

【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11月26日 (2015.11.26)

【公表番号】特表2015-507222(P2015-507222A)
 【公表日】平成27年3月5日 (2015.3.5)
 【年通号数】公開・登録公報2015-015
 【出願番号】特願2014-552206(P2014-552206)
 【国際特許分類】

G 1 0 L 19/22 (2013.01)

G 1 0 L 25/81 (2013.01)

G 1 0 L 19/20 (2013.01)

【 F I 】

G 1 0 L 19/22

G 1 0 L 25/81

G 1 0 L 19/20

【手続補正書】

【提出日】平成27年10月6日 (2015.10.6)

【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】

デジタルオーディオデバイス内の第 1 の分類器においてオーディオ信号の一部分を受信することと、

前記デジタルオーディオデバイスが、前記第 1 の分類器において前記オーディオ信号の前記一部分をスピーチとしてまたは音楽として分類することと、

前記オーディオ信号の前記一部分を処理することと、

を備え、ここにおいて、前記オーディオ信号の前記一部分を処理することは、

前記一部分が前記第 1 の分類器によってスピーチとして分類された場合、前記デジタルオーディオデバイスが、第 1 のコーディングモードを使用して前記スピーチを符号化すること、または

前記一部分が前記第 1 の分類器によって音楽として分類された場合、

前記一部分を前記デジタルオーディオデバイス内の第 2 の分類器に与えることと、

前記デジタルオーディオデバイスが、前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することと、

前記オーディオ信号の前記一部分を符号化することと、

を備え、ここにおいて、前記オーディオ信号の前記一部分を符号化することは、

前記一部分が前記第 2 の分類器においてスピーチとして分類された場合、前記デジタルオーディオデバイスが、第 2 のコーディングモードを使用して前記一部分を符号化すること、または

前記一部分が前記第 2 の分類器において音楽として分類された場合、前記デジタルオーディオデバイスが、第 3 のコーディングモードを使用して前記一部分を符号化すること

を備える、

方法。

【請求項 2】

前記オーディオ信号の前記一部分がフレームである、請求項 1 に記載の方法。

【請求項 3】

前記第 1 のコーディングモードが第 1 のスピーチコードを備え、前記第 2 のコーディングモードが第 2 のスピーチコードを備え、前記第 3 のコーディングモードが音楽コードを備える、請求項 1 に記載の方法。

【請求項 4】

前記第 1 のスピーチコードが符号励振線形予測 (C E L P) 型コードであり、前記第 2 のスピーチコードが C E L P / 変換ハイブリッドコードであり、前記音楽コードが変換コードである、請求項 3 に記載の方法。

【請求項 5】

前記一部分を前記第 2 の分類器に与えるより前に、前記第 2 の分類器が有効にされるかどうかを判断することと、前記第 2 の分類器が有効にされない場合、前記第 3 のコーディングモードを用いて前記一部分を符号化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 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することは、前記一部分が音楽の特性を有するか有声スピーチの特性を有するかを分類するために、前記一部分の複数の特徴を 1 つまたは複数のしきい値と比較することを備える、請求項 1 に記載の方法。

【請求項 7】

音楽の前記特性が音楽のブロードバンド擬似雑音特性を備え、有声スピーチの前記特性が、有声スピーチの音特性または有声スピーチの準定常特性のうちの少なくとも 1 つを備える、請求項 6 に記載の方法。

【請求項 8】

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することが、前記一部分の発声を第 1 のしきい値と比較すること、修正相関を第 2 のしきい値と比較すること、または長期ピッチ利得を第 3 のしきい値と比較することのうちの少なくとも 1 つを備える、請求項 1 に記載の方法。

【請求項 9】

前記発声が、スピーチとの無相関に対応する 0 から、スピーチとの高相関に対応する 1 にわたり、前記修正相関が、ランダム雑音に対応する 0 から、高度に構造化されたサウンドに対応する 1 にわたり、前記長期ピッチ利得が、過去の励起と現在の予測残差との間の正規化相互相関であり、前記長期ピッチ利得は、過去の部分の誤差が現在の部分を表すのに十分でないことを示す 0 から、前記過去の部分の残余誤差を使用することが前記現在の部分を完全に表すことができることを示す 1 にわたる、請求項 8 に記載の方法。

【請求項 10】

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することは、前記一部分中に信号アクティビティがあるかどうかを判断することと、信号アクティビティがない場合、符号化すべき有用な信号がないと判断することと、前記一部分をスピーチとして符号化することとを備える、請求項 1 に記載の方法。

【請求項 11】

デジタルオーディオデバイス内の第 1 の分類器においてオーディオ信号の一部分を受信するための手段と、

前記デジタルオーディオデバイスが、前記第 1 の分類器において前記オーディオ信号の前記一部分をスピーチとしてまたは音楽として分類するための手段と、

前記一部分が前記第 1 の分類器によってスピーチとして分類された場合、前記デジタルオーディオデバイスが、第 1 のコーディングモードを使用して前記スピーチを符号化する、または、前記一部分が前記第 1 の分類器によって音楽として分類されたとき、前記デジタルオーディオデバイスが、第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類する、ための手段と、

前記一部分が前記第 2 の分類器においてスピーチとして分類されたとき、前記デジタルオーディオデバイスが、第 2 のコーディングモードを使用して前記一部分を符号化する、または、前記一部分が前記第 2 の分類器において音楽として分類されたとき、前記デジタルオーディオデバイスが、第 3 のコーディングモードを使用して前記一部分を符号化する
ための手段と
を備える装置。

【請求項 1 2】

前記オーディオ信号の前記一部分がフレームである、請求項 1 1 に記載の装置。

【請求項 1 3】

前記第 1 のコーディングモードが第 1 のスピーチコーダを備え、前記第 2 のコーディングモードが第 2 のスピーチコーダを備え、前記第 3 のコーディングモードが音楽コーダを備える、請求項 1 1 に記載の装置。

【請求項 1 4】

前記第 1 のスピーチコーダが符号励振線形予測 (C E L P) 型コーダであり、前記第 2 のスピーチコーダが C E L P / 変換ハイブリッドコーダであり、前記音楽コーダが変換コーダである、請求項 1 3 に記載の装置。

【請求項 1 5】

前記一部分を前記第 2 の分類器に与えるより前に、前記第 2 の分類器が有効にされるかどうかを判断することと、前記第 2 の分類器が有効にされない場合、前記第 3 のコーディングモードを用いて前記一部分を符号化することとを行うための手段をさらに備える、請求項 1 1 に記載の装置。

【請求項 1 6】

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類するための前記手段は、前記一部分が音楽の特性を有するか有声スピーチの特性を有するかを分類するために、前記一部分の複数の特徴を 1 つまたは複数のしきい値と比較するための手段を備える、請求項 1 1 に記載の装置。

【請求項 1 7】

音楽の前記特性が音楽のブロードバンド擬似雑音特性を備え、有声スピーチの前記特性が、有声スピーチの音特性または有声スピーチの準定常特性のうちの少なくとも 1 つを備える、請求項 1 6 に記載の装置。

【請求項 1 8】

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類するための前記手段が、前記一部分の発声を第 1 のしきい値と比較するための手段と、修正相関を第 2 のしきい値と比較するための手段と、長期ピッチ利得を第 3 のしきい値と比較するための手段とのうちの少なくとも 1 つを備える、請求項 1 1 に記載の装置。

【請求項 1 9】

前記発声が、スピーチとの無相関に対応する 0 から、スピーチとの高相関に対応する 1 にわたり、前記修正相関が、ランダム雑音に対応する 0 から、高度に構造化されたサウンドに対応する 1 にわたり、前記長期ピッチ利得が、過去の励起と現在の予測残差との間の正規化相互相関であり、前記長期ピッチ利得は、過去の部分の誤差が現在の部分を表すのに十分でないことを示す 0 から、前記過去の部分の残余誤差を使用することが前記現在の部分を完全に表すことができることを示す 1 にわたる、請求項 1 8 に記載の装置。

【請求項 2 0】

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類するための前記手段は、前記一部分中に信号アクティビティがあるかどうかを判断することと、信号アクティビティがない場合、符号化すべき有用な信号がないと判断することと、前記一部分をスピーチとして符号化することとを行うための手段を備える、請求項 1 1 に記載の装置。

【請求項 2 1】

デジタルオーディオデバイス内の第 1 の分類器においてオーディオ信号の一部分を受信

することと、

前記デジタルオーディオデバイスが、前記第 1 の分類器において前記オーディオ信号の前記一部分をスピーチとしてまたは音として分類することと、

前記オーディオ信号の前記一部分を処理することと、

をプロセッサに行わせるためのコンピュータ可読命令を備え、ここにおいて、前記オーディオ信号の前記一部分を処理することは、

前記一部分が前記第 1 の分類器によってスピーチとして分類された場合、前記デジタルオーディオデバイスが、第 1 のコーディングモードを使用して前記スピーチを符号化すること、または、

前記一部分が前記第 1 の分類器によって音楽として分類された場合、

前記一部分を前記デジタルオーディオデバイス内の第 2 の分類器に与えることと、

前記デジタルオーディオデバイスが、前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することと、

前記オーディオ信号の前記一部分を符号化することと、

を備え、ここにおいて、前記オーディオ信号の前記一部分を符号化することは、

前記一部分が前記第 2 の分類器においてスピーチとして分類された場合、前記デジタルオーディオデバイスが、第 2 のコーディングモードを使用して前記一部分を符号化すること、または、

前記一部分が前記第 2 の分類器において音楽として分類された場合、前記デジタルオーディオデバイスが、第 3 のコーディングモードを使用して前記一部分を符号化すること

を備える、

非一時的コンピュータ可読媒体。

【請求項 2 2】

前記オーディオ信号の前記一部分がフレームである、請求項 2 1 に記載のコンピュータ可読媒体。

【請求項 2 3】

前記第 1 のコーディングモードが第 1 のスピーチコードを備え、前記第 2 のコーディングモードが第 2 のスピーチコードを備え、前記第 3 のコーディングモードが音楽コードを備える、請求項 2 1 に記載のコンピュータ可読媒体。

【請求項 2 4】

前記第 1 のスピーチコードが符号励振線形予測 (C E L P) 型コードであり、前記第 2 のスピーチコードが C E L P / 変換ハイブリッドコードであり、前記音楽コードが変換コードである、請求項 2 3 に記載のコンピュータ可読媒体。

【請求項 2 5】

前記一部分を前記第 2 の分類器に与えるより前に、前記第 2 の分類器が有効にされるかどうかを判断することと、前記第 2 の分類器が有効にされない場合、前記第 3 のコーディングモードを用いて前記一部分を符号化することとを前記プロセッサに行わせる命令をさらに備える、請求項 2 1 に記載のコンピュータ可読媒体。

【請求項 2 6】

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することは、前記一部分が音楽の特性を有するか有声スピーチの特性を有するかを分類するために、前記一部分の複数の特徴を 1 つまたは複数のしきい値と比較することを備える、請求項 2 1 に記載のコンピュータ可読媒体。

【請求項 2 7】

音楽の前記特性が音楽のブロードバンド擬似雑音特性を備え、有声スピーチの前記特性が、有声スピーチの音特性または有声スピーチの準定常特性のうちの少なくとも 1 つを備える、請求項 2 6 に記載のコンピュータ可読媒体。

【請求項 2 8】

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類すること

が、前記一部分の発声を第 1 のしきい値と比較すること、修正相関を第 2 のしきい値と比較すること、または長期ピッチ利得を第 3 のしきい値と比較することのうちの少なくとも 1 つを備える、請求項 2 1 に記載のコンピュータ可読媒体。

【請求項 2 9】

前記発声が、スピーチとの無相関に対応する 0 から、スピーチとの高相関に対応する 1 にわたり、前記修正相関が、ランダム雑音に対応する 0 から、高度に構造化されたサウンドに対応する 1 にわたり、前記長期ピッチ利得が、過去の励起と現在の予測残差との間の正規化相互相関であり、前記長期ピッチ利得は、過去の部分の誤差が現在の部分を表すのに十分でないことを示す 0 から、前記過去の部分の残余誤差を使用することが前記現在の部分を完全に表すことができることを示す 1 にわたる、請求項 2 8 に記載のコンピュータ可読媒体。

【請求項 3 0】

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することは、前記一部分中に信号アクティビティがあるかどうかを判断することと、信号アクティビティがない場合、符号化すべき有用な信号がないと判断することと、前記一部分をスピーチとして符号化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 3 1】

装置であって、

プロセッサと、

前記プロセッサと電子通信するメモリと、

前記メモリ内に記憶された命令と

を備え、前記命令は、

デジタルオーディオデバイス内の第 1 の分類器においてオーディオ信号の一部分を受信することと、

前記デジタルオーディオデバイスが、前記第 1 の分類器において前記オーディオ信号の前記一部分をスピーチとしてまたは音楽として分類することと、

前記オーディオ信号の前記一部分を処理することと、

を前記プロセッサによって実行可能であり、ここにおいて、前記オーディオ信号の前記一部分を処理することは、

前記一部分が前記第 1 の分類器によってスピーチとして分類された場合、前記デジタルオーディオデバイスが、第 1 のコーディングモードを使用して前記スピーチを符号化すること、または、

前記一部分が前記第 1 の分類器によって音楽として分類された場合、

前記一部分を前記デジタルオーディオデバイス内の第 2 の分類器に与えることと

、

前記デジタルオーディオデバイスが、前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することと、

前記オーディオ信号の前記一部分を符号化することと、

を備え、ここにおいて、前記オーディオ信号の前記一部分を符号化することは、

前記一部分が前記第 2 の分類器においてスピーチとして分類された場合、前記デジタルオーディオデバイスが第 2 のコーディングモードを使用して前記一部分を符号化すること、または、

前記一部分が前記第 2 の分類器において音楽として分類された場合、前記デジタルオーディオデバイスが第 3 のコーディングモードを使用して前記一部分を符号化すること

を備える、

装置。

【請求項 3 2】

前記オーディオ信号の前記一部分がフレームである、請求項 3 1 に記載のプロセッサ。

【請求項 3 3】

前記第 1 のコーディングモードが第 1 のスピーチコーダを備え、前記第 2 のコーディングモードが第 2 のスピーチコーダを備え、前記第 3 のコーディングモードが音楽コーダを備える、請求項 3 1 に記載のプロセッサ。

【請求項 3 4】

前記第 1 のスピーチコーダが符号励振線形予測 (C E L P) 型コーダであり、前記第 2 のスピーチコーダが C E L P / 変換ハイブリッドコーダであり、前記音楽コーダが変換コーダである、請求項 3 3 に記載のプロセッサ。

【請求項 3 5】

前記命令はまた、前記一部分を前記第 2 の分類器に与えるより前に、前記第 2 の分類器が有効にされるかどうかを判断することと、前記第 2 の分類器が有効にされない場合、前記第 3 のコーディングモードを用いて前記一部分を符号化することとを実行可能である、請求項 3 1 に記載のプロセッサ。

【請求項 3 6】

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することを実行可能である前記命令は、前記一部分が音楽の特性を有するか有声スピーチの特性を有するかを分類するために、前記一部分の複数の特徴を 1 つまたは複数のしきい値と比較することを実行可能である命令を備える、請求項 3 1 に記載のプロセッサ。

【請求項 3 7】

音楽の前記特性が音楽のブロードバンド擬似雑音特性を備え、有声スピーチの前記特性が、有声スピーチの音特性または有声スピーチの準定常特性のうちの少なくとも 1 つを備える、請求項 3 6 に記載のプロセッサ。

【請求項 3 8】

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することを実行可能である前記命令は、前記一部分の発声を第 1 のしきい値と比較することを実行可能である命令と、修正相関を第 2 のしきい値と比較することを実行可能である命令と、または長期ピッチ利得を第 3 のしきい値と比較することを実行可能である命令と、のうちの少なくとも 1 つを備える、請求項 3 1 に記載のプロセッサ。

【請求項 3 9】

前記発声が、スピーチとの無相関に対応する 0 から、スピーチとの高相関に対応する 1 にわたり、前記修正相関が、ランダム雑音に対応する 0 から、高度に構造化されたサウンドに対応する 1 にわたり、前記長期ピッチ利得が、過去の励起と現在の予測残差との間の正規化相互相関であり、前記長期ピッチ利得は、過去の部分の誤差が現在の部分を表すのに十分でないことを示す 0 から、前記過去の部分の残余誤差を使用することが前記現在の部分を完全に表すことができることを示す 1 にわたる、請求項 3 8 に記載のプロセッサ。

【請求項 4 0】

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することを実行可能である前記命令は、前記一部分中に信号アクティビティがあるかどうかを判断することと、信号アクティビティがない場合、符号化すべき有用な信号がないと判断することと、前記一部分をスピーチとして符号化することとを実行可能である命令を備える、請求項 3 1 に記載のプロセッサ。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0 0 6 9

【補正方法】変更

【補正の内容】

【0 0 6 9】

[0074] 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、上記で説明した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。

[C 1]

第 1 の分類器においてオーディオ信号の一部を受信することと、

前記第 1 の分類器において前記オーディオ信号の前記一部分をスピーチとしてまたは音楽として分類することと、

前記一部分が前記第 1 の分類器によってスピーチとして分類された場合、第 1 のコーディングモードを使用して前記スピーチを符号化することと、

前記一部分が前記第 1 の分類器によって音楽として分類された場合、

前記一部分を第 2 の分類器に与えることと、

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することと、

前記一部分が前記第 2 の分類器においてスピーチとして分類された場合、第 2 のコーディングモードを使用して前記一部分を符号化することと、

前記一部分が前記第 2 の分類器において音楽として分類された場合、第 3 のコーディングモードを使用して前記一部分を符号化することとを備える方法。

[C 2]

前記オーディオ信号の前記一部分がフレームである、C 1 に記載の方法。

[C 3]

前記第 1 のコーディングモードが第 1 のスピーチコードを備え、前記第 2 のコーディングモードが第 2 のスピーチコードを備え、前記第 3 のコーディングモードが音楽コードを備える、C 1 に記載の方法。

[C 4]

前記第 1 のスピーチコードが符号励振線形予測 (C E L P) 型コードであり、前記第 2 のスピーチコードが C E L P / 変換ハイブリッドコードであり、前記音楽コードが変換コードである、C 3 に記載の方法。

[C 5]

前記一部分を第 2 の分類器に与えるより前に、前記第 2 の分類器が有効にされるかどうかを判断することと、前記第 2 の分類器が有効にされない場合、前記第 3 のコーディングモードを用いて前記一部分を符号化することとをさらに備える、C 1 に記載の方法。

[C 6]

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することは、前記一部分が音楽の特性を有するか有声スピーチの特性を有するかを分類するために、前記一部分の複数の特徴を 1 つまたは複数のしきい値と比較することを備える、C 1 に記載の方法。

[C 7]

音楽の前記特性が音楽のブロードバンド擬似雑音特性を備え、有声スピーチの前記特性が、有声スピーチの音特性または有声スピーチの準定常特性のうちの少なくとも 1 つを備える、C 6 に記載の方法。

[C 8]

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することが、前記一部分の発声を第 1 のしきい値と比較すること、修正相関を第 2 のしきい値と比較すること、または長期ピッチ利得を第 3 のしきい値と比較することのうちの少なくとも 1 つを備える、C 1 に記載の方法。

[C 9]

前記発声が、スピーチとの無相関に対応する 0 から、スピーチとの高相関に対応する 1 にわたり、前記修正相関が、ランダム雑音に対応する 0 から、高度に構造化されたサウンドに対応する 1 にわたり、前記長期ピッチ利得が、過去の励起と現在の予測残差との間の正規化相互相関であり、前記長期ピッチ利得は、過去の部分の誤差が現在の部分を表すのに十分でないことを示す 0 から、前記過去の部分の残余誤差を使用することが前記現在の

部分を完全に表すことができることを示す 1 にわたる、C 8 に記載の方法。

[C 1 0]

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することは、前記一部分中に信号アクティビティがあるかどうかを判断することと、信号アクティビティがない場合、符号化すべき有用な信号がないと判断することと、前記一部分をスピーチとして符号化することとを備える、C 1 に記載の方法。

[C 1 1]

第 1 の分類器においてオーディオ信号の一部分を受信するための手段と、
前記第 1 の分類器において前記オーディオ信号の前記一部分をスピーチとしてまたは音楽として分類するための手段と、
前記一部分が前記第 1 の分類器によってスピーチとして分類された場合、第 1 のコーディングモードを使用して前記スピーチを符号化するための手段と、
前記一部分が前記第 1 の分類器によって音楽として分類されたとき、前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類するための手段と、
前記一部分が前記第 2 の分類器においてスピーチとして分類されたとき、第 2 のコーディングモードを使用して前記一部分を符号化するための手段と、
前記一部分が前記第 2 の分類器において音楽として分類されたとき、第 3 のコーディングモードを使用して前記一部分を符号化するための手段とを備える装置。

[C 1 2]

前記オーディオ信号の前記一部分がフレームである、C 1 1 に記載の装置。

[C 1 3]

前記第 1 のコーディングモードが第 1 のスピーチコードを備え、前記第 2 のコーディングモードが第 2 のスピーチコードを備え、前記第 3 のコーディングモードが音楽コードを備える、C 1 1 に記載の装置。

[C 1 4]

前記第 1 のスピーチコードが符号励振線形予測 (C E L P) 型コードであり、前記第 2 のスピーチコードが C E L P / 変換ハイブリッドコードであり、前記音楽コードが変換コードである、C 1 3 に記載の装置。

[C 1 5]

前記一部分を第 2 の分類器に与えるより前に、前記第 2 の分類器が有効にされるかどうかを判断することと、前記第 2 の分類器が有効にされない場合、前記第 3 のコーディングモードを用いて前記一部分を符号化することとを行うための手段をさらに備える、C 1 1 に記載の装置。

[C 1 6]

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類するための前記手段は、前記一部分が音楽の特性を有するか有声スピーチの特性を有するかを分類するために、前記一部分の複数の特徴を 1 つまたは複数のしきい値と比較するための手段を備える、C 1 1 に記載の装置。

[C 1 7]

音楽の前記特性が音楽のブロードバンド擬似雑音特性を備え、有声スピーチの前記特性が、有声スピーチの音特性または有声スピーチの準定常特性のうちの少なくとも 1 つを備える、C 1 6 に記載の装置。

[C 1 8]

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類するための前記手段が、前記一部分の発声を第 1 のしきい値と比較するための手段と、修正相関を第 2 のしきい値と比較するための手段と、長期ピッチ利得を第 3 のしきい値と比較するための手段とのうちの少なくとも 1 つを備える、C 1 1 に記載の装置。

[C 1 9]

前記発声が、スピーチとの無相関に対応する 0 から、スピーチとの高相関に対応する 1 にわたり、前記修正相関が、ランダム雑音に対応する 0 から、高度に構造化されたサウン

ドに対応する 1 にわたり、前記長期ピッチ利得が、過去の励起と現在の予測残差との間の正規化相互相関であり、前記長期ピッチ利得は、過去の部分の誤差が現在の部分を表すのに十分でないことを示す 0 から、前記過去の部分の残余誤差を使用することが前記現在の部分を完全に表すことができることを示す 1 にわたる、C 1 8 に記載の装置。

[C 2 0]

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類するための前記手段は、前記一部分中に信号アクティビティがあるかどうかを判断することと、信号アクティビティがない場合、符号化すべき有用な信号がないと判断することと、前記一部分をスピーチとして符号化することとを行うための手段を備える、C 1 1 に記載の装置。

[C 2 1]

第 1 の分類器においてオーディオ信号の一部分を受信することと、

前記第 1 の分類器において前記オーディオ信号の前記一部分をスピーチとしてまたは音楽として分類することと、

前記一部分が前記第 1 の分類器によってスピーチとして分類された場合、第 1 のコーディングモードを使用して前記スピーチを符号化することと、

前記一部分が前記第 1 の分類器によって音楽として分類された場合、

前記一部分を第 2 の分類器に与えることと、

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することと、

前記一部分が前記第 2 の分類器においてスピーチとして分類された場合、第 2 のコーディングモードを使用して前記一部分を符号化することと、

前記一部分が前記第 2 の分類器において音楽として分類された場合、第 3 のコーディングモードを使用して前記一部分を符号化することとをプロセッサに行わせるためのコンピュータ可読命令を備える非一時的コンピュータ可読媒体。

[C 2 2]

前記オーディオ信号の前記一部分がフレームである、C 2 1 に記載のコンピュータ可読媒体。

[C 2 3]

前記第 1 のコーディングモードが第 1 のスピーチコードを備え、前記第 2 のコーディングモードが第 2 のスピーチコードを備え、前記第 3 のコーディングモードが音楽コードを備える、C 2 1 に記載のコンピュータ可読媒体。

[C 2 4]

前記第 1 のスピーチコードが符号励振線形予測 (C E L P) 型コードであり、前記第 2 のスピーチコードが C E L P / 変換ハイブリッドコードであり、前記音楽コードが変換コードである、C 2 3 に記載のコンピュータ可読媒体。

[C 2 5]

前記一部分を第 2 の分類器に与えるより前に、前記第 2 の分類器が有効にされるかどうかを判断することと、前記第 2 の分類器が有効にされない場合、前記第 3 のコーディングモードを用いて前記一部分を符号化することとを前記プロセッサに行わせる命令をさらに備える、C 2 1 に記載のコンピュータ可読媒体。

[C 2 6]

前記第 2 の分類器において前記一部分をスピーチとしてまたは音楽として分類することは、前記一部分が音楽の特性を有するか有声スピーチの特性を有するかを分類するために、前記一部分の複数の特徴を 1 つまたは複数のしきい値と比較することを備える、C 2 1 に記載のコンピュータ可読媒体。

[C 2 7]

音楽の前記特性が音楽のブロードバンド擬似雑音特性を備え、有声スピーチの前記特性が、有声スピーチの音特性または有声スピーチの準定常特性のうちの少なくとも 1 つを備える、C 2 6 に記載のコンピュータ可読媒体。

[C 2 8]

前記第2の分類器において前記一部分をスピーチとしてまたは音楽として分類することが、前記一部分の発声を第1のしきい値と比較すること、修正相関を第2のしきい値と比較すること、または長期ピッチ利得を第3のしきい値と比較することのうちの少なくとも1つを備える、C 2 1に記載のコンピュータ可読媒体。

[C 2 9]

前記発声が、スピーチとの無相関に対応する0から、スピーチとの高相関に対応する1にわたり、前記修正相関が、ランダム雑音に対応する0から、高度に構造化されたサウンドに対応する1にわたり、前記長期ピッチ利得が、過去の励起と現在の予測残差との間の正規化相互相関であり、前記長期ピッチ利得は、過去の部分の誤差が現在の部分を表すのに十分でないことを示す0から、前記過去の部分の残余誤差を使用することが前記現在の部分を完全に表すことができることを示す1にわたる、C 2 8に記載のコンピュータ可読媒体。

[C 3 0]

前記第2の分類器において前記一部分をスピーチとしてまたは音楽として分類することは、前記一部分中に信号アクティビティがあるかどうかを判断することと、信号アクティビティがない場合、符号化すべき有用な信号がないと判断することと、前記一部分をスピーチとして符号化することとを備える、C 2 1に記載のコンピュータ可読媒体。

[C 3 1]

第1の分類器においてオーディオ信号の一部分を受信することと、

前記第1の分類器において前記オーディオ信号の前記一部分をスピーチとしてまたは音楽として分類することと、

前記一部分が前記第1の分類器によってスピーチとして分類された場合、第1のコーディングモードを使用して前記スピーチを符号化することと、

前記一部分が前記第1の分類器によって音楽として分類された場合、

前記一部分を第2の分類器に与えることと、

前記第2の分類器において前記一部分をスピーチとしてまたは音楽として分類することと、

前記一部分が前記第2の分類器においてスピーチとして分類された場合、第2のコーディングモードを使用して前記一部分を符号化することと、

前記一部分が前記第2の分類器において音楽として分類された場合、第3のコーディングモードを使用して前記一部分を符号化することとを行うように構成されたプロセッサ。

[C 3 2]

前記オーディオ信号の前記一部分がフレームである、C 3 1に記載のプロセッサ。

[C 3 3]

前記第1のコーディングモードが第1のスピーチコードを備え、前記第2のコーディングモードが第2のスピーチコードを備え、前記第3のコーディングモードが音楽コードを備える、C 3 1に記載のプロセッサ。

[C 3 4]

前記第1のスピーチコードが符号励振線形予測(CELP)型コードであり、前記第2のスピーチコードがCELP/変換ハイブリッドコードであり、前記音楽コードが変換コードである、C 3 3に記載のプロセッサ。

[C 3 5]

前記プロセッサは、前記一部分を第2の分類器に与えるより前に、前記第2の分類器が有効にされるかどうかを判断することと、前記第2の分類器が有効にされない場合、前記第3のコーディングモードを用いて前記一部分を符号化することとを行うようにさらに構成された、C 3 1に記載のプロセッサ。

[C 3 6]

前記第2の分類器において前記一部分をスピーチとしてまたは音楽として分類すること

は、前記一部分が音楽の特性を有するか有聲スピーチの特性を有するかを分類するために、前記一部分の複数の特徴を1つまたは複数のしきい値と比較することを備える、C 3 1に記載のプロセッサ。

[C 3 7]

音楽の前記特性が音楽のブロードバンド擬似雑音特性を備え、有聲スピーチの前記特性が、有聲スピーチの音特性または有聲スピーチの準定常特性のうちの少なくとも1つを備える、C 3 6に記載のプロセッサ。

[C 3 8]

前記第2の分類器において前記一部分をスピーチとしてまたは音楽として分類することが、前記一部分の発声を第1のしきい値と比較すること、修正相関を第2のしきい値と比較すること、または長期ピッチ利得を第3のしきい値と比較することのうちの少なくとも1つを備える、C 3 1に記載のプロセッサ。

[C 3 9]

前記発声が、スピーチとの無相関に対応する0から、スピーチとの高相関に対応する1にわたり、前記修正相関が、ランダム雑音に対応する0から、高度に構造化されたサウンドに対応する1にわたり、前記長期ピッチ利得が、過去の励起と現在の予測残差との間の正規化相互相関であり、前記長期ピッチ利得は、過去の部分の誤差が現在の部分を表すのに十分でないことを示す0から、前記過去の部分の残余誤差を使用することが前記現在の部分を完全に表すことができることを示す1にわたる、C 3 8に記載のプロセッサ。

[C 4 0]

前記第2の分類器において前記一部分をスピーチとしてまたは音楽として分類することは、前記一部分中に信号アクティビティがあるかどうかを判断することと、信号アクティビティがない場合、符号化すべき有用な信号がないと判断することと、前記一部分をスピーチとして符号化することとを備える、C 3 1に記載のプロセッサ。

[C 4 1]

オーディオ信号の一部分を受信し、前記オーディオ信号の前記一部分をスピーチとしてまたは音楽として分類し、前記一部分がスピーチとして分類された場合、第1のコーディングモードを使用して前記スピーチを符号化し、前記一部分が音楽として分類された場合、前記一部分を第2の分類器に与える、第1の分類器と、

前記一部分が前記第1の分類器によって音楽として分類された場合、前記一部分をスピーチとしてまたは音楽として分類し、前記一部分が前記第2の分類器においてスピーチとして分類された場合、第2のコーディングモードを使用して前記一部分を符号化し、前記一部分が前記第2の分類器において音楽として分類された場合、第3のコーディングモードを使用して前記一部分を符号化する、前記第2の分類器とを備えるシステム。

[C 4 2]

前記オーディオ信号の前記一部分がフレームである、C 4 1に記載のシステム。

[C 4 3]

前記第1のコーディングモードが第1のスピーチコードを備え、前記第2のコーディングモードが第2のスピーチコードを備え、前記第3のコーディングモードが音楽コードを備える、C 4 1に記載のシステム。

[C 4 4]

前記第1のスピーチコードが符号励振線形予測(CELP)型コードであり、前記第2のスピーチコードがCELP/変換ハイブリッドコードであり、前記音楽コードが変換コードである、C 4 3に記載のシステム。

[C 4 5]

前記一部分を第2の分類器に与えるより前に、前記第2の分類器が有効にされるかどうかを判断することと、前記第2の分類器が有効にされない場合、前記第3のコーディングモードを用いて前記一部分を符号化することとをさらに備える、C 4 1に記載のシステム。

[C 4 6]

前記第2の分類器において前記一部分をスピーチとしてまたは音楽として分類することは、前記一部分が音楽の特性を有するか有聲スピーチの特性を有するかを分類するために、前記一部分の複数の特徴を1つまたは複数のしきい値と比較することを備える、C41に記載のシステム。

[C47]

音楽の前記特性が音楽のブロードバンド擬似雑音特性を備え、有聲スピーチの前記特性が、有聲スピーチの音特性または有聲スピーチの準定常特性のうちの少なくとも1つを備える、C46に記載のシステム。

[C48]

前記第2の分類器において前記一部分をスピーチとしてまたは音楽として分類することが、前記一部分の発声を第1のしきい値と比較すること、修正相関を第2のしきい値と比較すること、または長期ピッチ利得を第3のしきい値と比較することのうちの少なくとも1つを備える、C41に記載のシステム。

[C49]

前記発声が、スピーチとの無相関に対応する0から、スピーチとの高相関に対応する1にわたり、前記修正相関が、ランダム雑音に対応する0から、高度に構造化されたサウンドに対応する1にわたり、前記長期ピッチ利得が、過去の励起と現在の予測残差との間の正規化相互相関であり、前記長期ピッチ利得は、過去の部分の誤差が現在の部分を表すのに十分でないことを示す0から、前記過去の部分の残余誤差を使用することが前記現在の部分を完全に表すことができることを示す1にわたる、C48に記載のシステム。

[C50]

前記第2の分類器において前記一部分をスピーチとしてまたは音楽として分類することは、前記一部分中に信号アクティビティがあるかどうかを判断することと、信号アクティビティがない場合、符号化すべき有用な信号がないと判断することと、前記一部分をスピーチとして符号化することとを備える、C41に記載のシステム。