

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7012917号
(P7012917)

(45)発行日 令和4年1月28日(2022.1.28)

(24)登録日 令和4年1月20日(2022.1.20)

(51)国際特許分類	F I			
G 1 0 L 15/04 (2013.01)	G 1 0 L	15/04	3 0 0 A	
G 1 0 L 25/84 (2013.01)	G 1 0 L	15/04	3 0 0 B	
	G 1 0 L	25/84		

請求項の数 17 (全23頁)

(21)出願番号	特願2021-559189(P2021-559189)	(73)特許権者	000006013 三菱電機株式会社 東京都千代田区丸の内二丁目7番3号
(86)(22)出願日	令和1年12月13日(2019.12.13)	(74)代理人	100116964 弁理士 山形 洋一
(86)国際出願番号	PCT/JP2019/048921	(74)代理人	100120477 弁理士 佐藤 賢改
(87)国際公開番号	WO2021/117219	(74)代理人	100135921 弁理士 篠原 昌彦
(87)国際公開日	令和3年6月17日(2021.6.17)	(74)代理人	100083840 弁理士 前田 実
審査請求日	令和3年10月4日(2021.10.4)	(72)発明者	花澤 利行 東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内
早期審査対象出願		審査官	渡部 幸和

最終頁に続く

(54)【発明の名称】 情報処理装置、検出方法、及び検出プログラム

(57)【特許請求の範囲】

【請求項1】

音信号を取得する取得部と、

前記音信号を複数の区間に分割し、前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、前記複数の区間のそれぞれの、音声らしさの度合いである音声度を前記音信号に基づいて算出し、前記複数の区間の中で、前記変動値が予め設定された閾値以下であり、かつ前記音声度が予め設定された閾値以下である区間を特定し、前記音信号に基づいて、特定された区間における前記音信号のパワーを算出し、特定された区間における前記音信号のパワーの中から最大値を特定し、前記最大値に基づく値を検出閾値に設定し、時間経過に伴う前記音信号のパワーの中で前記検出閾値以上の区間を検出対象区間として検出する制御部と、

を有する情報処理装置。

【請求項2】

音信号を取得する取得部と、

前記音信号を複数の区間に分割し、前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、前記複数の区間のそれぞれの、音声らしさの度合いである音声度を前記音信号に基づいて算出し、前記複数の区間の中で、前記変動値が予め設定された閾値以下の区間を特定し、特定された区間の前記音声度を小さい順に並べ、小さい順に予め設定された個数の区間を選択し、前記音信号に基づいて、選択された区間における前記音信号のパワーを算出し、選択された区間における前記音信号

のパワーの中から最大値を特定し、前記最大値に基づく値を検出閾値に設定し、時間経過に伴う前記音信号のパワーの中で前記検出閾値以上の区間を検出対象区間として検出する制御部と、

を有する情報処理装置。

【請求項 3】

音信号を取得する取得部と、

前記音信号を複数の区間に分割し、前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、前記複数の区間の中で、前記変動値が予め設定された閾値以下の区間を特定し、特定された区間毎に、区間における前記音信号のパワーに基づく値を仮検出閾値に設定し、時間経過に伴う前記音信号のパワーの中で、設定された仮検出閾値以上の区間の数を検出し、特定された区間毎に設定された仮検出閾値の中から、区間の数が最大であるときの仮検出閾値を検出閾値として検出し、前記検出閾値を用いて検出されたときの区間を、検出対象区間として検出する制御部と、

10

を有する情報処理装置。

【請求項 4】

音信号を取得する取得部と、

前記音信号を複数の区間に分割し、前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、前記複数の区間のそれぞれの、音声らしさの度合いである音声度を前記音信号に基づいて算出し、前記複数の区間の中で、前記変動値が予め設定された閾値以下の区間を特定し、特定された区間の前記音声度を小さい順に並べ、小さい順に予め設定された個数の区間を選択し、前記音信号に基づいて、選択された区間における前記音信号のパワーを算出し、選択された区間毎に、区間における前記音信号のパワーに基づく値を仮検出閾値に設定し、時間経過に伴う前記音信号のパワーの中で、設定された仮検出閾値以上の区間の数を検出し、選択された区間毎に設定された仮検出閾値の中から、区間の数が最大であるときの仮検出閾値を検出閾値として検出し、前記検出閾値を用いて検出されたときの区間を、検出対象区間として検出する制御部と、

20

を有する情報処理装置。

【請求項 5】

前記制御部は、前記検出対象区間を音声区間として検出する、
請求項 1 から 4 のいずれか 1 項に記載の情報処理装置。

30

【請求項 6】

音信号を取得する取得部と、

時間経過に伴う、かつ音声らしさの度合いである音声度を前記音信号に基づいて算出し、時間経過に伴う前記音声度の中で、前記音声度が予め設定された閾値未満の区間を、非定常雑音区間候補として、特定し、前記非定常雑音区間候補を複数の区間に分割し、前記音信号に基づいて、分割された複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、分割された複数の区間の中で、前記変動値が予め設定された閾値以下の区間を特定し、前記音信号に基づいて、特定された区間における前記音信号のパワーを算出し、特定された区間における前記音信号のパワーの中から最大値を特定し、特定された最大値に基づく値を検出閾値に設定し、前記非定常雑音区間候補内であり、かつ時間経過に伴う前記音信号のパワーの中で前記検出閾値以上の区間を非定常雑音区間として検出する制御部と、

40

を有する情報処理装置。

【請求項 7】

検出された区間を示す情報を出力する出力部をさらに有する、
請求項 1 から 6 のいずれか 1 項に記載の情報処理装置。

【請求項 8】

情報処理装置が、
音信号を取得し、
前記音信号を複数の区間に分割し、

50

前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、前記複数の区間のそれぞれの、音声らしさの度合いである音声度を前記音信号に基づいて算出し、

前記複数の区間の中で、前記変動値が予め設定された閾値以下であり、かつ前記音声度が予め設定された閾値以下である区間を特定し、

前記音信号に基づいて、特定された区間における前記音信号のパワーを算出し、

特定された区間における前記音信号のパワーの中から最大値を特定し、

前記最大値に基づく値を検出閾値に設定し、

時間経過に伴う前記音信号のパワーの中で前記検出閾値以上の区間を検出対象区間として検出する、

検出方法。

【請求項 9】

情報処理装置が、

音信号を取得し、

前記音信号を複数の区間に分割し、

前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、前記複数の区間のそれぞれの、音声らしさの度合いである音声度を前記音信号に基づいて算出し、

前記複数の区間の中で、前記変動値が予め設定された閾値以下の区間を特定し、

特定された区間の前記音声度を小さい順に並べ、

小さい順に予め設定された個数の区間を選択し、

前記音信号に基づいて、選択された区間における前記音信号のパワーを算出し、

選択された区間における前記音信号のパワーの中から最大値を特定し、

前記最大値に基づく値を検出閾値に設定し、

時間経過に伴う前記音信号のパワーの中で前記検出閾値以上の区間を検出対象区間として検出する、

検出方法。

【請求項 10】

情報処理装置が、

音信号を取得し、

前記音信号を複数の区間に分割し、

前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、

前記複数の区間の中で、前記変動値が予め設定された閾値以下の区間を特定し、

特定された区間毎に、区間における前記音信号のパワーに基づく値を仮検出閾値に設定し、時間経過に伴う前記音信号のパワーの中で、設定された仮検出閾値以上の区間の数を検出し、

特定された区間毎に設定された仮検出閾値の中から、区間の数が最大であるときの仮検出閾値を検出閾値として検出し、

前記検出閾値を用いて検出されたときの区間を、検出対象区間として検出する、

検出方法。

【請求項 11】

情報処理装置が、

音信号を取得し、

前記音信号を複数の区間に分割し、

前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、

前記複数の区間のそれぞれの、音声らしさの度合いである音声度を前記音信号に基づいて算出し、

前記複数の区間の中で、前記変動値が予め設定された閾値以下の区間を特定し、

10

20

30

40

50

特定された区間の前記音声を小さい順に並べ、小さい順に予め設定された個数の区間を選択し、

前記音信号に基づいて、選択された区間における前記音信号のパワーを算出し、
 選択された区間毎に、区間における前記音信号のパワーに基づく値を仮検出閾値に設定し、時間経過に伴う前記音信号のパワーの中で、設定された仮検出閾値以上の区間の数を検出し、

選択された区間毎に設定された仮検出閾値の中から、区間の数が最大であるときの仮検出閾値を検出閾値として検出し、

前記検出閾値を用いて検出されたときの区間を、検出対象区間として検出する、

検出方法。

10

【請求項 1 2】

情報処理装置が、

音信号を取得し、

時間経過に伴う、かつ音声らしさの度合いである音声を前記音信号に基づいて算出し、

時間経過に伴う前記音声度の中で、前記音声度が予め設定された閾値未満の区間を、非定常雑音区間候補として、特定し、

前記非定常雑音区間候補を複数の区間に分割し、

前記音信号に基づいて、分割された複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、

分割された複数の区間の中で、前記変動値が予め設定された閾値以下の区間を特定し、

20

前記音信号に基づいて、特定された区間における前記音信号のパワーを算出し、

特定された区間における前記音信号のパワーの中から最大値を特定し、

特定された最大値に基づく値を検出閾値に設定し、

前記非定常雑音区間候補内であり、かつ時間経過に伴う前記音信号のパワーの中で前記検出閾値以上の区間を非定常雑音区間として検出する、

検出方法。

【請求項 1 3】

情報処理装置に、

音信号を取得し、

前記音信号を複数の区間に分割し、

30

前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、前記複数の区間のそれぞれの、音声らしさの度合いである音声を前記音信号に基づいて算出し、

前記複数の区間の中で、前記変動値が予め設定された閾値以下であり、かつ前記音声度が予め設定された閾値以下である区間を特定し、

前記音信号に基づいて、特定された区間における前記音信号のパワーを算出し、

特定された区間における前記音信号のパワーの中から最大値を特定し、

前記最大値に基づく値を検出閾値に設定し、

時間経過に伴う前記音信号のパワーの中で前記検出閾値以上の区間を検出対象区間として検出する、

40

処理を実行させる検出プログラム。

【請求項 1 4】

情報処理装置に、

音信号を取得し、

前記音信号を複数の区間に分割し、

前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、前記複数の区間のそれぞれの、音声らしさの度合いである音声を前記音信号に基づいて算出し、

前記複数の区間の中で、前記変動値が予め設定された閾値以下の区間を特定し、

特定された区間の前記音声を小さい順に並べ、

50

小さい順に予め設定された個数の区間を選択し、
 前記音信号に基づいて、選択された区間における前記音信号のパワーを算出し、
 選択された区間における前記音信号のパワーの中から最大値を特定し、
 前記最大値に基づく値を検出閾値に設定し、
 時間経過に伴う前記音信号のパワーの中で前記検出閾値以上の区間を検出対象区間として検出する、
 処理を実行させる検出プログラム。

【請求項 15】

情報処理装置に、
 音信号を取得し、
 前記音信号を複数の区間に分割し、
 前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、
 前記複数の区間の中で、前記変動値が予め設定された閾値以下の区間を特定し、
 特定された区間毎に、区間における前記音信号のパワーに基づく値を仮検出閾値に設定し、時間経過に伴う前記音信号のパワーの中で、設定された仮検出閾値以上の区間の数を検出し、
 特定された区間毎に設定された仮検出閾値の中から、区間の数が最大であるときの仮検出閾値を検出閾値として検出し、
 前記検出閾値を用いて検出されたときの区間を、検出対象区間として検出する、
 処理を実行させる検出プログラム。

10

20

【請求項 16】

情報処理装置に、
 音信号を取得し、
 前記音信号を複数の区間に分割し、
 前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、
 前記複数の区間のそれぞれの、音声らしさの度合いである音声度を前記音信号に基づいて算出し、
 前記複数の区間の中で、前記変動値が予め設定された閾値以下の区間を特定し、
 特定された区間の前記音声度を小さい順に並べ、小さい順に予め設定された個数の区間を選択し、
 前記音信号に基づいて、選択された区間における前記音信号のパワーを算出し、
 選択された区間毎に、区間における前記音信号のパワーに基づく値を仮検出閾値に設定し、時間経過に伴う前記音信号のパワーの中で、設定された仮検出閾値以上の区間の数を検出し、
 選択された区間毎に設定された仮検出閾値の中から、区間の数が最大であるときの仮検出閾値を検出閾値として検出し、
 前記検出閾値を用いて検出されたときの区間を、検出対象区間として検出する、
 処理を実行させる検出プログラム。

30

40

【請求項 17】

情報処理装置に、
 音信号を取得し、
 時間経過に伴う、かつ音声らしさの度合いである音声度を前記音信号に基づいて算出し、
 時間経過に伴う前記音声度の中で、前記音声度が予め設定された閾値未満の区間を、非定常雑音区間候補として、特定し、
 前記非定常雑音区間候補を複数の区間に分割し、
 前記音信号に基づいて、分割された複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、
 分割された複数の区間の中で、前記変動値が予め設定された閾値以下の区間を特定し、

50

前記音信号に基づいて、特定された区間における前記音信号のパワーを算出し、
特定された区間における前記音信号のパワーの中から最大値を特定し、
特定された最大値に基づく値を検出閾値に設定し、
前記非定常雑音区間候補内であり、かつ時間経過に伴う前記音信号のパワーの中で前記
検出閾値以上の区間を非定常雑音区間として検出する、
処理を実行させる検出プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、情報処理装置、検出方法、及び検出プログラムに関する。

10

【背景技術】

【0002】

音声認識の技術が知られている。例えば、音声信号の中の音声区間に対して音声認識を行う技術が提案されている（特許文献1を参照）。

【先行技術文献】

【特許文献】

【0003】

【文献】特開平10-288994号公報

【発明の概要】

【発明が解決しようとする課題】

20

【0004】

ところで、音信号の中から検出対象を検出したい場合がある。例えば、雑音のパワーに基づく閾値を用いて、検出対象を検出する方法が考えられる。ここで、雑音のパワーが急激に上がる場合がある。雑音のパワーが閾値を超えた場合、当該方法では、検出対象が精度良く検出できない。

【0005】

本開示の目的は、検出対象を精度良く検出することである。

【課題を解決するための手段】

【0006】

本開示の一態様に係る情報処理装置が提供される。情報処理装置は、音信号を取得する取得部と、前記音信号を複数の区間に分割し、前記音信号に基づいて、前記複数の区間のそれぞれの、区間時間当たりの変動量である変動値を算出し、前記複数の区間のそれぞれの、音声らしさの度合いである音声度を前記音信号に基づいて算出し、前記複数の区間の中で、前記変動値が予め設定された閾値以下であり、かつ前記音声度が予め設定された閾値以下である区間を特定し、前記音信号に基づいて、特定された区間における前記音信号のパワーを算出し、特定された区間における前記音信号のパワーの中から最大値を特定し、前記最大値に基づく値を検出閾値に設定し、時間経過に伴う前記音信号のパワーの中で前記検出閾値以上の区間を検出対象区間として検出する制御部と、を有する。

30

【発明の効果】

【0007】

本開示によれば、検出対象を精度良く検出できる。

40

【図面の簡単な説明】

【0008】

【図1】実施の形態1の情報処理装置が有するハードウェアの構成を示す図である。

【図2】比較例を示す図である。

【図3】実施の形態1の情報処理装置が有する機能ブロック図である。

【図4】実施の形態1の情報処理装置が実行する処理の例を示すフローチャートである。

【図5】実施の形態1の情報処理装置が実行する処理の具体例を示す。

【図6】実施の形態2の情報処理装置が有する機能ブロック図である。

【図7】実施の形態2の情報処理装置が実行する処理の例を示すフローチャートである。

50

- 【図 8】実施の形態 2 の情報処理装置が実行する処理の具体例を示す。
- 【図 9】実施の形態 3 の情報処理装置が有する機能ブロック図である。
- 【図 10】実施の形態 3 の情報処理装置が実行する処理の例を示すフローチャートである。
- 【図 11】実施の形態 4 の情報処理装置が有する機能ブロック図である。
- 【図 12】実施の形態 4 の情報処理装置が実行する処理の例を示すフローチャート（その 1）である。
- 【図 13】実施の形態 4 の情報処理装置が実行する処理の例を示すフローチャート（その 2）である。
- 【図 14】実施の形態 4 の情報処理装置が実行する処理の具体例（その 1）を示す。
- 【図 15】実施の形態 4 の情報処理装置が実行する処理の具体例（その 2）を示す。 10
- 【図 16】実施の形態 4 の変形例を示すフローチャート（その 1）である。
- 【図 17】実施の形態 4 の変形例を示すフローチャート（その 2）である。
- 【図 18】実施の形態 5 の情報処理装置が有する機能ブロック図である。
- 【図 19】実施の形態 5 の情報処理装置が実行する処理の例を示すフローチャート（その 1）である。
- 【図 20】実施の形態 5 の情報処理装置が実行する処理の例を示すフローチャート（その 2）である。
- 【図 21】実施の形態 5 の情報処理装置が実行する処理の具体例（その 1）を示す。
- 【図 22】実施の形態 5 の情報処理装置が実行する処理の具体例（その 2）を示す。
- 【発明を実施するための形態】 20

【0009】

以下、図面を参照しながら実施の形態を説明する。以下の実施の形態は、例にすぎず、本開示の範囲内で種々の変更が可能である。

【0010】

実施の形態 1 .

図 1 は、実施の形態 1 の情報処理装置が有するハードウェアの構成を示す図である。情報処理装置 100 は、検出方法を実行する装置である。情報処理装置 100 は、プロセッサ 101、揮発性記憶装置 102、及び不揮発性記憶装置 103 を有する。

【0011】

プロセッサ 101 は、情報処理装置 100 全体を制御する。例えば、プロセッサ 101 は、CPU (Central Processing Unit)、FPGA (Field Programmable Gate Array) などである。プロセッサ 101 は、マルチプロセッサでもよい。情報処理装置 100 は、処理回路によって実現されてもよく、又は、ソフトウェア、ファームウェア若しくはそれらの組み合わせによって実現されてもよい。なお、処理回路は、単回路又は複回路でもよい。 30

【0012】

揮発性記憶装置 102 は、情報処理装置 100 の主記憶装置である。例えば、揮発性記憶装置 102 は、RAM (Random Access Memory) である。不揮発性記憶装置 103 は、情報処理装置 100 の補助記憶装置である。例えば、不揮発性記憶装置 103 は、HDD (Hard Disk Drive) 又は SSD (Solid State Drive) である。 40

【0013】

図 2 は、比較例を示す図である。図 2 の上段は、音の波形のグラフを示している。図 2 の上段の音の音信号をパワーで示したグラフが、図 2 の下段である。図 2 の範囲 900 は、雑音を示している。

【0014】

音信号の中から検出対象を検出したい場合がある。図 2 では、検出対象を音声とする。ここで、雑音のパワーは、音声のパワーより低い場合が多い。そこで、閾値を用いて、音声を検出する方法が考えられる。図 2 は、閾値 901 を示している。例えば、閾値 901 以上の区間が、検出対象区間として、検出される。すなわち、検出対象区間が、音声の区間 50

として、検出される。

【 0 0 1 5 】

ここで、雑音のパワーが急激に上がる場合がある。例えば、図 2 は、時刻 t 9 0 以降に雑音のパワーが急激に上がったことを示している。例えば、図 2 の範囲 9 0 2 は、雑音を示している。雑音のパワーが急激に上がった場合、時刻 t 9 0 以降の区間が検出対象区間として検出される。図 2 は、雑音のパワーが閾値を超えているため、音声以外に雑音も検出対象とされることを示している。

【 0 0 1 6 】

このように、図 2 の方法では、検出対象が精度良く検出できない。そこで、以下、検出対象が精度良く検出できる方法を説明する。

【 0 0 1 7 】

図 3 は、実施の形態 1 の情報処理装置が有する機能ブロック図である。情報処理装置 1 0 0 は、取得部 1 1 0、制御部 1 2 0、及び出力部 1 3 0 を有する。

【 0 0 1 8 】

取得部 1 1 0、制御部 1 2 0、及び出力部 1 3 0 の一部又は全部は、プロセッサ 1 0 1 によって実現してもよい。取得部 1 1 0、制御部 1 2 0、及び出力部 1 3 0 の一部又は全部は、プロセッサ 1 0 1 が実行するプログラムのモジュールとして実現してもよい。例えば、プロセッサ 1 0 1 が実行するプログラムは、検出プログラムとも言う。例えば、検出プログラムは、記録媒体に記録されている。

【 0 0 1 9 】

取得部 1 1 0 は、音信号を取得する。例えば、音信号の音は、会議が行われる会議室内の音、電話の通話などである。また、例えば、音信号は、録音データに基づく信号である。

【 0 0 2 0 】

制御部 1 2 0 は、音信号に基づいて、時間経過に伴う音信号のパワーを算出する。言い換えれば、制御部 1 2 0 は、音信号に基づいて、音信号の時系列のパワーを算出する。以下、音信号のパワーは、音信号パワーと呼ぶ。なお、音信号パワーは、情報処理装置 1 0 0 以外の装置によって算出されてもよい。

【 0 0 2 1 】

制御部 1 2 0 は、音信号を複数の区間に分割する。制御部 1 2 0 は、音信号を均等に分割してもよいし、音信号を不均等に分割してもよい。制御部 1 2 0 は、音信号に基づいて、複数の区間のそれぞれの変動値を算出する。なお、変動値とは、区間時間当たりの変動量である。変動値は、区間時間当たりの音信号のパワーの変動量と考えてもよい。また、区間時間とは、1 つの区間に対応する時間である。

【 0 0 2 2 】

制御部 1 2 0 は、複数の区間の中で、変動値が予め設定された閾値以下の区間を特定する。制御部 1 2 0 は、音信号に基づいて、特定された区間における音信号のパワーを算出する。すなわち、制御部 1 2 0 は、音信号に基づいて、特定された区間の音信号のパワーを算出する。制御部 1 2 0 は、特定された区間における音信号のパワーの中から最大値を特定する。制御部 1 2 0 は、最大値に基づく値を検出閾値に設定する。言い換えれば、制御部 1 2 0 は、最大値以上の値を検出閾値に設定する。例えば、制御部 1 2 0 は、予め決められた値を最大値に加算した値を検出閾値に設定する。制御部 1 2 0 は、音信号パワーの中で検出閾値以上の区間を検出対象区間として検出する。

【 0 0 2 3 】

出力部 1 3 0 は、検出対象区間を示す情報を出力する。例えば、出力部 1 3 0 は、検出対象区間を示す情報をディスプレイに出力する。また、例えば、出力部 1 3 0 は、情報処理装置 1 0 0 に接続可能な外部装置に検出対象区間を示す情報を出力する。また、例えば、出力部 1 3 0 は、印刷装置を介して、紙媒体に検出対象区間を示す情報を出力する。

【 0 0 2 4 】

次に、情報処理装置 1 0 0 が実行する処理を、フローチャートを用いて説明する。

図 4 は、実施の形態 1 の情報処理装置が実行する処理の例を示すフローチャートである。

10

20

30

40

50

(ステップ S 1 1) 取得部 1 1 0 は、音信号を取得する。
 (ステップ S 1 2) 制御部 1 2 0 は、音信号をフレーム単位に分割し、フレーム毎にパワーを算出する。なお、例えば、フレームは、1 0 m s e c である。
 つまり、ステップ S 1 2 の処理では、音信号パワーが算出される。これにより、例えば、音信号パワーが、グラフで表現できる。

【 0 0 2 5 】

(ステップ S 1 3) 制御部 1 2 0 は、音信号を複数の区間に分割する。例えば、制御部 1 2 0 は、グラフで表現される音信号パワーを複数の区間に分割してもよい。なお、ステップ S 1 2 の複数のフレームは、1 つの区間に属する。

【 0 0 2 6 】

(ステップ S 1 4) 制御部 1 2 0 は、音信号に基づいて、区間毎に変動値を算出する。また、制御部 1 2 0 は、音信号に基づいて、区間毎に分散値を算出してもよい。
 分散値の算出を説明する。まず、区間における音信号のパワー m が、式 (1) により、算出される。 P は、パワーである。 i は、フレーム番号である。また、 i は、1 ~ N の値である。

【 0 0 2 7 】

【数 1】

$$m = \frac{1}{N} \sum_{i=1}^N P_i \dots (1)$$

【 0 0 2 8 】

そして、分散値 v は、式 (2) を用いて算出される。

【 0 0 2 9 】

【数 2】

$$v = \frac{1}{N} \sum_{i=1}^N (P_i - m)^2 \dots (2)$$

【 0 0 3 0 】

(ステップ S 1 5) 制御部 1 2 0 は、変動値が予め設定された閾値以下の区間を特定する。分散値が算出された場合、制御部 1 2 0 は、分散値が予め設定された閾値以下の区間を特定する。

(ステップ S 1 6) 制御部 1 2 0 は、特定した区間における音信号のパワーを、式 (1) を用いて算出する。

(ステップ S 1 7) 制御部 1 2 0 は、区間毎に算出されたパワーの中で最大値のパワーを特定する。制御部 1 2 0 は、当該最大値以上の値を検出閾値に設定する。

【 0 0 3 1 】

(ステップ S 1 8) 制御部 1 2 0 は、音信号パワーの中で検出閾値以上の区間を、音声区間として検出する。

(ステップ S 1 9) 出力部 1 3 0 は、音声区間を示す情報を出力する。例えば、出力部 1 3 0 は、音声区間の開始時刻と終了時刻を出力する。

【 0 0 3 2 】

図 5 は、実施の形態 1 の情報処理装置が実行する処理の具体例を示す。図 5 は、制御部 1 2 0 によって算出された音信号パワー 1 1 のグラフを示している。例えば、図 5 のグラフの縦軸は、d B である。図 5 のグラフの横軸は、時間である。図 5 は、時刻 t_1 以降に雑音のパワーが急激に上がったことを示している。

【 0 0 3 3 】

また、図 5 のグラフは、音声度 1 2 を示している。音声度については、実施の形態 2 で説明する。

10

20

30

40

50

例えば、制御部 120 は、音信号パワー 11 を複数の区間に分割する。制御部 120 は、区間毎に変動値を算出する。制御部 120 は、変動値が予め設定された閾値以下の区間を特定する。例えば、制御部 120 は、変動値が予め設定された閾値以下の区間 13a ~ 13e を特定する。これにより、例えば、区間 14 は、除外される。なお、区間 14 は、音声区間である。よって、制御部 120 は、音声区間以外の区間を特定する。すなわち、制御部 120 は、雑音区間を特定する。ここで、以下の説明では、区間 13a ~ 13e が特定されたものとする。

【0034】

制御部 120 は、区間 13a ~ 13e のパワーを、式(1)を用いて算出する。制御部 120 は、区間 13a ~ 13e のパワーの中で最大値のパワーを特定する。制御部 120 は、当該最大値以上の値を検出閾値に設定する。図5は、検出閾値 15 を示している。制御部 120 は、音信号パワー 11 の中で検出閾値 15 以上の区間を、音声区間として検出する。例えば、制御部 120 は、区間 14 を検出する。出力部 130 は、音声区間を示す情報を出力する。

10

【0035】

実施の形態1によれば、情報処理装置 100 は、雑音のパワーが急激に上がった場合でも、検出閾値を雑音のパワー以上に設定する。そのため、情報処理装置 100 は、雑音区間を検出対象区間として、検出しない。例えば、情報処理装置 100 は、区間 13a ~ 13e を検出しない。そして、情報処理装置 100 は、音声区間を検出する。よって、情報処理装置 100 は、検出対象である音声を精度良く検出できる。

20

【0036】

実施の形態2 .

次に、実施の形態2を説明する。実施の形態2では、実施の形態1と相違する事項を主に説明する。そして、実施の形態2では、実施の形態1と共通する事項の説明を省略する。実施の形態2の説明では、図1, 3を参照する。

【0037】

図6は、実施の形態2の情報処理装置が有する機能ブロック図である。図3に示される構成と同じ図6の構成は、図3に示される符号と同じ符号を付している。

情報処理装置 100a は、制御部 120a を有する。制御部 120a については、後で説明する。

30

【0038】

図7は、実施の形態2の情報処理装置が実行する処理の例を示すフローチャートである。

(ステップS21) 取得部 110 は、音信号を取得する。

(ステップS22) 制御部 120a は、音信号をフレーム単位に分割し、フレーム毎にパワーを算出する。言い換えれば、制御部 120a は、音信号パワーを算出する。

【0039】

(ステップS23) 制御部 120a は、音信号をフレーム単位に分割し、フレーム毎に音声度を算出する。音声度は、音声らしさの度合いである。例えば、制御部 120a は、GMM (Gaussian mixture model)、DNN (Deep Neural Network) などを用いて、音声度を算出する。

40

【0040】

(ステップS24) 制御部 120a は、音信号を複数の区間に分割する。例えば、制御部 120a は、音信号パワーを複数の区間に分割してもよい。

(ステップS25) 制御部 120a は、音信号に基づいて、区間毎に変動値と音声度を算出する。例えば、制御部 120a は、複数の区間のうちの第1の区間の変動値と音声度を算出する。このように、制御部 120a は、同じ区間の変動値と音声度を算出する。

【0041】

ここで、区間の音声度の算出について説明する。例えば、制御部 120a は、1つの区間に属する複数のフレームのそれぞれの音声度の平均値を、当該区間の音声度として、算出する。制御部 120a は、同様に、区間毎に音声度を算出する。

50

【 0 0 4 2 】

このように、制御部 1 2 0 a は、複数の区間のそれぞれの音声度を音信号に基づいて算出する。詳細には、制御部 1 2 0 a は、複数の区間のそれぞれの音声度を、G M M、D N N などの予め決められた方法と音信号とに基づいて算出する。

【 0 0 4 3 】

(ステップ S 2 6) 制御部 1 2 0 a は、複数の区間の中で、変動値が予め設定された閾値以下であり、かつ音声度が音声度閾値以下である区間を特定する。なお、音声度閾値は、予め設定された閾値である。

(ステップ S 2 7) 制御部 1 2 0 a は、特定した区間における音信号のパワーを、式 (1) を用いて算出する。

10

【 0 0 4 4 】

(ステップ S 2 8) 制御部 1 2 0 a は、区間毎に算出されたパワーの中で最大値のパワーを特定する。制御部 1 2 0 a は、当該最大値以上の値を検出閾値に設定する。

(ステップ S 2 9) 制御部 1 2 0 a は、音信号の中で検出閾値以上の区間を、音声区間として検出する。

(ステップ S 3 0) 出力部 1 3 0 は、音声区間を示す情報を出力する。

【 0 0 4 5 】

図 8 は、実施の形態 2 の情報処理装置が実行する処理の具体例を示す。図 8 は、制御部 1 2 0 a によって算出された音信号パワー 2 1 のグラフを示している。また、図 8 は、制御部 1 2 0 a によって算出された音声度 2 2 のグラフを示している。このように、図 8 では、音信号パワー 2 1 のグラフと音声度 2 2 のグラフとが混在している状態を示している。音信号パワー 2 1 のグラフと音声度 2 2 のグラフとは、分かれていてもよい。図 8 の横軸は、時間を示す。

20

【 0 0 4 6 】

ここで、例えば、図 8 の縦軸が示す 0 に対応する音声度は、音声らしさの度合いが 5 0 % くらいであることを意味する。よって、例えば、0 よりも大きい値に対応する音声度の区間は、音声区間と考えてもよい。また、例えば、0 よりも小さい値に対応する音声度の区間は、雑音区間と考えてもよい。

制御部 1 2 0 a は、音信号パワー 2 1 を複数の区間に分割する。制御部 1 2 0 a は、区間毎に変動値を算出する。また、制御部 1 2 0 a は、区間毎に音声度を算出する。

30

【 0 0 4 7 】

制御部 1 2 0 a は、変動値が予め設定された閾値以下であり、かつ音声度が音声度閾値以下である区間を特定する。ここで、音声度が音声度閾値以下である区間を、説明する。図 8 は、音声度閾値 2 3 を示している。例えば、音声度が音声度閾値 2 3 以下である区間は、区間 2 4 a ~ 2 4 e である。例えば、変動値が当該閾値以下であり、かつ音声度が音声度閾値 2 3 以下である区間は、区間 2 5 a ~ 2 5 e である。以下の説明では、区間 2 5 a ~ 2 5 e が特定されたものとする。

【 0 0 4 8 】

制御部 1 2 0 a は、区間 2 5 a ~ 2 5 e のパワーを、式 (1) を用いて算出する。制御部 1 2 0 a は、区間 2 5 a ~ 2 5 e のパワーの中で最大値のパワーを特定する。制御部 1 2 0 a は、当該最大値以上の値を検出閾値に設定する。図 8 は、検出閾値 2 6 を示している。制御部 1 2 0 a は、音信号パワー 2 1 の中で検出閾値 2 6 以上の区間を、音声区間として検出する。出力部 1 3 0 は、音声区間を示す情報を出力する。

40

【 0 0 4 9 】

実施の形態 2 によれば、情報処理装置 1 0 0 a は、音声度を用いることで、“あー”などの音声の音信号パワーが一定の区間を誤って雑音区間とみなすことを防止できる。

【 0 0 5 0 】

実施の形態 3 .

次に、実施の形態 3 を説明する。実施の形態 3 では、実施の形態 1 , 2 と相違する事項を主に説明する。そして、実施の形態 3 では、実施の形態 1 , 2 と共通する事項の説明を省

50

略する。実施の形態 3 の説明では、図 1 , 3 , 7 を参照する。

【 0 0 5 1 】

図 9 は、実施の形態 3 の情報処理装置が有する機能ブロック図である。図 3 に示される構成と同じ図 9 の構成は、図 3 に示される符号と同じ符号を付している。

情報処理装置 1 0 0 b は、制御部 1 2 0 b を有する。制御部 1 2 0 b については、後で説明する。

【 0 0 5 2 】

図 1 0 は、実施の形態 3 の情報処理装置が実行する処理の例を示すフローチャートである。図 1 0 の処理では、ステップ S 2 6 a , 2 6 b , 2 7 a , 2 8 a が実行される点が、図 7 の処理と異なる。そのため、図 1 0 では、ステップ S 2 6 a , 2 6 b , 2 7 a , 2 8 a を説明する。図 1 0 における他のステップについては、図 7 のステップ番号と同じ番号を付することによって、処理の説明を省略する。なお、ステップ S 2 1 ~ 2 5 とステップ S 2 9 , 3 0 は、制御部 1 2 0 b によって実行される。

10

【 0 0 5 3 】

(ステップ S 2 6 a) 制御部 1 2 0 b は、複数の区間の中で、変動値が予め設定された閾値以下である区間を特定する。

(ステップ S 2 6 b) 制御部 1 2 0 b は、特定した区間の音声度を小さい順に並べる。なお、特定した区間の音声度は、ステップ S 2 5 で算出されている。

制御部 1 2 0 b は、小さい順に、予め設定された個数の区間を選択する。以下、予め設定された個数は、N 個と表現する。なお、N は、正の整数である。

20

このように、制御部 1 2 0 b は、小さい順に上位 N 個の区間を選択する。

【 0 0 5 4 】

(ステップ S 2 7 a) 制御部 1 2 0 b は、音信号に基づいて、上位 N 個の区間における音信号のパワーを算出する。具体的には、制御部 1 2 0 b は、上位 N 個の区間における音信号のパワーを、式 (1) を用いて算出する。

(ステップ S 2 8 a) 制御部 1 2 0 b は、上位 N 個の区間における音信号のパワーの中で最大値を特定する。制御部 1 2 0 b は、当該最大値以上の値を検出閾値に設定する。

【 0 0 5 5 】

ここで、実施の形態 2 のように音声度閾値が設定され、1 以上の区間が検出される。しかし、音声度閾値の値又は音声度によっては、1 以上の区間が検出されない場合が考えられる。このような場合、実施の形態 3 が有効である。実施の形態 3 によれば、N 個の区間が選択される。そして、情報処理装置 1 0 0 b は、ステップ S 2 9 で音声区間を検出する。これにより、情報処理装置 1 0 0 b は、検出対象である音声精度良く検出できる。

30

【 0 0 5 6 】

実施の形態 4 .

次に、実施の形態 4 を説明する。実施の形態 4 では、実施の形態 1 と相違する事項を主に説明する。そして、実施の形態 4 では、実施の形態 1 と共通する事項の説明を省略する。実施の形態 4 の説明では、図 1 , 3 を参照する。

【 0 0 5 7 】

図 1 1 は、実施の形態 4 の情報処理装置が有する機能ブロック図である。図 3 に示される構成と同じ図 1 1 の構成は、図 3 に示される符号と同じ符号を付している。

40

情報処理装置 1 0 0 c は、制御部 1 2 0 c を有する。制御部 1 2 0 c については、後で説明する。

【 0 0 5 8 】

図 1 2 は、実施の形態 4 の情報処理装置が実行する処理の例を示すフローチャート (その 1) である。

(ステップ S 3 1) 取得部 1 1 0 は、音信号を取得する。

(ステップ S 3 2) 制御部 1 2 0 c は、音信号をフレーム単位に分割し、フレーム毎にパワーを算出する。言い換えれば、制御部 1 2 0 c は、音信号パワーを算出する。

【 0 0 5 9 】

50

(ステップ S 3 3) 制御部 1 2 0 c は、音信号を複数の区間に分割する。例えば、制御部 1 2 0 c は、音信号パワーを複数の区間に分割してもよい。

(ステップ S 3 4) 制御部 1 2 0 c は、音信号に基づいて、区間毎に変動値を算出する。
【 0 0 6 0 】

(ステップ S 3 5) 制御部 1 2 0 c は、複数の区間の中で、変動値が予め設定された閾値以下である区間を特定する。

(ステップ S 3 6) 制御部 1 2 0 c は、特定した区間における音信号のパワーを、式 (1) を用いて算出する。そして、制御部 1 2 0 c は、処理をステップ S 4 1 に進める。

【 0 0 6 1 】

図 1 3 は、実施の形態 4 の情報処理装置が実行する処理の例を示すフローチャート (その 2) である。 10

(ステップ S 4 1) 制御部 1 2 0 c は、ステップ S 3 5 で特定した区間の中から 1 つの区間を選択する。

(ステップ S 4 2) 制御部 1 2 0 c は、選択した区間における音信号のパワー以上を仮検出閾値に設定する。なお、選択した区間における音信号のパワーは、ステップ S 3 6 で算出されている。

【 0 0 6 2 】

(ステップ S 4 3) 制御部 1 2 0 c は、音信号パワーの中で仮検出閾値以上の区間の数を検出する。

(ステップ S 4 4) 制御部 1 2 0 c は、ステップ S 3 5 で特定した全ての区間を選択したか否かを判定する。全ての区間を選択した場合、制御部 1 2 0 c は、処理をステップ S 4 5 に進める。選択していない区間がある場合、制御部 1 2 0 c は、処理をステップ S 4 1 に進める。 20

【 0 0 6 3 】

このように、制御部 1 2 0 c は、ステップ S 3 5 で特定した区間毎に、区間における音信号のパワーに基づく値を仮検出閾値に設定し、音信号パワーの中で仮検出閾値以上の区間の数を検出する。

【 0 0 6 4 】

(ステップ S 4 5) 制御部 1 2 0 c は、ステップ S 3 5 で特定した区間毎に設定された仮検出閾値の中から、ステップ S 4 3 で検出された区間の数が最大であるときの仮検出閾値を検出閾値として検出する。 30

(ステップ S 4 6) 制御部 1 2 0 c は、ステップ S 4 5 で検出された仮検出閾値を用いて検出されたときの区間を、音声区間として検出する。言い換えれば、制御部 1 2 0 c は、検出閾値を用いて検出されたときの区間を、音声区間として検出する。

(ステップ S 4 7) 出力部 1 3 0 は、音声区間を示す情報を出力する。

【 0 0 6 5 】

図 1 4 は、実施の形態 4 の情報処理装置が実行する処理の具体例 (その 1) を示す。図 1 4 は、制御部 1 2 0 c によって算出された音信号パワー 3 1 のグラフを示している。図 1 4 は、制御部 1 2 0 c がステップ S 3 5 で特定した区間 3 2 a ~ 3 2 e を示している。 40

【 0 0 6 6 】

制御部 1 2 0 c は、区間 3 2 a ~ 3 2 e の中から区間 3 2 a を選択する。制御部 1 2 0 c は、区間 3 2 a のパワー以上を仮検出閾値に設定する。図 1 4 は、設定された仮検出閾値 3 3 を示している。制御部 1 2 0 c は、音信号パワー 3 1 の中で、仮検出閾値 3 3 以上の区間を検出する。例えば、制御部 1 2 0 c は、区間 A 1 ~ A 3 を検出する。すなわち、制御部 1 2 0 c は、3 つの区間を検出する。

【 0 0 6 7 】

図 1 5 は、実施の形態 4 の情報処理装置が実行する処理の具体例 (その 2) を示す。次に、制御部 1 2 0 c は、区間 3 2 b を選択する。制御部 1 2 0 c は、区間 3 2 b のパワー以上を仮検出閾値に設定する。図 1 5 は、設定された仮検出閾値 3 4 を示している。制御部 1 2 0 c は、音信号パワー 3 1 の中で、仮検出閾値 3 4 以上の区間を検出する。例えば、 50

制御部 120c は、区間 B1 ~ B21 を検出する。すなわち、制御部 120c は、21 の区間を検出する。

【0068】

制御部 120c は、区間 32c ~ 32e も同様の処理を実行する。

制御部 120c は、ステップ S43 で検出された区間の数が最大であるときの仮検出閾値を検出する。制御部 120c は、ステップ S45 で検出された仮検出閾値を用いて検出されたときの区間を、音声区間として検出する。

【0069】

実施の形態 4 によれば、情報処理装置 100c は、複数の仮検出閾値を用いて音声区間を検出する。言い換えれば、情報処理装置 100c は、仮検出閾値を変動させて音声区間を検出する。例えば、実施の形態 1 のように、検出閾値を一意に決めるよりも仮検出閾値を変動させた方が、音声区間を検出する精度を高めることができる。

なお、検出した区間数が一番多いものを最終的な検出結果とする理由は、ノイズパワー（すなわち、雑音のパワー）が不適切な場合、実際の音声区間数よりも検出される区間数が減少するからである。すなわち、ノイズパワーが不適切に低い場合、複数の音声区間がまとめて 1 区間として検出されるので検出数が少なくなる。一方、ノイズパワーが不適切に高い場合、パワーの小さい音声区間が検出漏れになるため、やはり検出数が少なくなる。

【0070】

実施の形態 4 の変形例。

次に、実施の形態 4 の変形例を説明する。

図 16 は、実施の形態 4 の変形例を示すフローチャート（その 1）である。図 16 の処理では、ステップ S32a, 34a, 35a, 36a が実行される点が、図 12 の処理と異なる。そのため、図 16 では、ステップ S32a, 34a, 35a, 36a を説明する。図 16 における他のステップについては、図 12 のステップ番号と同じ番号を付することによって、処理の説明を省略する。

【0071】

（ステップ S32a）制御部 120c は、音信号をフレーム単位に分割し、フレーム毎に音声度を算出する。

（ステップ S34a）制御部 120c は、音信号に基づいて、区間毎に変動値と音声度を算出する。

（ステップ S35a）制御部 120c は、特定した区間の音声度を小さい順に並べる。制御部 120c は、小さい順に上位 N 個の区間を選択する。

【0072】

（ステップ S36a）制御部 120b は、音信号に基づいて、上位 N 個の区間における音信号のパワーを算出する。具体的には、制御部 120b は、上位 N 個の区間における音信号のパワーを、式（1）を用いて算出する。そして、制御部 120b は、処理をステップ S41a に進める。

【0073】

図 17 は、実施の形態 4 の変形例を示すフローチャート（その 2）である。図 17 の処理では、ステップ S41a, 42a, 44a が実行される点が、図 13 の処理と異なる。そのため、図 17 では、ステップ S41a, 42a, 44a を説明する。図 17 における他のステップについては、図 13 のステップ番号と同じ番号を付することによって、処理の説明を省略する。

【0074】

（ステップ S41a）制御部 120c は、上位 N 個の区間の中から 1 つの区間を選択する。

（ステップ S42a）制御部 120c は、選択した区間における音信号のパワー以上を仮検出閾値に設定する。なお、選択した区間における音信号のパワーは、ステップ S36a で算出されている。

【0075】

（ステップ S44a）制御部 120c は、上位 N 個の区間を選択したか否かを判定する。

上位N個の区間を選択した場合、制御部120cは、処理をステップS45に進める。選択していない区間がある場合、制御部120cは、処理をステップS41aに進める。このように、制御部120cは、上位N個の区間毎に、区間における音信号のパワーに基づく値を仮検出閾値に設定し、音信号パワーの中で仮検出閾値以上の区間の数を検出する。
【0076】

実施の形態4の変形例によれば、情報処理装置100cは、音声区間を検出する精度を高めることができる。

【0077】

実施の形態5

次に、実施の形態5を説明する。実施の形態5では、実施の形態1と相違する事項を主に説明する。そして、実施の形態5では、実施の形態1と共通する事項の説明を省略する。実施の形態5の説明では、図1, 3を参照する。

実施の形態1~4では、音声区間を検出対象区間として検出する場合を説明した。実施の形態5では、非正常雑音区間を検出対象区間として検出する場合を説明する。

【0078】

図18は、実施の形態5の情報処理装置が有する機能ブロック図である。図3に示される構成と同じ図18の構成は、図3に示される符号と同じ符号を付している。

情報処理装置100dは、制御部120dと出力部130dを有する。制御部120dと出力部130dについては、後で説明する。

【0079】

図19は、実施の形態5の情報処理装置が実行する処理の例を示すフローチャート(その1)である。

(ステップS51) 取得部110は、音信号を取得する。

(ステップS52) 制御部120dは、音信号をフレーム単位に分割し、フレーム毎にパワーを算出する。言い換えれば、制御部120dは、音信号パワーを算出する。

【0080】

(ステップS53) 制御部120dは、音信号をフレーム単位に分割し、フレーム毎に音声度を算出する。言い換えれば、制御部120aは、時間経過に伴う音声度を、GMM、DNNなどの予め決められた方法と音信号とに基づいて算出する。ここで、時間経過に伴う音声度は、時系列の音声度と表現してもよい。

【0081】

(ステップS54) 制御部120dは、音声度が音声度閾値以上の区間を特定する。これにより、制御部120dは、音声区間を特定する。なお、音声区間が特定されない場合、制御部120dは、音声度閾値を下げてよい。

(ステップS55) 制御部120dは、特定した区間以外の区間を特定する。これにより、制御部120dは、非正常雑音区間候補を特定する。

【0082】

また、制御部120dは、ステップS54とステップS55に変えて、次の処理を実行してもよい。制御部120dは、音声度が音声度閾値未満の区間を特定する。これにより、制御部120dは、非正常雑音区間候補を特定する。そして、制御部120dは、処理をステップS61に進める。

【0083】

図20は、実施の形態5の情報処理装置が実行する処理の例を示すフローチャート(その2)である。以下の説明では、1つの非正常雑音区間候補が特定されたものとする。また、複数の非正常雑音区間候補が特定された場合、図20の処理が非正常雑音区間候補の数だけ繰り返される。

【0084】

(ステップS61) 制御部120dは、1つの非正常雑音区間候補を複数の区間に分割する。なお、制御部120dは、非正常雑音区間候補を均等に分割してもよいし、非正常雑音区間候補を不均等に分割してもよい。

10

20

30

40

50

(ステップ S 6 2) 制御部 1 2 0 d は、音信号に基づいて、複数の区間のそれぞれの変動値を算出する。

【 0 0 8 5 】

(ステップ S 6 3) 制御部 1 2 0 d は、複数の区間の中で、変動値が予め設定された閾値以下である区間を特定する。

(ステップ S 6 4) 制御部 1 2 0 d は、音信号に基づいて、特定した区間における音信号のパワーを算出する。具体的には、制御部 1 2 0 d は、特定した区間における音信号のパワーを、式 (1) を用いて算出する。

【 0 0 8 6 】

(ステップ S 6 5) 制御部 1 2 0 d は、特定された区間における音信号のパワーの中から最大値を特定する。制御部 1 2 0 d は、当該最大値以上の値を検出閾値に設定する。

10

(ステップ S 6 6) 制御部 1 2 0 d は、非定常雑音区間候補内であり、かつ音信号パワーの中で検出閾値以上の区間を、非定常雑音区間として検出する。

(ステップ S 6 7) 出力部 1 3 0 d は、検出対象区間である非定常雑音区間を示す情報を出力する。例えば、出力部 1 3 0 は、非定常雑音区間の開始時刻と終了時刻を出力する。

【 0 0 8 7 】

図 2 1 は、実施の形態 5 の情報処理装置が実行する処理の具体例 (その 1) を示す。図 2 1 は、制御部 1 2 0 d によって算出された音信号パワー 4 1 のグラフを示している。また、図 2 1 は、音声度 4 2 のグラフを示している。さらに、図 1 2 は、音声度閾値 4 3 を示している。

20

制御部 1 2 0 d は、音声度が音声度閾値 4 3 以上の区間を特定する。図 2 1 は、特定された区間である音声区間を示している。

【 0 0 8 8 】

図 2 2 は、実施の形態 5 の情報処理装置が実行する処理の具体例 (その 2) を示す。制御部 1 2 0 d は、特定した区間以外の区間を特定する。図 2 2 は、特定された区間である非定常雑音区間候補を示している。

【 0 0 8 9 】

例えば、制御部 1 2 0 d は、非定常雑音区間候補 1 を複数の区間に分割する。制御部 1 2 0 d は、区間毎に変動値を算出する。制御部 1 2 0 d は、変動値が予め設定された閾値以下である区間を特定する。制御部 1 2 0 d は、特定した区間における音信号のパワーを算出する。制御部 1 2 0 d は、区間毎に算出されたパワーの中で最大値のパワーを特定する。制御部 1 2 0 d は、当該最大値以上の値を検出閾値に設定する。制御部 1 2 0 d は、非定常雑音区間候補 1 内であり、かつ音信号パワー 4 1 の中で検出閾値以上の区間を、非定常雑音区間として検出する。

30

情報処理装置 1 0 0 d は、同様に、非定常雑音区間候補 2 ~ 6 の中から非定常雑音区間を検出することができる。

【 0 0 9 0 】

実施の形態 5 によれば、情報処理装置 1 0 0 d は、音声度を用いることで音声を安定して検出することができる。また、情報処理装置 1 0 0 d は、非定常雑音の検出に関して、音声以外の区間を対象として、非定常雑音候補区間ごとに検出閾値を設定するので、高精度に非定常雑音を検出できる。

40

【 0 0 9 1 】

以上に説明した各実施の形態における特徴は、互いに適宜組み合わせることができる。

【 符号の説明 】

【 0 0 9 2 】

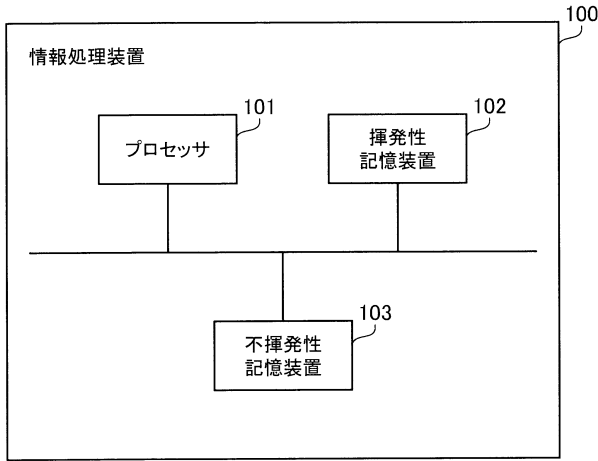
1 1 音信号パワー、 1 2 音声度、 1 3 a ~ 1 3 e 区間、 1 4 区間、 1 5 検出閾値、 2 1 音信号パワー、 2 2 音声度、 2 3 音声度閾値、 2 4 a ~ 2 4 e 区間、 2 5 a ~ 2 5 e 区間、 2 6 検出閾値、 3 1 音信号パワー、 3 2 a ~ 3 2 e 区間、 3 3 仮検出閾値、 3 4 仮検出閾値、 4 1 音信号パワー、 4 2 音声度、 4 3 音声度閾値、 1 0 0 , 1 0 0 a , 1 0 0 b , 1 0 0 c , 1 0 0 d 情報処理

50

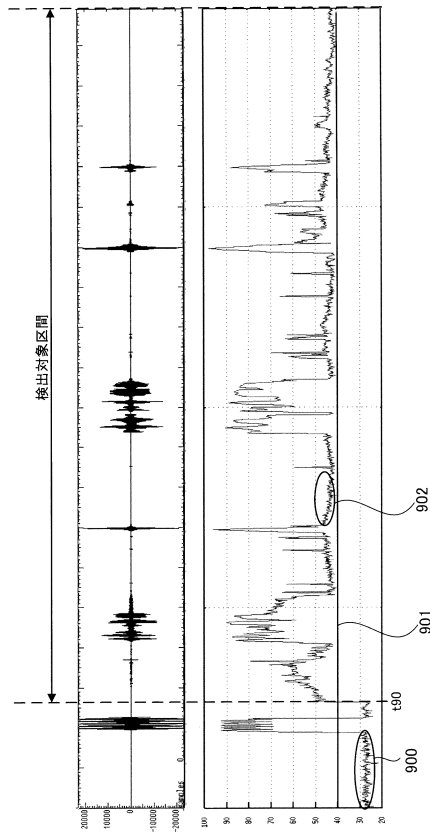
装置、 101 プロセッサ、 102 揮発性記憶装置、 103 不揮発性記憶装置、
110 取得部、 120, 120a, 120b, 120c, 120d 制御部、 130
, 130d 出力部、 900 範囲、 901 閾値、 902 範囲。

【図面】

【図1】



【図2】



10

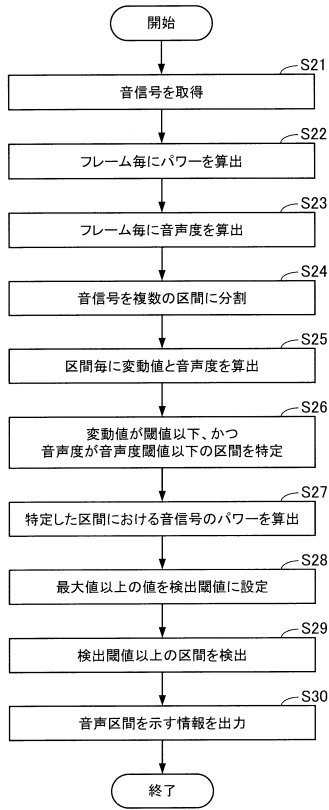
20

30

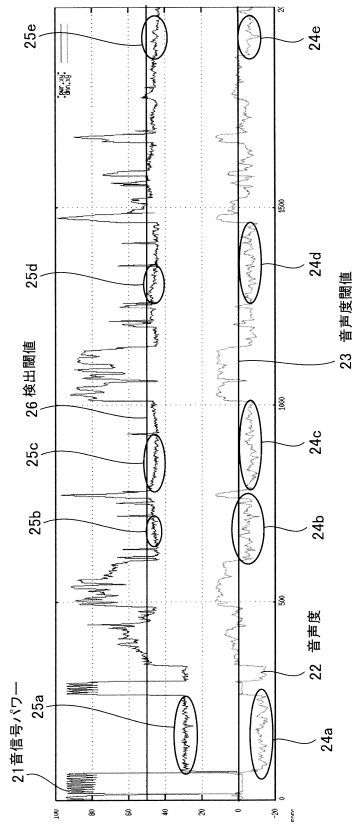
40

50

【図 7】



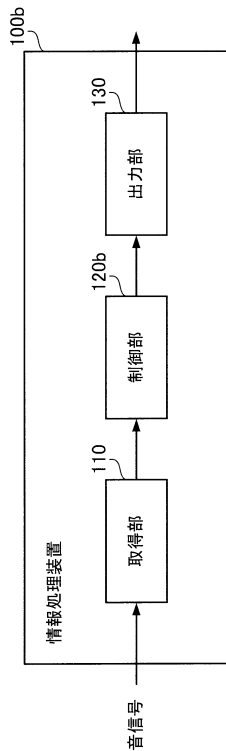
【図 8】



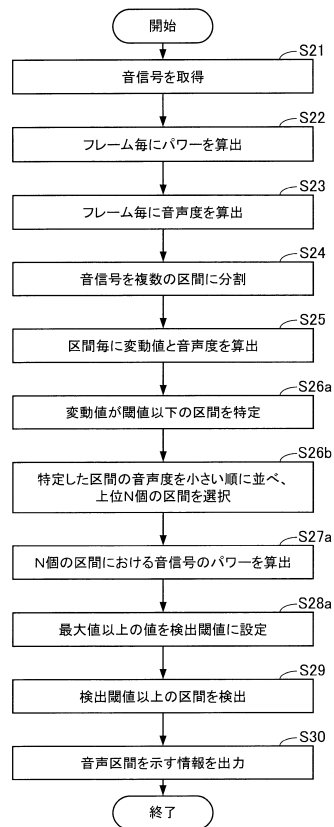
10

20

【図 9】



【図 10】

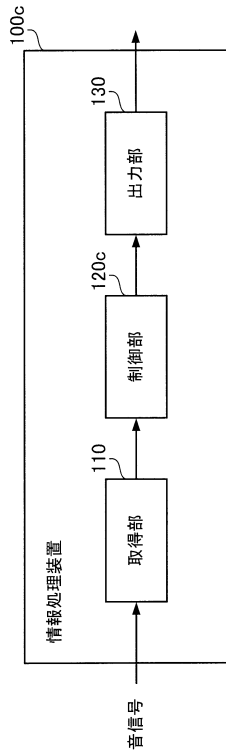


30

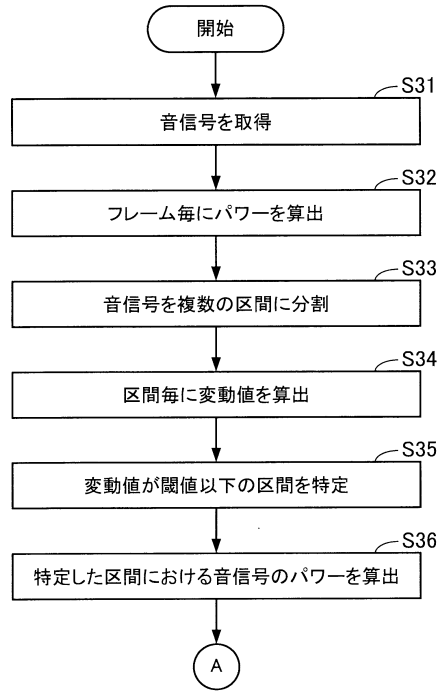
40

50

【図 1 1】



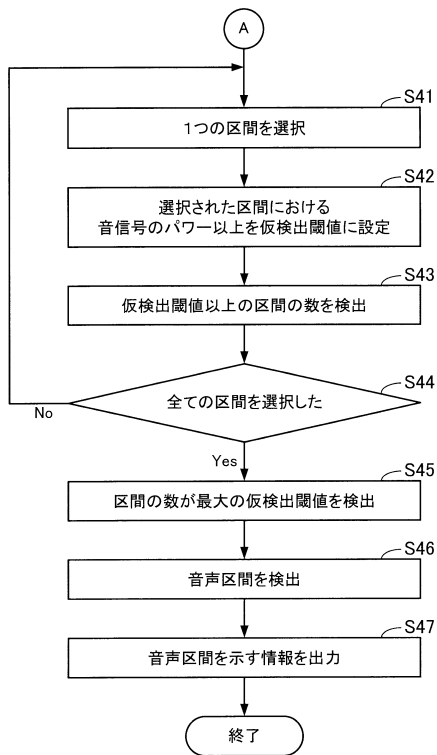
【図 1 2】



10

20

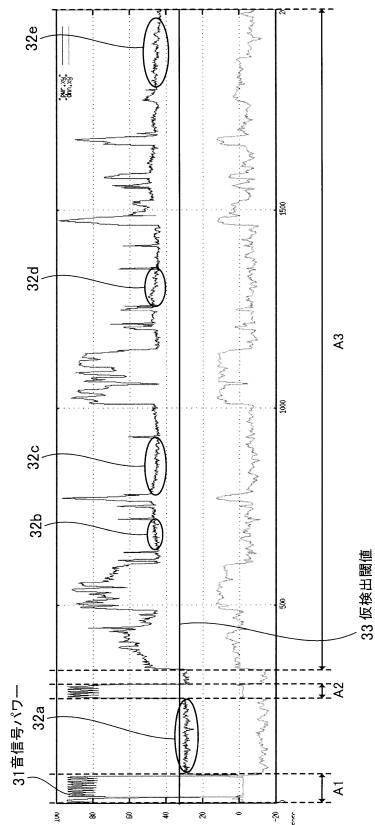
【図 1 3】



30

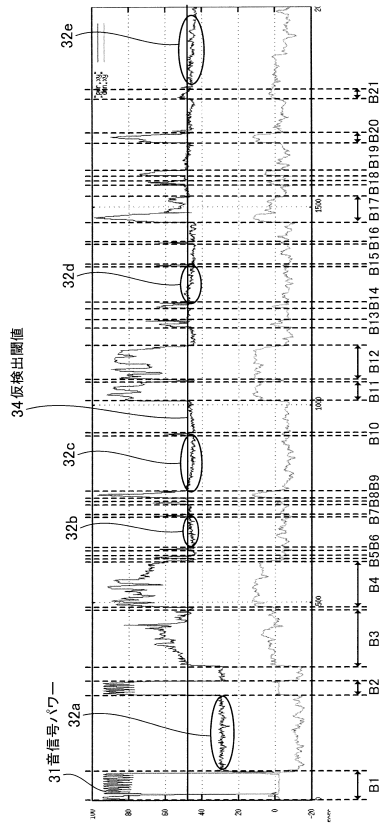
40

【図 1 4】

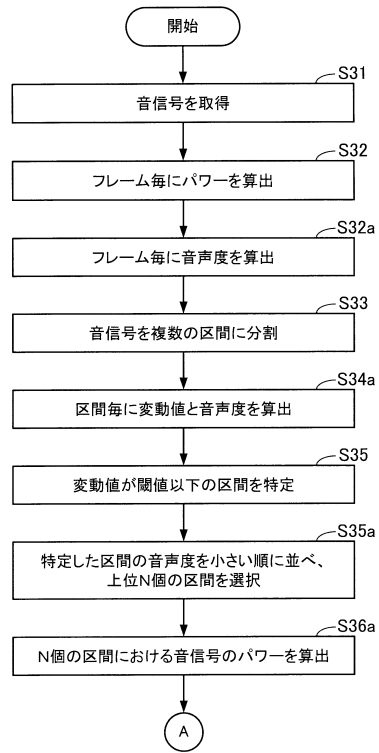


50

【図15】



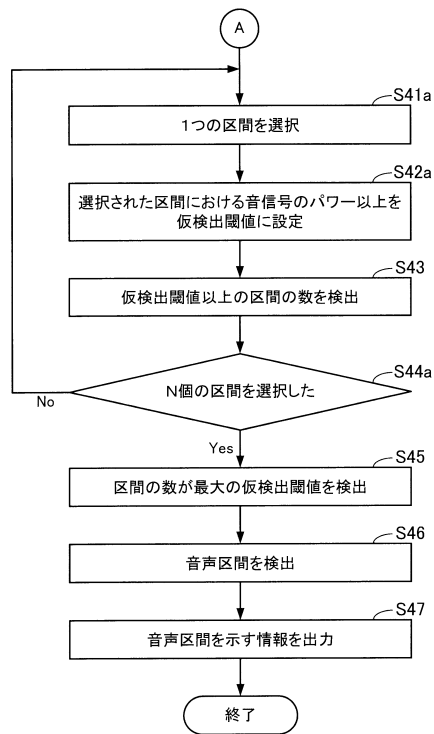
【図16】



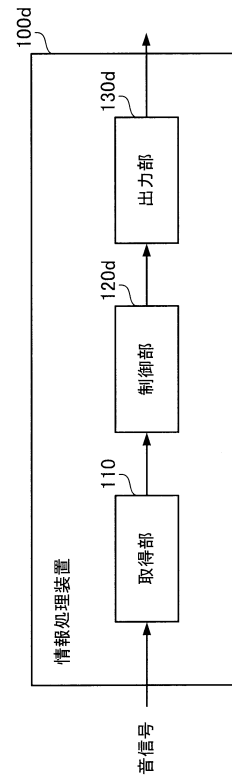
10

20

【図17】



【図18】

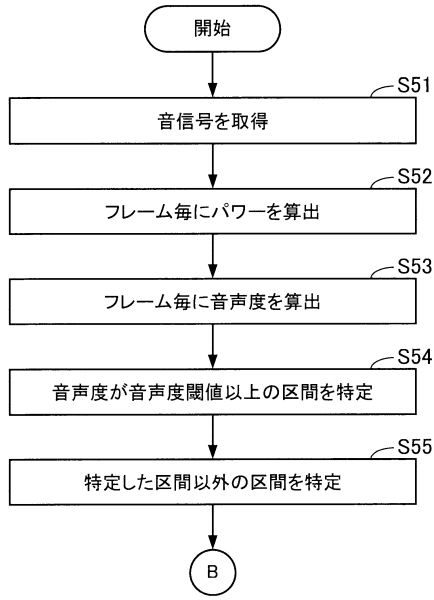


30

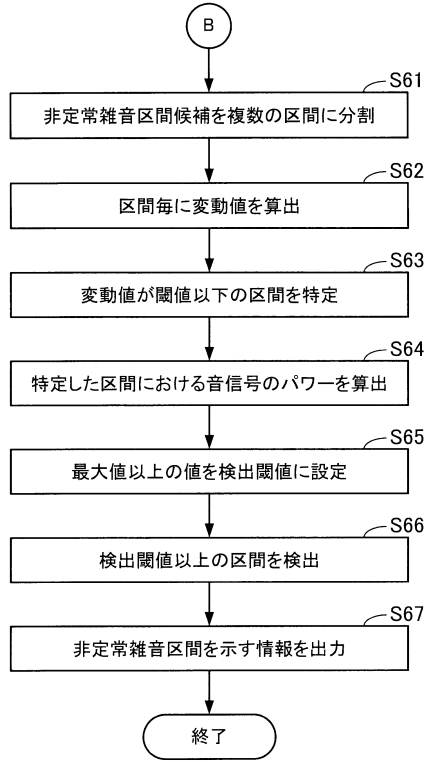
40

50

【図 19】



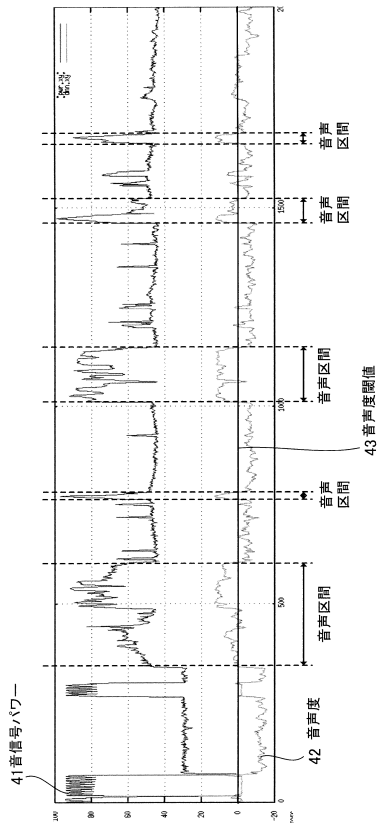
【図 20】



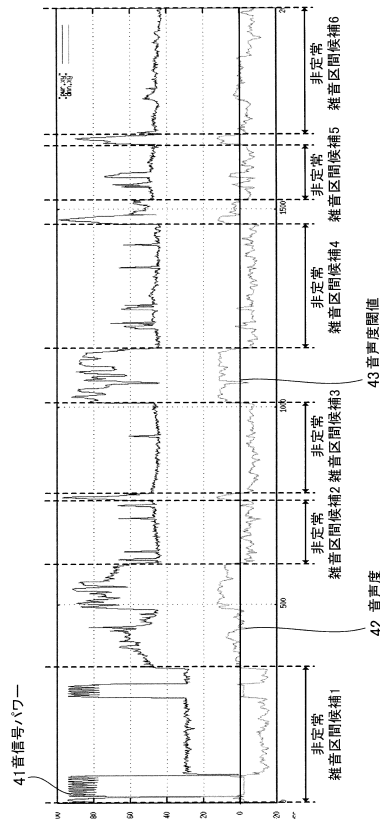
10

20

【図 21】



【図 22】



30

40

50

フロントページの続き

- (56)参考文献 特開昭 6 2 - 2 6 5 6 9 9 (J P , A)
特許第 3 8 1 2 8 8 7 (J P , B 2)
国際公開第 2 0 0 9 / 0 7 8 0 9 3 (W O , A 1)
国際公開第 2 0 1 2 / 0 3 6 3 0 5 (W O , A 1)
- (58)調査した分野 (Int.Cl. , D B 名)
G 1 0 L 1 5 / 0 0 - 2 5 / 9 3