

【公報種別】特許法第 17 条の 2 の規定による補正の掲載
 【部門区分】第 6 部門第 2 区分
 【発行日】平成 27 年 5 月 7 日 (2015.5.7)

【公表番号】特表 2014-515833 (P2014-515833A)
 【公表日】平成 26 年 7 月 3 日 (2014.7.3)
 【年通号数】公開・登録公報 2014-035
 【出願番号】特願 2013-556670 (P2013-556670)
 【国際特許分類】

G 1 0 L 19/018 (2013.01)

G 0 6 T 7/00 (2006.01)

G 1 0 L 21/0272 (2013.01)

【F I】

G 1 0 L 19/018

G 0 6 T 7/00 3 0 0 H

G 1 0 L 21/0272

【手続補正書】

【提出日】平成 27 年 3 月 3 日 (2015.3.3)

【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】

音声信号内の音声ソースを分離するための方法であって、
 コンピュータシステムにおいて、

音声の複数のサブセットから、音声を包含する音声信号を受信するステップであって、
 複数のサブセットの異なるサブセットの音声、異なるソースの特徴を備え、異なるソース
 を起源とすることを特徴とする、受信するステップと、

音声信号を音声データの複数のセグメントに分離するステップと、

複数のセグメントと 1 又はそれ以上の音声フィンガープリントを比較するステップと、
 1 又はそれ以上の音声フィンガープリントと複数のセグメントとの比較に基づいて、音声
 データの複数のサブセットの少なくとも 1 つの第 2 のサブセットに対応する複数のセグメ
 ントの少なくとも 1 つの第 2 のセグメントを、音声信号の複数のサブセットの第 1 のサブ
 セットに対応する複数のセグメントの少なくとも 1 つの第 1 のセグメントから分離するス
 テップと、

を有する事の特徴とする方法。

【請求項 2】

少なくとも 1 つの第 1 のセグメントを包含し、少なくとも 1 つの第 2 のセグメントを実
 質的に欠いている、改善された音声データを出力するステップを
 更に有することを特徴とする請求項 1 に記載の方法。

【請求項 3】

少なくとも 1 つの第 2 のセグメントを少なくとも 1 つの第 1 のセグメントから分離した
 後、改善された音声データを出力する前に、改善された音声データを生成するように複数
 の第 1 のセグメントを結合するステップを
 更に有することを特徴とする請求項 2 に記載の方法。

【請求項 4】

前記改善された音声データを出力するステップが、改善された音声データを包含する音

声ストリームを出力するステップを包含することを特徴とする請求項 2 に記載の方法。

【請求項 5】

少なくとも 1 つの第 2 のセグメントを第 1 のセグメントから分離するステップが、少なくとも 1 つの第 2 のセグメントの音声から実質的に独立した第 1 のセグメントの音声を包含する改善された音声データを得るように、1 又はそれ以上のマスク又はフィルタを複数のセグメントに適用するステップを包含することを特徴とする請求項 1 に記載の方法。

【請求項 6】

前記音声信号を受信するステップが、リアルタイムで音声データを受信することを包含することを特徴とする請求項 1 に記載の方法。

【請求項 7】

前記音声信号を受信するステップが、ライブマイクロフォンからデータを受信するステップを包含することを特徴とする請求項 6 に記載の方法。

【請求項 8】

前記少なくとも 1 つの第 2 のセグメントがノイズからなることを特徴とする請求項 1 に記載の方法。

【請求項 9】

前記音声信号を複数のセグメントに分離するステップが、周波数数列によって少なくとも 1 つのセグメントを識別することを包含することを特徴とする請求項 1 に記載の方法。

【請求項 10】

前記音声信号を複数のセグメントに分離するステップが、ベースラインの上で連続偏差として少なくとも 1 つのセグメントを識別することを包含することを特徴とする請求項 1 に記載の方法。

【請求項 11】

音声信号を受信するステップが、音声信号を

第 1 の声、

第 2 の声、

第 1 の楽器、

第 2 の楽器、

バックグラウンドのノイズ

のいずれかの組み合わせから受信することからなることを特徴とする請求項 1 に記載の方法。

【請求項 12】

音声信号からの異なるソースから音声データのサブセットを分離するための方法を実行するためのコンピュータプログラムをストアするコンピュータ読み取り可能媒体であって、

コンピュータ読み取り可能媒体が、

コンピュータ記憶媒体と、

前記コンピュータ記憶媒体にストアされたコンピュータ実行可能命令と、

を有し、

コンピュータシステムによって実行されるとき、コンピュータ実行可能命令が、コンピュータシステムによって、

音声信号にアクセスすることができ、音声信号が複数の音声のサブセットを集合的に包含し、複数のサブセットの異なるサブセットからの音声、異なるソース特徴を有し、異なるソースを起源とすることを特徴とし、

音声信号を複数の音声データのセグメントに分離することができ、

複数のセグメントを 1 又はそれ以上の音声フィンガープリントと比較し、複数の雪面との欠くセグメントと、1 又はそれ以上の音声フィンガープリントの少なくとも 1 つの対応する音声フィンガープリントとの間の類似性値を決定することができ、

類似性値が複数のサブセットの第 1 のサブセットに対応する可能性があるかどうか対応するセグメントを判断するために、閾値を上回っている各類似性値を比較することがで

き、

改善された音声データを出力することができるように構成され、

改善された音声データが第 1 のサブセットに対応する可能性があるように判断された複数のセグメントのセグメントのアセンブリを包含し、改善された音声データが、第 1 のサブセットに対応する可能性がないように判断されたセグメントを実質的に欠いていることを特徴とするコンピュータ読み取り可能媒体。

【請求項 13】

コンピュータ記憶媒体にストアされた 1 又はそれ以上の音声フィンガープリントを更に有することを特徴とする請求項 12 に記載のコンピュータ読み取り可能媒体。

【請求項 14】

コンピュータシステムによって、コンピュータ実行可能命令が、第 1 のサブセットに対応する可能性があるように判断されたセグメントを包含し、ノイズを実質的に欠くように改善された音声データを出力するようにできるように構成されることを特徴とする請求項 12 に記載のコンピュータ読み取り可能媒体。

【請求項 15】

コンピュータシステムによって、コンピュータ実行可能命令が、第 1 のサブセットに対応する可能性があるように判断されたセグメントを包含し、特定の声に対応する可能性があるように判断されたセグメントを実質的に欠くように改善された音声データを出力するようにできるように構成されることを特徴とする請求項 12 に記載のコンピュータ読み取り可能媒体。

【請求項 16】

複数のセグメントが、音符、進行、または、音節の組み合わせを包含することを特徴とする請求項 12 に記載のコンピュータ読み取り可能媒体。

【請求項 17】

音声信号内の 1 又はそれ以上のサンプルを分離するためのコンピュータシステムであって、

1 又はそれ以上のプロセッサと、

前記 1 又はそれ以上のプロセッサに通信可能に結合されたデータストレージと、

コンピュータ実行可能命令と

を有し、

コンピュータ実行可能命令が 1 又はそれ以上のプロセッサによって実行されるとき、コンピュータシステムによって、

ネットワークを介して電子デバイスからネットワークによって受信される音声データを受信することができ、

音声データを音声データの複数のセグメントにスライスすることができ、

複数のセグメントの少なくともいくつかのセグメントと、音声のソースの特定のサブセットに関連する 1 又はそれ以上の音声フィンガープリントとを比較することができ、

1 又はそれ以上の音声フィンガープリントと少なくともいくつかのセグメントとの比較に基づいて、

特定の音声のサブセットを起源とする可能性があるセグメントとして少なくともいくつかのセグメントのサブセットを分類することができ、

特定のソースのサブセットを起源とする可能性がないその他の少なくともいくつかのセグメントから、特定のソースのサブセットを起源とする可能性のあるセグメントを分離することができ、

特定のソースのサブセットを起源とする可能性があるセグメントに基づいて改善された音声データをアセンブリすることができ、

改善された音声データを出力することができることを特徴とするコンピュータシステム。

【請求項 18】

コンピュータシステムが、ポータブル電子デバイスからなることを特徴とする請求項 17 に記載のコンピュータシステム。

【請求項 19】

コンピュータシステムが、移動電話からなることを特徴とする請求項 17 に記載のコンピュータシステム。

【請求項 20】

部分的な音声信号をアセンブリするための方法であって、

2 又はそれ以上のソースから結合された音声のサブセットを包含する音声信号を受信するステップと、

音声データの複数のセグメントに音声信号を分離するステップであって、複数のセグメントの各セグメントが開始時間と終了時間とを有する事を特徴とする、分離するステップと、

2 又はそれ以上のソースから結合された音声のサブセットの音声の第 1 のサブセットに関連する音声データの複数のフィンガープリントに関連する複数のセグメントの各セグメントに関する類似性値を得るステップと、

複数のセグメントの第 1 のセットのセグメントを形成するステップであって、第 1 のセットが、第 1 のサブセットに関連する音声データの複数のフィンガープリントのうちの 1 つのフィンガープリントの所定の閾値を上回る類似性値を有する複数のセグメントの全てのセグメントを包含することを特徴とする、第 1 のセットのセグメントを形成するステップと、

複数のセグメントの第 2 のセットのセグメントを形成するステップであって、第 2 のセットが、第 1 のセットではない複数のセグメントの少なくともいくつかのセグメントを包含することを特徴とする、第 2 のセットのセグメントを形成するステップと、

アセンブリされた音声信号を出力するステップであって、アセンブリされた音声信号が、第 1 のセットのセグメントを包含し、第 2 のセットを実質的に包含しないことを特徴とする出力するステップと、
を有することを特徴とする方法。

【請求項 21】

音声信号を受信するステップが、音声データにアクセスするステップを包含することを特徴とする請求項 20 に記載の方法。

【請求項 22】

前記音声信号を受信するステップが、別のソースから、2 又はそれ以上の音声のサブセットを包含する音声信号を受信することを包含することを特徴とする請求項 21 に記載の方法。

【請求項 23】

第 2 のセットのセグメントを形成するステップが、第 2 のセットのノイズに相当するセグメントを包含することを包含することを特徴とする請求項 21 に記載の方法。

【請求項 24】

音声の第 1 のサブセットを起源とする可能性があるようにセグメントの第 1 のセットを分類するステップと、

音声の少なくとも 1 つの第 2 のサブセットを起源とする可能性があるようにセグメントの第 2 のセットを分類するステップと、
を更に有することを特徴とする請求項 21 に記載の方法。