【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21年10月8日(2009.10.8)

【公開番号】特開2007-65659(P2007-65659A)
【公開日】平成19年3月15日(2007.3.15)
【年通号数】公開・登録公報2007-010
【出願番号】特願2006-230500(P2006-230500)
【国際特許分類】
Ｇ１０Ｌ　19/00　(2006.01)
Ｇ１０Ｌ　19/02　(2006.01)
Ｇ１０Ｌ　11/00　(2006.01)
【ＦＩ】
Ｇ１０Ｌ　19/00　２３０
Ｇ１０Ｌ　19/02　１９０
Ｇ１０Ｌ　11/00　１０１Ｇ

【手続補正書】
【提出日】平成21年8月25日(2009.8.25)
【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】
オーディオフレームからオーディオ指紋を抽出するための方法であって、
前記オーディオフレームを複数の周波数バンドにフィルタ処理して、対応する複数のフィルタ処理したオーディオ信号を生成するステップと、
前記フィルタ処理したオーディオ信号を非直線時間スケールで再サンプリングするステップと、
前記周波数バンド毎に前記再サンプリングしたオーディオ信号を変換して、前記周波数バンドに対する特徴ベクトルを生成するステップと、
前記特徴ベクトルのセットに基づいて前記オーディオ指紋を計算するステップと、
を具備する方法。
【請求項２】
前記オーディオフレームを複数の周波数バンドにフィルタ処理するステップは、前記複数の周波数バンド毎の前記オーディオフレームをバンドパスフィルタ処理するステップを含む請求項１の方法。
【請求項３】
前記オーディオフレームを複数の周波数バンドにフィルタ処理するステップは、前記オーディオサンプルに高速フーリエ変換（ＦＦＴ）を実行するステップを含む請求項１の方法。
【請求項４】
前記オーディオフレームは、ＭＰ３フォーマットで格納されるオーディオファイルの一部である請求項１の方法。
【請求項５】
前記フィルタ処理したオーディオ信号は、対数時間スケールで再サンプリングされる請求項１の方法。
【請求項６】

前記周波数バンドは、周波数軸で直線的に離間している請求項１の方法。
【請求項７】
前記周波数バンドはオーバーラップしている請求項１の方法。
【請求項８】
特定周波数バンドの、前記再サンプリングしてフィルタ処理したオーディオ信号を変換するステップは、前記再サンプリングしたオーディオ信号に高速フーリエ変換（ＦＦＴ）を実行するステップで構成される請求項１の方法。
【請求項９】
前記オーディオ指紋を計算するステップは、前記周波数バンドに対する前記特徴ベクトルを差分エンコードするステップを含む請求項１の方法。
【請求項１０】
前記オーディオ指紋に対するインデックス値を計算するステップを更に具備し、前記インデックス値は前記オーディオ指紋の一部を含む請求項１の方法。
【請求項１１】
前記インデックス値は、前記変換したオーディオ信号の低周波数成分のセットに対応する前記オーディオ指紋の一部を含む請求項１０の方法。
【請求項１２】
前記オーディオ指紋の一部を無関連化するステップを更に具備し、前記オーディオ指紋の前記無関連化した部分は、音響モデルにより重要でないと決定された周波数範囲に対応する請求項１の方法。
【請求項１３】
前記音響モデルは心理音響モデルである請求項１２の方法。
【請求項１４】
前記音響モデルはオーディオエンコード処理の特性を模す請求項１２の方法。
【請求項１５】
前記音響モデルは環境特性を模す請求項１２の方法。
【請求項１６】
前記音響モデルはオーディオ信号の特性を模す請求項１２の方法。
【請求項１７】
オーディオ指紋のデータベースを作成するための方法であって、
請求項１に記載された方法に従ってオーディオフレームからオーディオ指紋を抽出することに基づき、複数のオーディオ指紋を生成するステップと、
前記生成された各オーディオ指紋に対するインデックス値を計算するステップであって、前記インデックス値は、前記オーディオ指紋の一部を含むものと、
それらのインデックス値に従って前記データベース内の前記オーディオ指紋をグループ化するステップと
を具備する方法。
【請求項１８】
オーディオ断片のエッジを検出するための方法であって、
前記オーディオ断片を表す入力信号を受信するステップと、
前記受信入力信号の直近部分を有限長バッファに格納するステップと、
前記入力信号を受信している間の複数の時間毎に、請求項１の方法に従って、前記バッファに格納された前記受信した入力信号の一部についての指紋を計算するステップと、
前記計算された指紋それぞれと、前記オーディオ断片と関係付けられる参照指紋との間のビットエラー率（ＢＥＲ）を計算するステップと、
時間を通じて前記計算されたＢＥＲに基づいて前記オーディオ断片のエッジを識別するステップと、
を具備する方法。
【請求項１９】
前記オーディオ断片のエッジを識別するステップは、

時間を通じて低下していくＢＥＲの期間中に、前記計算されたＢＥＲの一部を一致させる最適ラインを見出すステップと、
前記最適ラインと、ＢＥＲが５０％に等しくなる水平ラインとの交点として前記オーディオ断片の開始エッジを識別するステップと、を含む請求項１８の方法。
【請求項２０】
前記オーディオ断片のエッジを識別するステップは、
時間を通じて増加していくＢＥＲの期間中に、前記計算されたＢＥＲの一部を一致させる最適ラインを見出すステップと、
前記最適ラインと、ＢＥＲが５０％に等しくなる水平ラインとの交点から、前記バッファ長を減算した点として、前記オーディオ断片の終了エッジを識別するステップと、を含む請求項１８の方法。
【請求項２１】
オーディオフレームからオーディオ指紋を抽出するための方法であって、
前記オーディオフレームに対するスペクトログラムを計算するステップと、
前記スペクトログラム内の複数の周波数バンドに対して非直線時間スケールで前記スペクトログラムをサンプリングするステップと、
前記サンプリングされた周波数バンドそれぞれからの前記サンプルを用いて、長時間特徴ベクトルを抽出するステップと、
前記特徴ベクトルに基づいて前記オーディオ指紋を生成するステップと、
を具備する方法。
【請求項２２】
前記スペクトログラムを対数時間スケールでサンプリングする請求項２１の方法。
【請求項２３】
前記周波数バンド毎に前記再サンプリングしたオーディオ信号を変換する前記ステップは、前記再サンプリングしたオーディオ信号に離散コサイン変換（ＤＣＴ）を施すことを含む請求項１の方法。
【請求項２４】
前記周波数バンドは、周波数軸に沿って対数的な中間周波数分布を持つことを特徴とする請求項１の方法。
【請求項２５】
前記フィルタ処理したオーディオ信号を非直線時間スケールで再サンプリングする前記ステップは、該フィルタ処理したオーディオ信号を指数的時間スケールで再サンプリングすることを特徴とする請求項１の方法。
【請求項２６】
前記オーディオ指紋を計算する前記ステップは、コードブック表に従ってビット値を割り当てることにより前記周波数バンドにつき前記特徴ベクトルをエンコードすることを含む請求項１の方法。
【請求項２７】
前記コードブック表にコードブック値が設定されるようにトレーニングすることと、
設定された前記コードブック値を基礎にした前記コードブック表をチューニングすることと
を更に具備する請求項２６の方法。
【請求項２８】
オーディオフレームからオーディオ指紋を抽出するために、コンピュータに、
前記オーディオフレームを複数の周波数バンドにフィルタ処理して、対応する複数のフィルタ処理したオーディオ信号を生成するステップと、
前記フィルタ処理されたオーディオ信号を非直線時間スケールで再サンプリングするステップと、
前記周波数バンド毎に前記再サンプリングされたオーディオ信号を変換して、前記周波数バンドに対する特徴ベクトルを生成するステップと、

前記特徴ベクトルのセットに基づいて前記オーディオ指紋を計算するステップと、
を実行させることからなるプログラム。
【請求項２９】
前記オーディオフレームを複数の周波数バンドにフィルタ処理するステップは、前記複数の周波数バンド毎の前記オーディオフレームをバンドパスフィルタ処理するステップを含む請求項２８のプログラム。
【請求項３０】
前記オーディオフレームを複数の周波数バンドにフィルタ処理するステップは、前記オーディオサンプルに高速フーリエ変換（ＦＦＴ）を実行するステップを含む請求項２８のプログラム。
【請求項３１】
前記オーディオフレームは、ＭＰ３フォーマットで格納されるオーディオファイルの一部であり、前記フィルタ処理したオーディオ信号は、前記オーディオファイルに関係付けられるＭＰ３ハイブリッドフィルタバンクから得られる請求項２８のプログラム。
【請求項３２】
前記フィルタ処理されたオーディオ信号は、対数時間スケールで再サンプリングされる請求項２８のプログラム。
【請求項３３】
前記周波数バンドは、周波数軸で直線的に離間している請求項２８のプログラム。
【請求項３４】
前記周波数バンドはオーバーラップしている請求項２８のプログラム。
【請求項３５】
特定周波数バンドの前記再サンプリングし、フィルタ処理したオーディオ信号を変換するステップは、前記再サンプリングしたオーディオ信号に高速フーリエ変換（ＦＦＴ）を実行するステップを含む請求項２８のプログラム。
【請求項３６】
前記オーディオ指紋を計算するステップは、前記周波数バンドに対する前記特徴ベクトルを差分エンコードするステップを含む請求項２８のプログラム。
【請求項３７】
前記オーディオ指紋に対するインデックス値を計算するステップを更に含み、前記インデックス値は前記オーディオ指紋の一部を含む請求項２８のプログラム。
【請求項３８】
前記インデックス値は、前記変換したオーディオ信号の低周波数成分のセットに対応する前記オーディオ指紋の一部を含む請求項３７のプログラム。
【請求項３９】
前記オーディオ指紋の一部を無関連化するステップを更に含み、前記オーディオ指紋の前記無関連化した部分は、音響モデルにより重要でないと決定した周波数範囲に対応する請求項２８のプログラム。
【請求項４０】
前記音響モデルは心理音響モデルである請求項３９のプログラム。
【請求項４１】
前記音響モデルはオーディオエンコード処理の特性を模す請求項３９のプログラム。
【請求項４２】
前記音響モデルは環境特性を模す請求項３９のプログラム。
【請求項４３】
前記音響モデルはオーディオ信号の特性を模す請求項３９のプログラム。
【請求項４４】
前記周波数バンド毎に前記再サンプリングしたオーディオ信号を変換する前記ステップは、前記再サンプリングしたオーディオ信号に離散コサイン変換（ＤＣＴ）を施すことを含む請求項２８のプログラム。

【請求項４５】
前記周波数バンドは、周波数軸に沿って対数的な中間周波数分布を持つことを特徴とする請求項２８のプログラム。
【請求項４６】
前記フィルタ処理したオーディオ信号を非直線時間スケールで再サンプリングする前記ステップは、該フィルタ処理したオーディオ信号を指数的時間スケールで再サンプリングすることを特徴とする請求項２８のプログラム。
【請求項４７】
前記オーディオ指紋を計算する前記ステップは、コードブック表に従ってビット値を割り当てることにより前記周波数バンドにつき前記特徴ベクトルをエンコードすることを含む請求項２８のプログラム。
【請求項４８】
前記コードブック表にコードブック値が設定されるようにトレーニングするステップと、
設定された前記コードブック値を基礎にした前記コードブック表をチューニングするステップと
を、前記コンピュータに更に実行させることを特徴する請求項４７のプログラム。