答して、前記
合成音から前記特定の音を分離することと、前記特定の音が前記少なくとも 1 つのメディア
音に対応する、
を備える、請求項 10 に記載の方法。

【請求項 17】

モバイルデバイスによって実行されるべき機能を決定するためのサーバであって、
音特徴を受信するように構成された受信機と、前記音特徴は、前記モバイルデバイスによってキャプチャされたメディア音から抽出され、
前記音特徴に基づいて、基準音特徴のセット中の少なくとも 1 つの基準音特徴を識別するように構成された識別ユニットと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも 1 つと、複数の機能のうちの少なくとも 1 つとに関連し、
前記少なくとも 1 つの基準音特徴は広告に関連し、
前記少なくとも 1 つの識別された基準音特徴に基づいて前記モバイルデバイスによって
実行されるべき前記機能を決定するように構成された機能決定ユニットと、前記機能決定
ユニットは、前記モバイルデバイスにおいて前記広告に基づいて販売促進情報を表示する
ような前記機能を決定するように構成される、
を備える、サーバ。

【請求項 18】

前記メディア音が環境音に対応する、請求項 17 に記載のサーバ。

【請求項 19】

前記少なくとも 1 つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記機能決定ユニットは、
前記キーワードを前記検索エンジンに送信し、
前記検索エンジンから検索結果を受信し、
前記モバイルデバイス上に前記検索結果を出力するような
前記機能を決定するように構成された、請求項 17 に記載のサーバ。

【請求項 20】

前記機能決定ユニットは、前記モバイルデバイスの構成を調整するような前記機能を決定するように構成された、請求項 17 に記載のサーバ。

【請求項 21】

前記機能決定ユニットは、
実行されるべき前記機能の通知を出力し、
前記機能を実行するための入力を受信し、
前記受信された入力に応答して前記機能を実行するような
前記機能を決定するように構成された、請求項 17 に記載のサーバ。

【請求項 22】

前記音特徴は、前記キャプチャされたメディア音の音レベルが所定の音レベルよりも大きい場合、前記キャプチャされたメディア音から抽出される、請求項 17 に記載のサーバ。

【請求項 23】

前記音特徴は、前記キャプチャされたメディア音から中間音特徴を抽出することと、前記音特徴を生成するために前記中間音特徴に少なくとも 1 つの基礎フィルタを適用することによって、前記キャプチャされたメディア音から抽出される、請求項 17 に記載のサーバ。

【請求項 24】

複数の基礎フィルタから、前記複数の基礎フィルタの各々によって 1 ペアのサンプル音から抽出された音特徴間の類似度の程度に基づいて、前記少なくとも 1 つの基礎フィルタを選択するように構成された基礎フィルタ生成器をさらに備える、請求項 23 に記載のサーバ。

【請求項 25】

前記少なくとも 1 つの基準音特徴は、前記モバイルデバイスによって実行されるべき前記機能とコンテキスト情報とを含むメタデータに関連する、請求項 17 に記載のサーバ。

【請求項 26】

モバイルデバイスであって、
前記モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャするように構成された音センサーと、
前記キャプチャされたメディア音から音特徴を抽出するように構成された音特徴抽出器と、
前記音特徴に基づいて基準音特徴のセット中の少なくとも 1 つの基準音特徴を識別するように構成された識別ユニットと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも 1 つと、複数の機能のうちの少なくとも 1 つとに関連し、
前記少なくとも 1 つの識別された基準音特徴に基づいて前記モバイルデバイスによって実行されるべき機能を決定するように構成された機能決定ユニットと、
前記モバイルデバイスにおいて前記決定された機能を実行するように構成された制御ユニットと、前記少なくとも 1 つの基準音特徴は広告に関連し、前記制御ユニットは、前記モバイルデバイスにおいて前記広告に基づいて販売促進情報を表示するように構成される

を備える、モバイルデバイス。

【請求項 27】

前記メディア音が環境音に対応する、請求項 26 に記載のモバイルデバイス。

【請求項 28】

前記少なくとも 1 つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記制御ユニットは、
前記キーワードを前記検索エンジンに送信することと、
前記検索エンジンから検索結果を受信することと、
前記モバイルデバイス上に前記検索結果を出力することと、

を行うように構成された、請求項 26 に記載のモバイルデバイス。

【請求項 29】

前記制御ユニットは、前記モバイルデバイスの構成を調整するように構成された、請求項 26 に記載のモバイルデバイス。

【請求項 30】

前記制御ユニットは、
実行されるべき前記機能の通知を出力することと、
前記機能を実行するための入力を受信することと、
前記受信された入力に応答して前記機能を実行することと、
を行うように構成された、請求項 26 に記載のモバイルデバイス。

【請求項 31】

基準音特徴の前記セット中の各基準音特徴は、スピーカーから導出された音に関連し、特定の音特徴に関連する特定のキャプチャされたメディア音がスピーカーからキャプチャされない場合、前記音特徴抽出器から抽出された前記特定の音特徴は、前記識別ユニットによって識別されない、請求項 26 に記載のモバイルデバイス。

【請求項 32】

モバイルデバイスにおいて機能を実行するための命令を記憶す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、

前記モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャすることと、

前記キャプチャされたメディア音から音特徴を抽出することと、

前記抽出された音特徴に基づいて基準音特徴のセット中の少なくとも 1 つの基準音特徴を識別することによって、前記モバイルデバイスにおいて実行されるべき前記機能を決定することと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも 1 つと、複数の機能のうちの少なくとも 1 つとに関連し、

前記モバイルデバイスにおいて前記決定された機能を実行することと、前記少なくとも 1 つの基準音特徴は広告に関連し、前記決定された機能を実行することは、前記モバイルデバイスにおいて前記広告に基づいて販売促進情報を表示することを含む、

の動作をプロセッサに実行させる、非一時的コンピュータ可読記憶媒体。

【請求項 33】

前記メディア音が環境音に対応する、請求項 32 に記載の媒体。

【請求項 34】

前記少なくとも 1 つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記決定された機能を実行することは、

前記キーワードを前記検索エンジンに送信することと、

前記検索エンジンから検索結果を受信することと、

前記モバイルデバイス上に前記検索結果を出力することと、

を含む、請求項 32 に記載の媒体。

【請求項 35】

前記決定された機能を実行することは、前記モバイルデバイスの構成を調整することを含む、請求項 32 に記載の媒体。

【請求項 36】

前記決定された機能を実行することは、

実行されるべき前記機能の通知を出力することと、

前記機能を実行するための入力を受信することと、

前記受信された入力に応答して前記機能を実行することと、

を含む、請求項 32 に記載の媒体。

【請求項 37】

前記キャプチャされたメディア音の音レベルが所定の音レベルよりも大きい場合、前記キャプチャされたメディア音から前記音特徴を抽出することと、前記モバイルデバイスに

よって実行されるべき前記機能を決定することとが実行される、請求項 3 2 に記載の媒体。

【請求項 3 8】

前記キャプチャされたメディア音から前記音特徴を抽出することは、
前記キャプチャされたメディア音から中間音特徴を抽出することと、
前記音特徴を生成するために前記中間音特徴に少なくとも 1 つの基礎フィルタを適用することと、
を含む、請求項 3 2 に記載の媒体。

【請求項 3 9】

前記少なくとも 1 つの基礎フィルタは、複数の基礎フィルタから、前記複数の基礎フィルタの各々によって 1 ペアのサンプル音から抽出された音特徴間の類似度の程度に基づいて、選択される、請求項 3 8 に記載の媒体。

【請求項 4 0】

前記少なくとも 1 つの基準音特徴は、前記モバイルデバイスにおいて実行されるべき前記機能とコンテキスト情報とを含むメタデータに関連する、請求項 3 2 に記載の媒体。

【請求項 4 1】

モバイルデバイスによって実行されるべき機能を決定するためのサーバであって、
音特徴を受信する手段と、前記音特徴は、前記モバイルデバイスによってキャプチャされたメディア音から抽出され、
前記音特徴に基づいて、基準音特徴のセット中の少なくとも 1 つの基準音特徴を識別する手段と、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも 1 つと、複数の機能のうちの少なくとも 1 つとに関連し、
前記少なくとも 1 つの識別された基準音特徴に基づいて前記モバイルデバイスによって実行されるべき前記機能を決定する手段と、前記少なくとも 1 つの基準音特徴は広告に関連し、前記決定する手段は、前記モバイルデバイスにおいて前記広告に基づいて販売促進情報を表示するような前記機能を決定するように構成される、
を備える、サーバ。

【請求項 4 2】

前記メディア音が環境音に対応する、請求項 4 1 に記載のサーバ。

【請求項 4 3】

前記少なくとも 1 つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記決定する手段は、
前記キーワードを前記検索エンジンに送信し、
前記検索エンジンから検索結果を受信し、
前記モバイルデバイス上に前記検索結果を出力するような
前記機能を決定するように構成された、請求項 4 1 に記載のサーバ。

【請求項 4 4】

前記決定する手段は、前記モバイルデバイスの構成を調整するような前記機能を決定するように構成された、請求項 4 1 に記載のサーバ。

【請求項 4 5】

前記決定する手段は、
実行されるべき前記機能の通知を出力し、
前記機能を実行するための入力を受信し、
前記受信された入力に応答して前記機能を実行するような
前記機能を決定するように構成された、請求項 4 1 に記載のサーバ。

【請求項 4 6】

前記音特徴は、前記キャプチャされたメディア音の音レベルが所定の音レベルよりも大きい場合、前記キャプチャされたメディア音から抽出される、請求項 4 1 に記載のサーバ。

【請求項 4 7】

前記音特徴は、前記キャプチャされたメディア音から中間音特徴を抽出することと、前記音特徴を生成するために前記中間音特徴に少なくとも１つの基礎フィルタを適用することとによって、前記キャプチャされたメディア音から抽出される、請求項４１に記載のサーバ。

【請求項４８】

複数の基礎フィルタから、前記複数の基礎フィルタの各々によって１ペアのサンプル音から抽出された音特徴間の類似度の程度に基づいて、前記少なくとも１つの基礎フィルタを選択する手段をさらに備える、請求項４７に記載のサーバ。

【請求項４９】

前記少なくとも１つの基準音特徴は、前記モバイルデバイスによって実行されるべき前記機能とコンテキスト情報とを含むメタデータに関連する、請求項４１に記載のサーバ。

【請求項５０】

モバイルデバイスであって、

前記モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャする手段と、

前記キャプチャされたメディア音から音特徴を抽出する手段と、

前記音特徴に基づいて、基準音特徴のセット中の少なくとも１つの基準音特徴を識別する手段と、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも１つと、複数の機能のうちの少なくとも１つとに関連し、

前記少なくとも１つの識別された基準音特徴に基づいて前記モバイルデバイスによって実行されるべき機能を決定する手段と、

前記モバイルデバイスにおいて前記決定された機能を実行する手段と、前記少なくとも１つの基準音特徴は広告に関連し、前記実行する手段は、前記広告に基づいて販売促進情報を表示するように構成される、

を備える、モバイルデバイス。

【請求項５１】

前記メディア音が環境音に対応する、請求項５０に記載のモバイルデバイス。

【請求項５２】

前記少なくとも１つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記実行する手段は、

前記キーワードを前記検索エンジンに送信することと、

前記検索エンジンから検索結果を受信することと、

前記モバイルデバイス上に前記検索結果を出力することと

を行うように構成された、請求項５０に記載のモバイルデバイス。

【請求項５３】

前記実行する手段は、前記モバイルデバイスの構成を調整するように構成された、請求項５０に記載のモバイルデバイス。

【請求項５４】

前記実行する手段は、

実行されるべき前記機能の通知を出力することと、

前記機能を実行するための入力を受信することと、

前記受信された入力に応答して前記機能を実行することと、

を行うように構成された、請求項５０に記載のモバイルデバイス。

【手続補正２】

【補正対象書類名】明細書

【補正対象項目名】００９０

【補正方法】変更

【補正の内容】

【００９０】

[00104]主題について構造的機能および／または方法論的行為に特有の言語で説明した

が、添付の特許請求の範囲において定義される主題は、必ずしも上記で説明した特定の特徴または行為に限定されるとは限らないことを理解されたい。むしろ、上記の特定の特徴および行為は、特許請求の範囲を実施することの例示的な形態として開示される。

以下に、本願出願の当初の特許請求の範囲に記載された発明を付記する。

[1] モバイルデバイスにおいて機能を実行するための方法であって、
前記モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャすることと、

前記キャプチャされたメディア音から音特徴を抽出することと、
前記抽出された音特徴に基づいて、基準音特徴のセット中の少なくとも1つの基準音特徴を識別することによって、前記モバイルデバイスにおいて実行されるべき前記機能を決定することと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも1つと、複数の機能のうちの少なくとも1つとに関連し、

前記モバイルデバイスにおいて前記決定された機能を実行することと、
を備える、方法。

[2] 前記少なくとも1つの基準音特徴は広告に関連し、前記決定された機能を実行することは、前記モバイルデバイス上に前記広告を表示することを含む、[1]に記載の方法。

[3] 前記少なくとも1つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記決定された機能を実行することは、
前記キーワードを前記検索エンジンに送信することと、
前記検索エンジンから検索結果を受信することと、
前記モバイルデバイス上に前記検索結果を出力することと、
を含む、[1]に記載の方法。

[4] 前記決定された機能を実行することは、前記モバイルデバイスの構成を調整することを含む、[1]に記載の方法。

[5] 前記決定された機能を実行することは、
実行されるべき前記機能の通知を出力することと、
前記機能を実行するための入力を受信することと、
前記受信された入力に応答して前記機能を実行することと、
を含む、[1]に記載の方法。

[6] 前記キャプチャされたメディア音の音レベルが所定の音レベルよりも大きい場合、前記キャプチャされたメディア音から前記音特徴を抽出することと、前記モバイルデバイスによって実行されるべき前記機能を決定することとが実行される、[1]に記載の方法。

[7] 前記キャプチャされたメディア音から前記音特徴を抽出することは、
前記キャプチャされたメディア音から中間音特徴を抽出することと、
前記音特徴を生成するために、前記中間音特徴に少なくとも1つの基礎フィルタを適用することと、
を含む、[1]に記載の方法。

[8] 前記少なくとも1つの基礎フィルタは、複数の基礎フィルタから、前記複数の基礎フィルタの各々によって1ペアのサンプル音から抽出された音特徴間の類似度の程度に基づいて、選択される、[7]に記載の方法。

[9] 前記少なくとも1つの基準音特徴は、前記モバイルデバイス上で実行されるべき前記機能とコンテキスト情報とを含むメタデータに関連する、[1]に記載の方法。

[10] モバイルデバイスにおいて機能を実行するための方法であって、
複数の音センサーによって、少なくとも1つのメディア音を含む合成音をキャプチャすることと、

前記合成音から前記少なくとも1つのメディア音を分離することと、
前記少なくとも1つの分離されたメディア音から少なくとも1つの音特徴を抽出することと、

前記少なくとも1つの抽出された音特徴に基づいて、基準音特徴のセット中の少なくとも1つの基準音特徴を識別することによって、前記モバイルデバイスにおいて実行されるべき前記機能を決定することと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも1つと、複数の機能のうちの少なくとも1つとに関連し、前記モバイルデバイスにおいて前記決定された機能を実行することと、
を備える、方法。

[1 1] 前記少なくとも1つの基準音特徴は広告に関連し、前記決定された機能を実行することは、前記モバイルデバイス上に前記広告を表示することを含む、[1 0]に記載の方法。

[1 2] 前記少なくとも1つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記決定された機能を実行することは、
前記キーワードを前記検索エンジンに送信することと、
前記検索エンジンから検索結果を受信することと、
前記モバイルデバイス上に前記検索結果を出力することと、
を含む、[1 0]に記載の方法。

[1 3] 前記決定された機能を実行することは、前記モバイルデバイスの構成を調整することを含む、[1 0]に記載の方法。

[1 4] 前記決定された機能を実行することは、
実行されるべき前記機能の通知を出力することと、
前記機能を実行するための入力を受信することと、
前記受信された入力に応答して前記機能を実行することと、
を含む、[1 0]に記載の方法。

[1 5] 前記少なくとも1つの分離されたメディア音から前記少なくとも1つの音特徴を抽出することは、
前記少なくとも1つの分離されたメディア音から1つのメディア音を選択することと、
前記選択されたメディア音から音特徴を抽出することと、
を含む、[1 0]に記載の方法。

[1 6] 前記少なくとも1つの分離されたメディア音から前記少なくとも1つの音特徴を抽出することは、すべての分離されたメディア音から音特徴を抽出することを含む、
[1 0]に記載の方法。

[1 7] モバイルデバイスによって実行されるべき機能を決定するためのサーバであって、

音特徴を受信するように構成された受信機と、前記音特徴は、前記モバイルデバイスによってキャプチャされたメディア音から抽出され、

前記音特徴に基づいて、基準音特徴のセット中の少なくとも1つの基準音特徴を識別するように構成された識別ユニットと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも1つと、複数の機能のうちの少なくとも1つとに関連し、

前記少なくとも1つの識別された基準音特徴に基づいて前記モバイルデバイスによって実行されるべき前記機能を決定するように構成された機能決定ユニットと、
を備える、サーバ。

[1 8] 前記少なくとも1つの基準音特徴は広告に関連し、前記機能決定ユニットは、前記モバイルデバイス上に前記広告を表示するような前記機能を決定するように構成された、[1 7]に記載のサーバ。

[1 9] 前記少なくとも1つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記機能決定ユニットは、
前記キーワードを前記検索エンジンに送信し、
前記検索エンジンから検索結果を受信し、
前記モバイルデバイス上に前記検索結果を出力するような
前記機能を決定するように構成された、[1 7]に記載のサーバ。

[2 0] 前記機能決定ユニットは、前記モバイルデバイスの構成を調整するような前記機能を決定するように構成された、[1 7]に記載のサーバ。

[2 1] 前記機能決定ユニットは、
実行されるべき前記機能の通知を出力し、
前記機能を実行するための入力を受信し、
前記受信された入力に応答して前記機能を実行するような
前記機能を決定するように構成された、[1 7]に記載のサーバ。

[2 2] 前記音特徴は、前記キャプチャされたメディア音の音レベルが所定の音レベルよりも大きい場合、前記キャプチャされたメディア音から抽出される、[1 7]に記載のサーバ。

[2 3] 前記音特徴は、前記キャプチャされたメディア音から中間音特徴を抽出することと、前記音特徴を生成するために前記中間音特徴に少なくとも1つの基礎フィルタを適用することとによって、前記キャプチャされたメディア音から抽出される、[1 7]に記載のサーバ。

[2 4] 複数の基礎フィルタから、前記複数の基礎フィルタの各々によって1ペアのサンプル音から抽出された音特徴間の類似度の程度に基づいて、前記少なくとも1つの基礎フィルタを選択するように構成された基礎フィルタ生成器をさらに備える、[2 3]に記載のサーバ。

[2 5] 前記少なくとも1つの基準音特徴は、前記モバイルデバイスによって実行されるべき前記機能とコンテキスト情報とを含むメタデータに関連する、[1 7]に記載のサーバ。

[2 6] モバイルデバイスであって、
前記モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャするように構成された音センサーと、
前記キャプチャされたメディア音から音特徴を抽出するように構成された音特徴抽出器と、

前記音特徴に基づいて基準音特徴のセット中の少なくとも1つの基準音特徴を識別するように構成された識別ユニットと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも1つと、複数の機能のうちの少なくとも1つとに関連し、

前記少なくとも1つの識別された基準音特徴に基づいて前記モバイルデバイスによって実行されるべき機能を決定するように構成された機能決定ユニットと、

前記モバイルデバイスにおいて前記決定された機能を実行するように構成された制御ユニットと、
を備える、モバイルデバイス。

[2 7] 前記少なくとも1つの基準音特徴は広告に関連し、前記制御ユニットは、前記モバイルデバイス上に前記広告を表示するように構成された、[2 6]に記載のモバイルデバイス。

[2 8] 前記少なくとも1つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記制御ユニットは、

前記キーワードを前記検索エンジンに送信することと、
前記検索エンジンから検索結果を受信することと、
前記モバイルデバイス上に前記検索結果を出力することと、
を行うように構成された、[2 6]に記載のモバイルデバイス。

[2 9] 前記制御ユニットは、前記モバイルデバイスの構成を調整するように構成された、[2 6]に記載のモバイルデバイス。

[3 0] 前記制御ユニットは、
実行されるべき前記機能の通知を出力することと、
前記機能を実行するための入力を受信することと、
前記受信された入力に応答して前記機能を実行することと、
を行うように構成された、[2 6]に記載のモバイルデバイス。

[3 1] 前記音センサーは、所定の時間間隔で環境音を受信することと、前記受信された環境音が前記メディア音であるかどうかを決定することとによって、前記メディア音をキャプチャするようにさらに構成された、[2 6]に記載のモバイルデバイス。

[3 2] モバイルデバイスにおいて機能を実行するための命令を記憶す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、

前記モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャすることと、

前記キャプチャされたメディア音から音特徴を抽出することと、

前記抽出された音特徴に基づいて基準音特徴のセット中の少なくとも1つの基準音特徴を識別することによって、前記モバイルデバイスにおいて実行されるべき前記機能を決定することと、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも1つと、複数の機能のうちの少なくとも1つとに関連し、

前記モバイルデバイスにおいて前記決定された機能を実行することと、

の動作をプロセッサに実行させる、非一時的コンピュータ可読記憶媒体。

[3 3] 前記少なくとも1つの基準音特徴は広告に関連し、前記決定された機能を実行することは、前記モバイルデバイス上に前記広告を表示することを含む、[3 2]に記載の媒体。

[3 4] 前記少なくとも1つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記決定された機能を実行することは、

前記キーワードを前記検索エンジンに送信することと、

前記検索エンジンから検索結果を受信することと、

前記モバイルデバイス上に前記検索結果を出力することと、

を含む、[3 2]に記載の媒体。

[3 5] 前記決定された機能を実行することは、前記モバイルデバイスの構成を調整することを含む、[3 2]に記載の媒体。

[3 6] 前記決定された機能を実行することは、

実行されるべき前記機能の通知を出力することと、

前記機能を実行するための入力を受信することと、

前記受信された入力に応答して前記機能を実行することと、

を含む、[3 2]に記載の媒体。

[3 7] 前記キャプチャされたメディア音の音レベルが所定の音レベルよりも大きい場合、前記キャプチャされたメディア音から前記音特徴を抽出することと、前記モバイルデバイスによって実行されるべき前記機能を決定することとが実行される、[3 2]に記載の媒体。

[3 8] 前記キャプチャされたメディア音から前記音特徴を抽出することは、

前記キャプチャされたメディア音から中間音特徴を抽出することと、

前記音特徴を生成するために前記中間音特徴に少なくとも1つの基礎フィルタを適用することと、

を含む、[3 2]に記載の媒体。

[3 9] 前記少なくとも1つの基礎フィルタは、複数の基礎フィルタから、前記複数の基礎フィルタの各々によって1ペアのサンプル音から抽出された音特徴間の類似度の程度に基づいて、選択される、[3 8]に記載の媒体。

[4 0] 前記少なくとも1つの基準音特徴は、前記モバイルデバイスにおいて実行されるべき前記機能とコンテキスト情報とを含むメタデータに関連する、[3 2]に記載の媒体。

[4 1] モバイルデバイスによって実行されるべき機能を決定するためのサーバであって、

音特徴を受信する手段と、前記音特徴は、前記モバイルデバイスによってキャプチャされたメディア音から抽出され、

前記音特徴に基づいて、基準音特徴のセット中の少なくとも1つの基準音特徴を識別す

る手段と、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも1つと、複数の機能のうちの少なくとも1つとに関連し、

前記少なくとも1つの識別された基準音特徴に基づいて前記モバイルデバイスによって実行されるべき前記機能を決定する手段と、

を備える、サーバ。

[4 2] 前記少なくとも1つの基準音特徴は広告に関連し、前記決定する手段は、前記モバイルデバイス上に前記広告を表示するような前記機能を決定するように構成された、[4 1]に記載のサーバ。

[4 3] 前記少なくとも1つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記決定する手段は、

前記キーワードを前記検索エンジンに送信し、

前記検索エンジンから検索結果を受信し、

前記モバイルデバイス上に前記検索結果を出力するような前記機能を決定するように構成された、[4 1]に記載のサーバ。

[4 4] 前記決定する手段は、前記モバイルデバイスの構成を調整するような前記機能を決定するように構成された、[4 1]に記載のサーバ。

[4 5] 前記決定する手段は、

実行されるべき前記機能の通知を出力し、

前記機能を実行するための入力を受信し、

前記受信された入力に応答して前記機能を実行するような前記機能を決定するように構成された、[4 1]に記載のサーバ。

[4 6] 前記音特徴は、前記キャプチャされたメディア音の音レベルが所定の音レベルよりも大きい場合、前記キャプチャされたメディア音から抽出される、[4 1]に記載のサーバ。

[4 7] 前記音特徴は、前記キャプチャされたメディア音から中間音特徴を抽出することと、前記音特徴を生成するために前記中間音特徴に少なくとも1つの基礎フィルタを適用することによって、前記キャプチャされたメディア音から抽出される、[4 1]に記載のサーバ。

[4 8] 複数の基礎フィルタから、前記複数の基礎フィルタの各々によって1ペアのサンプル音から抽出された音特徴間の類似度の程度に基づいて、前記少なくとも1つの基礎フィルタを選択する手段をさらに備える、[4 7]に記載のサーバ。

[4 9] 前記少なくとも1つの基準音特徴は、前記モバイルデバイスによって実行されるべき前記機能とコンテキスト情報とを含むメタデータに関連する、[4 1]に記載のサーバ。

[5 0] モバイルデバイスであって、

前記モバイルデバイスの外部の音出力デバイスからのメディア音をキャプチャする手段と、

前記キャプチャされたメディア音から音特徴を抽出する手段と、

前記音特徴に基づいて、基準音特徴のセット中の少なくとも1つの基準音特徴を識別する手段と、基準音特徴の前記セット中の各基準音特徴は、複数のメディア音のうちの少なくとも1つと、複数の機能のうちの少なくとも1つとに関連し、

前記少なくとも1つの識別された基準音特徴に基づいて前記モバイルデバイスによって実行されるべき機能を決定する手段と、

前記モバイルデバイスにおいて前記決定された機能を実行する手段と、

を備える、モバイルデバイス。

[5 1] 前記少なくとも1つの基準音特徴は広告に関連し、前記実行する手段は、前記モバイルデバイス上に前記広告を表示するように構成された、[5 0]に記載のモバイルデバイス。

[5 2] 前記少なくとも1つの基準音特徴は、検索エンジンによって検索されるべきキーワードに関連し、前記実行する手段は、

前記キーワードを前記検索エンジンに送信することと、
前記検索エンジンから検索結果を受信することと、
前記モバイルデバイス上に前記検索結果を出力することと
を行うように構成された、[5 0] に記載のモバイルデバイス。

[5 3] 前記実行する手段は、前記モバイルデバイスの構成を調整するように構成された、[5 0] に記載のモバイルデバイス。

[5 4] 前記実行する手段は、
実行されるべき前記機能の通知を出力することと、
前記機能を実行するための入力を受信することと、
前記受信された入力に応答して前記機能を実行することと、
を行うように構成された、[5 0] に記載のモバイルデバイス。

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/023607

A. CLASSIFICATION OF SUBJECT MATTER

INV. H04M1/27
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)
H04M

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Dean A. Rutter: "SIRI A Primer", 31 December 2011 (2011-12-31), XP55061828, Retrieved from the Internet: URL: http://venturewerks.com/Siri-A-Primer.pdf [retrieved on 2013-05-03] page 2, paragraph 1 - page 5, paragraph 1 ----- -/--	1-54

☒ 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 doubt 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

"Z" document member of the same patent family

Date of the actual completion of the international search

6 May 2013

Date of mailing of the international search report

15/05/2013

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

Kim-Mayser, Michael

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/023607

(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>EP 1 176 503 A2 (SONY CORP [JP]) 30 January 2002 (2002-01-30) abstract paragraph [0014] - paragraph [0029] paragraph [0033] paragraph [0037] paragraph [0041] paragraph [0055] - paragraph [0056] paragraph [0068] - paragraph [0071] paragraph [0080] - paragraph [0082] paragraph [0085] - paragraph [0112] paragraph [0117] - paragraph [0239] figures 1-31</p>	1-54
X	<p>US 2006/129637 A1 (YOSHIDA ICHIRO [JP]) 15 June 2006 (2006-06-15) abstract figures 3a,3b,4 paragraph [0002] paragraph [0005] - paragraph [0006] paragraph [0020] - paragraph [0021] paragraph [0025] paragraph [0037] - paragraph [0042]</p>	1-54
A	<p>WO 2011/088053 A2 (APPLE INC [US]; GRUBER THOMAS ROBERT [US]; CHEYER ADAM JOHN [US]; KITT) 21 July 2011 (2011-07-21) the whole document</p>	1-54

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2013/023607

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1176503	A2	30-01-2002	EP 1176503 A2 30-01-2002
		JP 2002041276 A	08-02-2002
		US 2002010589 A1	24-01-2002

US 2006129637	A1	15-06-2006	JP 2006154926 A 15-06-2006
		US 2006129637 A1	15-06-2006

WO 2011088053	A2	21-07-2011	AU 2011205426 A1 23-08-2012
			CA 2787351 A1 21-07-2011
			CA 2791791 A1 21-07-2011
			CA 2792412 A1 21-07-2011
			CA 2792442 A1 21-07-2011
			CA 2792570 A1 21-07-2011
			CA 2793002 A1 21-07-2011
			CA 2793118 A1 21-07-2011
			CA 2793248 A1 21-07-2011
			CA 2793741 A1 21-07-2011
			CA 2793743 A1 21-07-2011
			CN 102792320 A 21-11-2012
			EP 2526511 A2 28-11-2012
			GB 2490444 A 31-10-2012
			KR 20120120316 A 01-11-2012
			KR 20120136417 A 18-12-2012
			KR 20120137424 A 20-12-2012
			KR 20120137425 A 20-12-2012
			KR 20120137434 A 20-12-2012
			KR 20120137435 A 20-12-2012
			KR 20120137440 A 20-12-2012
			KR 20120138826 A 26-12-2012
			KR 20120138827 A 26-12-2012
			KR 20130000423 A 02-01-2013
			US 2012016678 A1 19-01-2012
			US 2012245944 A1 27-09-2012
			US 2013110505 A1 02-05-2013
			US 2013110515 A1 02-05-2013
			US 2013110518 A1 02-05-2013
			US 2013110519 A1 02-05-2013
			US 2013110520 A1 02-05-2013
			US 2013111348 A1 02-05-2013
			US 2013111487 A1 02-05-2013
			WO 2011088053 A2 21-07-2011

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, 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, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, 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, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC

(74)代理人 100153051
弁理士 河野 直樹

(74)代理人 100140176
弁理士 砂川 克

(74)代理人 100158805
弁理士 井関 守三

(74)代理人 100179062
弁理士 井上 正

(74)代理人 100124394
弁理士 佐藤 立志

(74)代理人 100112807
弁理士 岡田 貴志

(74)代理人 100111073
弁理士 堀内 美保子

(72)発明者 キム、テス
アメリカ合衆国、カリフォルニア州 9 2 1 2 1 - 1 7 1 4、サン・ディエゴ、モアハウス・ドライブ 5 7 7 5

(72)発明者 ホワン、キュ・ウォン
アメリカ合衆国、カリフォルニア州 9 2 1 2 1 - 1 7 1 4、サン・ディエゴ、モアハウス・ドライブ 5 7 7 5

(72)発明者 ジン、ミンホ
アメリカ合衆国、カリフォルニア州 9 2 1 2 1 - 1 7 1 4、サン・ディエゴ、モアハウス・ドライブ 5 7 7 5

F ターム(参考) 5K127 AA36 BA03 BA16 CA25 CA36 DA11 DA19 FA09 GA29 GB72

HA10 JA02 JA59 KA02 KA04
5K201 AA05 BA07 BB09 CA01 DC05 EA04 EB06 EC06 ED04