

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7017873号

(P7017873)

(45)発行日 令和4年2月9日(2022.2.9)

(24)登録日 令和4年2月1日(2022.2.1)

(51)国際特許分類

H 0 4 R 3/02 (2006.01)

F I

H 0 4 R 3/02

請求項の数 16 (全21頁)

(21)出願番号	特願2017-136325(P2017-136325)	(73)特許権者	321003371 L I N E 株式会社 東京都新宿区四谷一丁目6番1号
(22)出願日	平成29年7月12日(2017.7.12)	(74)代理人	100107766 弁理士 伊東 忠重
(65)公開番号	特開2018-19396(P2018-19396A)	(74)代理人	100070150 弁理士 伊東 忠彦
(43)公開日	平成30年2月1日(2018.2.1)	(74)代理人	100135079 弁理士 宮崎 修
審査請求日	令和2年7月2日(2020.7.2)	(72)発明者	カン インギュ 大韓民国 13591 京畿道城南市盆唐 区ファンサエル路360番街42 11階 ラインプラス株式会社内
(31)優先権主張番号	10-2016-0095045	審査官	富澤 直樹
(32)優先日	平成28年7月26日(2016.7.26)		
(33)優先権主張国・地域又は機関	韓国(KR)		

最終頁に続く

(54)【発明の名称】 音質改善方法、音質改善方法を実行させるためのコンピュータプログラム、および電子機器

(57)【特許請求の範囲】

【請求項1】

コンピュータによって実現される電子機器における音質改善方法を実行させるためのコンピュータプログラムであって、

前記音質改善方法は、

前記電子機器へのマイク入力信号および前記電子機器からのスピーカ出力信号を分析してソフトウェアの音質改善機能の要否を決定する段階、および

前記の決定されたソフトウェアの音質改善機能の要否に基づき、前記ソフトウェアの音質改善機能を活性化または非活性化する段階を含み、

前記ソフトウェアの音質改善機能の要否を決定する段階は、

前記マイク入力信号および前記スピーカ出力信号を分析してエコー区間を決定する段階、

前記マイク入力信号および前記スピーカ出力信号の全体区間から前記の決定されたエコー区間を除いた残りの区間のうち、前記マイク入力信号のエネルギーが前記マイク入力信号の平均エネルギーと既定の第1加重値との積よりも大きい区間を、ユーザ入力区間として決定する段階、

前記の決定されたユーザ入力区間における平均エネルギーを算出する段階、および

前記の算出されたユーザ入力区間における平均エネルギーが既定の範囲に属するかどうかを利用して、前記ソフトウェアの音質改善機能としてA G Cモジュールの要否を決定する段階

を含む、コンピュータプログラム。

【請求項 2】

前記ソフトウェアの音質改善機能は、A E C モジュール、N S モジュール、および A G C モジュールを含み、

前記ソフトウェアの音質改善機能の要否を決定する段階では、前記 A E C モジュール、前記 N S モジュール、および前記 A G C モジュールのうち少なくとも 1 つのモジュールの要否を決定し、

前記ソフトウェアの音質改善機能を活性化または非活性化する段階では、前記の決定された少なくとも 1 つのモジュールの要否に基づき、前記少なくとも 1 つのモジュールを活性化または非活性化する、請求項 1 に記載のコンピュータプログラム。

10

【請求項 3】

前記ソフトウェアの音質改善機能の要否を決定する段階は、

前記の決定されたエコー区間における平均エネルギーを算出する段階、および前記の算出されたエコー区間における平均エネルギーと既定の第 1 閾値とを比較した結果を利用して、前記ソフトウェアの音質改善機能として A E C モジュールの要否を決定する段階

を更に含む、請求項 1 または 2 に記載のコンピュータプログラム。

【請求項 4】

前記エコー区間を決定する段階では、

前記スピーカ出力信号の活性化区間において、前記マイク入力信号と前記スピーカ出力信号との相関度を分析することにより、前記エコー区間を決定する、請求項 1 に記載のコンピュータプログラム。

20

【請求項 5】

前記 A E C モジュールの要否を決定する段階では、前記マイク入力信号と前記スピーカ出力信号の間の遅延の連続性の有無を更に利用して、前記 A E C モジュールの要否を決定し、前記遅延の連続性は、前記マイク入力信号と前記スピーカ出力信号のそれぞれを既定の大きさのフレーム単位に分割したとき、前記フレーム単位ごとの遅延の値が同様に保たれることを意味する、請求項 3 に記載のコンピュータプログラム。

【請求項 6】

前記ソフトウェアの音質改善機能の要否を決定する段階は、

前記全体区間から前記の決定されたエコー区間および前記の決定されたユーザ入力区間を除いた残りの区間のうち、前記マイク入力信号のエネルギーが前記マイク入力信号の平均エネルギーと既定の第 2 加重値との積よりも小さい区間を、雑音区間として決定する段階、前記の決定された雑音区間における平均エネルギーを算出する段階、および前記の算出された雑音区間における平均エネルギーと既定の第 2 閾値とを比較した結果を利用して、前記ソフトウェアの音質改善機能として N S モジュールの要否を決定する段階を更に含む、請求項 1 乃至 5 のうちいずれか 1 項に記載のコンピュータプログラム。

30

【請求項 7】

電子機器における音質改善方法であって、

前記電子機器へのマイク入力信号および前記電子機器からのスピーカ出力信号を分析してソフトウェアの音質改善機能の要否を決定する段階、および

前記の決定されたソフトウェアの音質改善機能の要否に基づき、前記ソフトウェアの音質改善機能を活性化または非活性化する段階

40

を含み、

前記ソフトウェアの音質改善機能の要否を決定する段階は、

前記マイク入力信号および前記スピーカ出力信号を分析してエコー区間を決定する段階、前記マイク入力信号および前記スピーカ出力信号の全体区間から前記の決定されたエコー区間を除いた残りの区間のうち、前記マイク入力信号のエネルギーが前記マイク入力信号の平均エネルギーと既定の第 1 加重値との積よりも大きい区間を、ユーザ入力区間として決定する段階、

50

前記の決定されたユーザ入力区間における平均エネルギーを算出する段階、および
前記の算出されたユーザ入力区間における平均エネルギーが既定の範囲に属するかどうかを利用して、前記ソフトウェアの音質改善機能としてA G Cモジュールの要否を決定する段階

を含む、音質改善方法。

【請求項 8】

前記ソフトウェアの音質改善機能は、A E Cモジュール、N Sモジュール、およびA G Cモジュールを含み、

前記ソフトウェアの音質改善機能の要否を決定する段階では、前記A E Cモジュール、前記N Sモジュール、および前記A G Cモジュールのうち少なくとも1つのモジュールの要否を決定し、

前記ソフトウェアの音質改善機能を活性化または非活性化する段階では、前記の決定された少なくとも1つのモジュールの要否に基づき、前記少なくとも1つのモジュールを活性化または非活性化する、請求項 7 に記載の音質改善方法。

【請求項 9】

前記ソフトウェアの音質改善機能の要否を決定する段階は、

前記の決定されたエコー区間における平均エネルギーを算出する段階、および

前記の算出されたエコー区間における平均エネルギーと既定の第1閾値とを比較した結果を利用して、前記ソフトウェアの音質改善機能としてA E Cモジュールの要否を決定する段階

を更に含む、請求項 7 または 8 に記載の音質改善方法。

【請求項 10】

前記A E Cモジュールの要否を決定する段階では、

前記マイク入力信号と前記スピーカ出力信号の間の遅延の連続性の有無を更に利用して、

前記A E Cモジュールの要否を決定し、

前記遅延の連続性は、前記マイク入力信号と前記スピーカ出力信号のそれぞれを既定の大きさのフレーム単位に分割したとき、前記フレーム単位ごとの遅延の値が同様に保たれることを意味する、請求項 9 に記載の音質改善方法。

【請求項 11】

前記ソフトウェアの音質改善機能の要否を決定する段階は、

前記全体区間から前記の決定されたエコー区間および前記の決定されたユーザ入力区間を除いた残りの区間のうち、前記マイク入力信号のエネルギーが前記マイク入力信号の平均エネルギーと既定の第2加重値との積よりも小さい区間を、雑音区間として決定する段階、前記の決定された雑音区間における平均エネルギーを算出する段階、および

前記の算出された雑音区間における平均エネルギーと既定の第2閾値とを比較した結果を利用して、前記ソフトウェアの音質改善機能としてN Sモジュールの要否を決定する段階を更に含む、請求項 7 乃至 10 のうちいずれか1項に記載の音質改善方法。

【請求項 12】

電子機器であって、

コンピュータで読取可能な命令を格納するメモリ、および

前記コンピュータで読取可能な命令を実行するための少なくとも1つのプロセッサを含み、

前記少なくとも1つのプロセッサは、

前記電子機器へのマイク入力信号および前記電子機器からのスピーカ出力信号を分析してソフトウェアの音質改善機能の要否を決定し、

前記の決定されたソフトウェアの音質改善機能の要否に基づき、前記ソフトウェアの音質改善機能を活性化または非活性化するものであり、

前記少なくとも1つのプロセッサは、前記ソフトウェアの音質改善機能の要否を決定するために、

前記マイク入力信号および前記スピーカ出力信号を分析してエコー区間を決定し、

10

20

30

40

50

前記マイク入力信号および前記スピーカ出力信号の全体区間から前記の決定されたエコー区間を除いた残りの区間のうち、前記マイク入力信号のエネルギーが前記マイク入力信号の平均エネルギーと既定の第1加重値との積よりも大きい区間を、ユーザ入力区間として決定し、

前記の決定されたユーザ入力区間における平均エネルギーを算出し、

前記の算出されたユーザ入力区間における平均エネルギーが既定の範囲に属するかどうかを利用して、前記ソフトウェアの音質改善機能としてA G Cモジュールの要否を決定するものである、電子機器。

【請求項13】

前記ソフトウェアの音質改善機能は、A E Cモジュール、N Sモジュール、およびA G Cモジュールを含み、

前記少なくとも1つのプロセッサは、

前記A E Cモジュール、前記N Sモジュール、および前記A G Cモジュールのうち少なくとも1つのモジュールの要否を決定することにより、前記ソフトウェアの音質改善機能の要否を決定し、

前記の決定された少なくとも1つのモジュールの要否に基づき、前記少なくとも1つのモジュールを活性化または非活性化することにより、前記ソフトウェアの音質改善機能を活性化または非活性化するものである、請求項12に記載の電子機器。

【請求項14】

前記少なくとも1つのプロセッサは、前記ソフトウェアの音質改善機能の要否を決定するために、

前記の決定されたエコー区間における平均エネルギーを算出し、

前記の算出されたエコー区間における平均エネルギーと既定の第1閾値とを比較した結果を利用して、前記ソフトウェアの音質改善機能としてA E Cモジュールの要否を決定するものである、請求項12または13に記載の電子機器。

【請求項15】

前記少なくとも1つのプロセッサは、前記A E Cモジュールの要否を決定する際に、前記マイク入力信号と前記スピーカ出力信号の間の遅延の連続性の有無を更に利用して、前記A E Cモジュールの要否を決定するものであり、

前記遅延の連続性は、前記マイク入力信号と前記スピーカ出力信号のそれぞれを既定の大きさのフレーム単位に分割したとき、前記フレーム単位ごとの遅延の値が同様に保たれることを意味する、請求項14に記載の電子機器。

【請求項16】

前記少なくとも1つのプロセッサは、前記ソフトウェアの音質改善機能の要否を決定するために、

前記全体区間から前記の決定されたエコー区間および前記の決定されたユーザ入力区間を除いた残りの区間のうち、前記マイク入力信号のエネルギーが前記マイク入力信号の平均エネルギーと既定の第2加重値との積よりも小さい区間を、雑音区間として決定し、

前記の決定された雑音区間における平均エネルギーを算出し、

前記の算出された雑音区間における平均エネルギーと既定の第2閾値とを比較した結果を利用して、前記ソフトウェアの音質改善機能としてN Sモジュールの要否を決定するものである、請求項12乃至15のうちいずれか1項に記載の電子機器。

【発明の詳細な説明】

【技術分野】

【0001】

以下の説明は、音質改善方法、音質改善方法を実行させるためのコンピュータプログラム、および電子機器に関する。

【背景技術】

【0002】

最近、様々な種類のマルチメディア機器が市販されており、マルチメディア機器をサポー

10

20

30

40

50

トするための様々なアプリケーションも開発されている。このような様々なアプリケーションのうち、マルチメディア機器に備えられているマイクから入ってくるオーディオ信号を利用するアプリケーションでは、音質改善 (Audio Quality Enhancement) 機能が必須的に採用されている。それは、スピーカから出るエコーや周囲の雑音がユーザの音声入力に混ざってマイクに入ってくるからである。そのようなアプリケーションとして、例えば、電話やカラオケ、音声の録音や映像の録画、音声や音楽の認識のためのアプリケーションがある。

【0003】

一方、最近のマルチメディア機器の中には、ハードウェア自体において音質改善機能を有するものがあり、ハードウェア自体において音質改善機能を有するマルチメディア機器の種類や数はますます増えている傾向にある。また、ソフトウェア的に音質改善機能を提供する必要があるアプリケーションも存在する。

10

【0004】

音質改善機能は、マイクから入ってくる入力からエコーや雑音を取り除くため、ユーザの音声入力が損傷を受けるわけであり、音質改善機能が数回に渡って実行されると、損傷がそれだけ大きくなるという問題がある。したがって、ソフトウェアにおける音質改善機能を提供する必要があるアプリケーションの開発者は、ハードウェア自体において音質改善機能を有するマルチメディア機器を一つ一つ確認し、そのリストを生成および管理しなければならないという煩雑さを抱えている。また、ハードウェアの音質改善機能を有するマルチメディア機器にインストールされたアプリケーションにおいて、ソフトウェアの音質改善機能を提供しないために、その機能をオフ (off) とするように管理しなければならないという煩雑さもある。

20

【0005】

或いは、このような煩雑さを避けるために、ハードウェアの音質改善機能とソフトウェアの音質改善機能のいずれも機能することを放置する場合もあるが、上述したように、音質の損傷を受けることが避けられないという問題がある。

【先行技術文献】

【特許文献】

【0006】

【文献】国際公開第WO/2015/065001号

30

米国特許出願公開第2014/0019540A1号明細書

米国特許出願公開第2013/0332543A1号明細書

米国特許出願公開第2013/0260893A1号明細書

【発明の概要】

【発明が解決しようとする課題】

【0007】

マイクの入力信号とスピーカの出力信号を分析してリアルタイムでソフトウェアの音質改善機能の要否を判断し、ソフトウェアの音質改善機能を選択的に活性化または非活性化することができる、音質改善方法およびシステムを提供する。

【課題を解決するための手段】

40

【0008】

コンピュータによって実現される電子機器における音質改善方法を実行させるためのコンピュータプログラムであって、前記音質改善方法は、前記電子機器へのマイク入力信号および前記電子機器からのスピーカ出力信号を分析してソフトウェアの音質改善機能の要否を決定する段階、および前記の決定されたソフトウェアの音質改善機能の要否に基づき、ソフトウェアの音質改善機能を活性化または非活性化する段階を含む、コンピュータプログラムを提供する。

【0009】

電子機器における音質改善方法であって、前記電子機器へのマイク入力信号および前記電子機器からのスピーカ出力信号を分析してソフトウェアの音質改善機能の要否を決定する

50

段階、および前記の決定されたソフトウェアの音質改善機能の要否に基づき、ソフトウェアの音質改善機能を活性化または非活性化する段階を含む、音質改善方法を提供する。

【0010】

電子機器であって、コンピュータで読取可能な命令 (instructions) を格納するメモリ、および前記コンピュータで読取可能な命令を実行するための少なくとも1つのプロセッサを含み、前記少なくとも1つのプロセッサは、前記電子機器へのマイク入力信号および前記電子機器からのスピーカ出力信号を分析してソフトウェア音の質改善機能の要否を決定し、前記の決定されたソフトウェアの音質改善機能の要否に基づき、ソフトウェアの音質改善機能を活性化または非活性化する、電子機器を提供する。

【発明の効果】

10

【0011】

マイクの入力信号とスピーカの出力信号を分析してリアルタイムでソフトウェアの音質改善機能の要否を判断し、ソフトウェアの音質改善機能を選択的に活性化または非活性化することができる。

【図面の簡単な説明】

【0012】

【図1】本発明の一実施形態における電子機器の内部構成の例を示した図である。

【図2】本発明の一実施形態における電子機器のプロセッサが含む構成要素の例を示したブロック図である。

【図3】本発明の一実施形態における電子機器が実行する音質改善方法の例を示したフローチャートである。

20

【図4】本発明の一実施形態において、ソフトウェアの音質改善機能の要否を決定するための方法の例を示したフローチャートである。

【図5】本発明の一実施形態において、ソフトウェアの音質改善機能を活性化する過程の例を示した図である。

【発明を実施するための形態】

【0013】

以下、実施形態について、添付の図面を参照しながら詳しく説明する。

【0014】

本発明の実施形態に係る音質改善システムは、以下で説明される電子機器によって実現されて良く、本発明の実施形態に係る音質改善方法は、電子機器によって実行されて良い。例えば、電子機器には、本発明の一実施形態に係るコンピュータプログラムとして実現されるアプリケーションがインストールされて、駆動されて良く、電子機器は、駆動するアプリケーションの制御にしたがって、本発明の一実施形態に係る音質改善方法を実行して良い。

30

【0015】

このような電子機器は、コンピュータ装置によって実現される固定端末や移動端末であって良い。例えば、電子機器は、スマートフォンや、携帯電話、ナビゲーション、コンピュータ、ノートパソコン、デジタル放送用端末、PDA (Personal Digital Assistant)、PMP (Portable Multimedia Player)、タブレット等のように、少なくとも1つのプロセッサと少なくとも1つのメモリ、更に、データを格納するための永久格納所を少なくとも含む装置であって良い。

40

【0016】

図1は、本発明の一実施形態における電子機器の内部構成の例を示した図である。電子機器100は、プロセッサ110や、バス120、メモリ130、通信モジュール140、入力／出力インタフェース150を含んで良い。

【0017】

プロセッサ110は、基本的な算術や、ロジック、入出力演算を実行することにより、コンピュータプログラムの命令を処理するように構成されて良い。例えば、命令は、メモリ130または通信モジュール140より、そしてバス120を介してプロセッサ110に

50

提供されて良い。例えば、プロセッサ 110 は、メモリ 130 のような記録装置に格納されたプログラムコードにしたがって受信される命令を実行するように構成されて良い。

【0018】

バス 120 は、電子機器 100 の構成要素同士の通信やデータ送信を可能にする。例えば、バス 120 は、高速シリアルバス (High-speed Serial Bus)、パラレルバス (Parallel Bus)、SAN (Storage Area Network)、及び／又は、その他の好適な通信技術から構成されて良い。

【0019】

メモリ 130 は、コンピュータで読取可能な記録媒体であって、RAM (random access memory) や、ROM (read only memory)、ディスクドライブのような永久大容量記憶装置 (permanent mass storage device) を含んで良い。ここで、ROM と永久大容量記憶装置は、メモリ 130 から分離され、別の永久記憶装置として電子機器 100 に含まれても良い。また、メモリ 130 には、オペレーティングシステムや、少なくとも 1 つのプログラムコード (一例として、電気機器 100 にインストールされ駆動されるブラウザや、特定のサービスを提供するために電子機器 100 にインストールされたアプリケーション等のためのコード) が格納されて良い。このようなソフトウェア構成要素は、メモリ 130 とは別のコンピュータで読取可能な記録媒体からロードされても良い。このように別のコンピュータで読取可能な記録媒体は、フロッピー (登録商標) ドライブや、ディスク、テープ、DVD / CD-ROM ドライブ、メモリカード等のコンピュータで読取可能な記録媒体を含んで良い。他の実施形態において、ソフトウェア構成要素は、コンピュータで読取可能な記録媒体ではなく、通信モジュール 140 を通じてメモリ 130 にロードされても良い。例えば、開発者や、アプリケーションのインストールファイルを配布するファイル配布システムがネットワーク 170 を介して提供するファイルによって電子機器 100 にインストールされるコンピュータプログラム (一例として、上述したアプリケーション) がメモリ 130 にロードされて良い。

【0020】

通信モジュール 140 は、電子機器 100 をコンピュータネットワークに接続するためのコンピュータハードウェア構成要素であって良い。例えば、通信モジュール 140 は、電子機器 100 がネットワークを介してネットワーク上の他の電子機器と通信するための機能を提供して良い。ここで、コンピュータネットワークを利用した通信方式には、特に限定されることはなく、コンピュータネットワークが含むことができる通信網 (一例として、移動通信網、有線インターネット、無線インターネット、放送網) を活用する通信方式だけでなく、機器同士の近距離無線通信が含まれても良い。例えば、コンピュータネットワークは、PAN (personal area network) や、LAN (local area network)、CAN (campus area network)、MAN (metropolitan area network)、WAN (wide area network)、BBN (broadband network) など、インターネット等のネットワークのうち 1 つ以上の任意のネットワークを含んで良い。また、コンピュータネットワークは、バスネットワークや、スターネットワーク、リングネットワーク、メッシュネットワーク、スターバスネットワーク、ツリーまたは階層的 (hierarchical) ネットワーク等を含むネットワークトポロジのうち任意の 1 つ以上を含んで良いが、これらに限定されることはない。

【0021】

入力／出力インタフェース 150 は、入力／出力装置 160 とのインタフェースのための手段であって良い。例えば、入力装置は、キーボードや、マウス、マイクロフォン (microphone、以下「マイク」とする) 等の装置を、出力装置は、ディスプレイやスピーカのような装置を含んで良い。他の例として、入力／出力インタフェース 150 は、タッチスクリーンのように入力と出力のための機能が 1 つに統合された装置とのインタフェースのための手段であっても良い。実施形態によって、入力／出力装置 160 は、電子

10

20

30

40

50

機器１００とは別の構成要素として、電子機器１００と通信するように構成されても良く、電子機器１００に含まれる１つの装置として構成されても良い。例えば、パーソナルコンピュータのように、マイクやスピーカが別の装置としてパーソナルコンピュータの本体に接続される実施形態が良く、スマートフォンのように、スマートフォン本体にマイクやスピーカが含まれる実施形態でも良い。

【００２２】

電子機器１００のプロセッサ１１０は、メモリ１３０にロードされたコンピュータプログラムの命令を処理するにあたり、キーボードやマウス、マイク、タッチスクリーン等の入力装置より電子機器１００に入力される各種の信号や情報を処理し、サービス画面やコンテンツ、オーディオ信号等の各種の信号や情報については、入力／出力インタフェース１５０を介してディスプレイやスピーカ等の出力装置より出力されるように電子機器１００を制御して良い。

10

【００２３】

また、他の実施形態において、電子機器１００は、図１の構成要素よりも多くの構成要素を含んでも良い。ところが、殆どの従来技術の構成要素を図に示す必要はない。例えば、電子機器１００は、上述した入力／出力装置１６０のうち少なくとも一部を含むように実現されても良く、トランシーバや、GPS (Global Positioning System) モジュール、カメラ、各種のセンサ、データベース等の他の構成要素を更にも含んでも良い。より具体的な例として、電子機器１００がスマートフォンである場合、一般的にスマートフォンが含んでいる加速度センサやジャイロセンサ、カメラ、物理的な各種ボタン、タッチパネルを利用したボタン、入力／出力ポート、振動のための振動器等の多様な構成要素が電子機器１００に更に含まれるように実現されても良い。

20

【００２４】

このような電子機器１００にインストールされたコンピュータプログラムは、電子機器１００においてソフトウェアの音質改善機能の可否を判断し、ソフトウェアの音質改善機能を選択的に活性化することができる。ここで、音質改善機能は、AEC (Acoustic Echo Cancellation) モジュールや、NS (Noise Suppression: NS) モジュール、AGC (Automatic Gain Control) モジュール等から構成されて良い。このような音質改善機能についての具体的な説明は、以下で更に詳しく説明する。

30

【００２５】

図２は、本発明の一実施形態における電子機器のプロセッサが含む構成要素の例を示したブロック図であり、図３は、本発明の一実施形態における電子機器が実行する音質改善方法の例を示したフローチャートである。上述したように、電子機器１００には、本実施形態に係る音質改善システムが実現されて良く、このために、電子機器１００のプロセッサ１１０は、図２に示すように、マイク信号処理部２１０、スピーカ信号処理部２２０、決定部２３０、および活性化部２４０を含んで良い。

【００２６】

ここで、プロセッサ１１０の構成要素は、電子機器１００にインストールされて駆動されるコンピュータプログラムのコード（又は、ブラウザやオペレーティングシステム）が提供する命令 (instruction) にしたがってプロセッサ１１０によって実行されるプロセッサ１１０の相異なる機能 (different functions) の表現であって良い。例えば、プロセッサ１１０がマイク信号を処理するために電子機器１００を制御するプロセッサ１１０の機能を表現したものとして、マイク信号処理部２１０が用いられて良い。

40

【００２７】

このようなプロセッサ１１０およびプロセッサ１１０の構成要素は、メモリ１３０の含むオペレーティングシステムのコードや少なくとも１つのプログラムのコードによる命令を実行するように実現されて良い。特に、プロセッサ１１０およびプロセッサ１１０の構成要素は、図３の音質改善方法が含む段階３１０～３６０を実行するように電子機器１００

50

を制御することができる。

【0028】

段階310において、マイク信号処理部210は、電子機器100のマイクから入力されるマイク入力信号を処理するように、電子機器100を制御することができる。ここで、マイクは、電子機器100に含まれる構成要素であっても良く、電子機器100とはネットワーク（一例として、USB（Universal Serial Bus）やブルートゥース（登録商標）等）を介して接続される別の装置であっても良い。

【0029】

段階320において、スピーカ信号処理部220は、電子機器100のスピーカから出力されるスピーカ出力信号を処理するように、電子機器100を制御することができる。スピーカについても、マイクと同様に、電子機器100に含まれる構成要素であっても良く、電子機器100とネットワークを介して接続される別の装置であっても良い。

【0030】

段階330において、決定部230は、電子機器100へのマイク入力信号および電子機器100からのスピーカ出力信号を分析して、ソフトウェアの音質改善機能の要否を決定することができる。ソフトウェアの音質改善機能の要否を決定するためのより具体的な方法については、以下で図4を参照しつつ詳しく説明する。

【0031】

段階340において、決定部230は、段階330で決定されたソフトウェアの音質改善機能の要否に基づき、段階350または段階360が選択的に実行されるようにして良い。例えば、決定部230は、ソフトウェアの音質改善機能が必要であると決定された場合、活性化部240へ、ソフトウェアの音質改善機能を活性化するための命令を伝達することにより、に段階350が実行される。一方、決定部230は、ソフトウェアの音質改善機能が必要ないと決定された場合には、活性化部240へ、ソフトウェアの音質改善機能を非活性化するための命令を伝達することにより、段階360が実行される。

【0032】

段階350において、活性化部240は、ソフトウェアの音質改善機能を活性化することができる。一例として、ソフトウェアの音質改善機能は、上述したように、AEC（Acoustic Echo Cancellat ion）モジュールや、NS（Noise Suppression：NS）モジュール、AGC（Automatic Gain Control）モジュールを含み、それぞれのモジュールはソフトウェア的に実現されて良い。この際、段階330では、AECモジュールや、NSモジュール、AGCモジュールそれぞれの要否が決定され、活性化部240は、段階350において決定された必要性に対応するモジュールを選択的に活性化することができる。

【0033】

段階360において、活性化部240は、ソフトウェアの音質改善機能を非活性化することができる。例えば、ソフトウェアの音質改善機能が既に活性化されていて、段階330でソフトウェアの音質改善機能が必要ないと決定された場合、活性化部240は、段階360でソフトウェアの音質改善機能を非活性化することができる。上述したように、AECモジュールや、NSモジュール、AGCモジュールのそれぞれに対する活性化が可能であると同様に、AECモジュールや、NSモジュール、およびAGCモジュールのそれぞれに対する非活性化も可能である。

【0034】

このようなソフトウェアの音質改善機能の要否に関する決定と、ソフトウェアの音質改善機能の活性化または非活性化は、電子機器100にインストールされたアプリケーションの目的に応じて、音質改善が求められる間に繰り返されて良い。例えば、段階330～360は、終了命令が入力されるまで繰り返して実行されても良い。

【0035】

図4は、本発明の一実施形態において、ソフトウェアの音質改善機能の要否を決定するための方法の例を示したフローチャートである。本実施形態において、図4の段階410～

10

20

30

40

50

470は、図3を参照しつつ説明した段階330に含まれて良い。以下では、マイク入力信号を「Y」、スピーカ出力信号を「X」と称する。

【0036】

段階410において、決定部230は、エコー（echo）区間を決定することができる。

【0037】

決定部230は、Xの活性化区間において、YとXの相互相関度分析によってエコー区間を決定することができる。Xの活性化区間を決定（Voice Activity Detection：VAD）するための方法は種々ある。例えば、Xの平均エネルギーよりも相対的に高い平均エネルギーをもつ区間を、Xの活性化区間として決定して良い。X、即ち、スピーカ出力信号の活性化区間を決定するために、公知の多様な活性化区間の決定方法のうちの1つが活用されて良い。

10

【0038】

決定部230は、リアルタイム処理のために、XとYを既定の大きさのフレーム単位であるTの単位に分割して良い。この際、分割されたXのエネルギー E_x は、下記の数式（1）にて計算されて良い。

【0039】

【数1】

$$E_{x_f} = \sum_{n=f*T}^{(f+1)*T} x^2(n) \quad \dots (1)$$

20

【0040】

ここで、fは、分割されたフレームのインデックスを、Tは、既定の大きさのフレーム処理単位を、それぞれ意味して良い。例えば、処理単位Tとして10msecを設定し、サンプリングレートが16,000Hzであると仮定すると、Xは160サンプルフレームに分割される。

30

【0041】

ここで、Xの平均エネルギー

【0042】

【数2】

$$\overline{E_{x_f}}$$

40

は、以下の数式（2）にて計算されて良い。

【0043】

【数3】

50

$$\overline{Ex_f} = 0.99\overline{Ex_{f-1}} + 0.01Ex_f \dots (2)$$

$$Ex_f \text{ が } \overline{Ex_f}$$

よりも大きければ、 X_f は活性化区間であると言える。

【0044】

Xの活性化区間における相関度分析は、下記の数式（3）のように、YとXの相互相関関数にて得られて良い。

【0045】

【数4】

$$\sum_{n=0}^T y(n)x(n-d) \dots (3)$$

ここで、 d は、遅延を意味して良い。

【0046】

遅延 d は、負数または正数であって良く、 d の範囲は、音響エコー遅延（acoustic echo delay）やシステム遅延（system delay）を含んで良い。音響エコー遅延は、信号がスピーカ出力からマイクへ流入するまでの音響環境で発生する遅延を含んで良い。また、システム遅延は、装置バッファ遅延（device buffer delay）のように、マイクに流入した信号が相関度分析端に伝達されるまでのハードウェア及びソフトウェアにおける全ての遅延を含んで良い。即ち、決定部230で信号を受信するまでの全ての遅延が、システム遅延に含まれて良い。

【0047】

このような数式（3）を正規化すると、以下の数式（4）のように表現される。

【0048】

【数5】

$$R(d) = \frac{\frac{1}{T} \sum_{n=0}^T y(n)x(n-d)}{\frac{1}{T} y^2(n)} \dots (4)$$

$R(d)$ は、XとYが類似するほど、大きい値を有する。

【0049】

ここで、XとYの2つの信号の相互相関度を分析した結果、 $R(d)$ のうち最大値を有するインデックスの d が、以下の数式（5）のように、XとY2つの信号の間の遅延（D）を意味する。

10

20

30

40

50

【0050】

$D = \arg \max (R(d)) \dots (5)$

【0051】

信号Yでのエコー区間は、R(D)によって決定されて良い。DがXの活性化区間の間に同様の値で保たれるのであれば、エコー区間は、x(n)が活性化区間であり、Dが同様の値で保たれる場合、以下の数式(6)のように定義されて良い。

【0052】

$y(n+D) \dots (6)$

【0053】

x(n)が非活性化区間であるか、または、Dが一定の値で保たれない場合には、数式(6)のy(n+D)は、非エコー区間を示す。

10

【0054】

段階420において、決定部230は、ユーザ入力区間を決定することができる。

【0055】

まず、決定部230は、マイク入力信号およびスピーカ出力信号の全体区間からエコー区間を除いた、Yに対する活性化区間を決定することができる。

【0056】

例えば、決定部230は、Yに対するエネルギー

【0057】

【数6】

20

$$Ey_f \propto \alpha * \overline{Ey_f}$$

よりも大きい場合、Y_fが活性化区間であると判断し、このような活性化区間をユーザ入力区間として決定することができる。ここで、Yに対する平均エネルギー

【0058】

【数7】

30

$$\overline{Ey_f}$$

は、Y_fがユーザ入力区間である場合には以下の数式(7)のように、その他の場合には以下の数式(8)のように、それぞれ計算されて良い。例えば、αは、第1加重値として2.0のようにその値が予め設定されて良いが、これに限定されるものではない。上述したように、活性化区間の決定方法が本例示に限定されるものではない。

【0059】

【数8】

40

$$\overline{Ey_f} = 0.99\overline{Ey_{f-1}} + 0.01Ey_f \dots (7)$$

$$\overline{Ey_f} = \overline{Ey_{f-1}} \dots (8)$$

50

【 0 0 6 0 】

段階 4 3 0 において、決定部 2 3 0 は、雑音区間を決定することができる。

【 0 0 6 1 】

例えば、決定部 2 3 0 は、マイク入力信号およびスピーカ出力信号の全体区間からエコー区間とユーザ入力区間を除いた残りの区間のうち、 Y_f のエネルギー

【 0 0 6 2 】

【数 9】

$$E_{y_f} \beta * \overline{E_{y_f}}$$

10

よりも小さい場合、 Y_f を雑音区間として決定することができる。例えば、 β は、第 2 加重値として 1 のようにその値が予め設定されて良い。ただし、雑音区間の決定方法や用いられた係数は、単に例示に過ぎず、これに限定されるものではない。

【 0 0 6 3 】

段階 4 4 0 において、決定部 2 3 0 は、エコー区間、ユーザ入力区間、及び雑音区間のそれぞれにおける平均エネルギーを測定することができる。

【 0 0 6 4 】

一例として、エコー区間における平均エネルギー

20

【 0 0 6 5 】

【数 1 0】

$$\overline{Echo_f}$$

は、 Y_f がエコー区間である場合は以下の数式 (9) のように計算され、その他の場合は以下の数式 (10) のように計算されて良い。

【 0 0 6 6 】

30

【数 1 1】

$$\overline{Echo_f} = 0.99 \overline{Echo_{f-1}} + 0.01 E_{y_f} \dots (9)$$

$$\overline{Echo_f} = \overline{Echo_{f-1}} \dots (10)$$

【 0 0 6 7 】

40

また、ユーザ入力区間における平均エネルギー

【 0 0 6 8 】

【数 1 2】

$$\overline{E_{user_f}}$$

は、 Y_f がユーザ入力区間である場合は以下の数式 (11) のように計算され、その他の場合は以下の数式 (12) のように計算されて良い。

50

【 0 0 6 9 】

【 数 1 3 】

$$\overline{Euser}_f = 0.99\overline{Euser}_{f-1} + 0.01Ey_f \dots (11)$$

$$\overline{Euser}_f = \overline{Euser}_{f-1} \dots (12)$$

10

【 0 0 7 0 】

また、雑音区間における平均エネルギー

【 0 0 7 1 】

【 数 1 4 】

$$\overline{Enoise}_f$$

20

は、 Y_f が雑音区間である場合は以下の数式 (13) のように計算され、その他の場合は以下の数式 (14) のように計算されて良い。

【 0 0 7 2 】

【 数 1 5 】

$$\overline{Enoise}_f = 0.99\overline{Enoise}_{f-1} + 0.01Ey_f \dots (13)$$

$$\overline{Enoise}_f = \overline{Enoise}_{f-1} \dots (14)$$

30

【 0 0 7 3 】

数式 (9)、数式 (11)、および数式 (13) において平均エネルギーを求めるために用いられた係数 0.99 と 0.01 は一例に過ぎず、これに限定されるものではない。

【 0 0 7 4 】

ここで、エネルギーは、人間が感じる音の大きさの単位であるデシベル ($dB = 10 \log(E/T)$) で表現されて良い。

【 0 0 7 5 】

40

段階 450 において、決定部 230 は、遅延 D とエコー区間における平均エネルギーのうち少なくとも 1 つを利用して、AEC モジュールの要否を決定することができる。例えば、決定部 230 は、段階 410 で決定された X と Y 2 つの信号の間の遅延 D により、AEC モジュールの要否を決定することができる。例えば、遅延 D がエコー区間の間に、連続する k 個のフレーム (k は 2 以上の自然数であり、一例として、k は 2) において同様の値で保たれなければ、2 つの信号の相関度が低いと決定することができる。これで、ハードウェア AEC が提供されているか、それとも、エコーが入ってこない環境であると判断することができる。実施形態によっては、k は 3 以上の値を有することもある。この場合、決定部 230 は、AEC モジュールが必要でないと判断することができる。また、決定部 230 は、段階 440 で決定されたエコー区間における平均エネルギー

50

【 0 0 7 6 】

【 数 1 6 】

 $\overline{Echo}_f$

が既定の第1デシベル値（一例として、30 dB、第1閾値）よりも小さい場合、ハードウェアAECが提供されているか、それとも、エコーが入ってこない環境であると判断することができる。したがって、この場合にも、決定部230は、AECモジュールが必要でないとは判断することができる。このように、AECモジュールが必要でないとは判断されれば、決定部230は、AECモジュールの非活性化を行うための信号を生成して、活性化部240に送信することができる。この場合、活性化部240は、段階360において、AECモジュールを非活性化（AECモジュールが活性化の状態である場合）することができる。

10

【 0 0 7 7 】

一方、決定部230は、遅延Dがk個の連続するフレームにおいて同様の値を有し、かつ、平均エネルギー

【 0 0 7 8 】

【 数 1 7 】

20

 $\overline{Echo}_f$

が既定の第1デシベル値以上である場合は、AECモジュールが必要な状況であると判断することができる。この場合、決定部230は、AECモジュールを活性化するための信号を生成し、活性化部240は、生成された信号の伝達を受け、段階350において、AECモジュールの活性化（AECモジュールが非活性化の状態である場合）を行うことができる。

30

【 0 0 7 9 】

即ち、AECモジュールの要否を決定する際に、マイク入力信号とスピーカ出力信号の間の遅延の連続性の有無や、エコー区間における平均エネルギーと既定の第1閾値との比較結果が利用される。

【 0 0 8 0 】

段階460において、決定部230は、雑音区間における平均エネルギーを利用して、NSモジュールの要否を決定することができる。

【 0 0 8 1 】

例えば、決定部230は、段階440で決定された雑音区間における平均エネルギー

【 0 0 8 2 】

【 数 1 8 】

40

 $\overline{Enoise}_f$

が既定の第2デシベル値（一例として、20 dB、第2閾値）よりも小さい場合、ハードウェアNSが提供されているか、それとも、雑音が入ってこない環境であると判断することができる。この場合、決定部230は、NSモジュールが必要でないとは判断することができる。このように、NSモジュールが必要でないとは判断されれば、決定部230は、N

50

S モジュールの非活性化を行うための信号を生成して、活性化部 240 に送信することができる。この場合、活性化部 240 は、段階 360 において、NS モジュールの非活性化（NS モジュールが活性化の状態である場合）を行うことができる。

【0083】

【数19】

$$\overline{Enoise_f}$$

10

が既定の第2デシベル値以上である場合、決定部230は、NS モジュールが必要な状況であると判断することができる。この場合、決定部230は、NS モジュールを活性化するための信号を生成し、活性化部240は、生成された信号にしたがい、段階350において、NS モジュールを活性化（NS モジュールが非活性化の状態である場合）することができる。

【0084】

段階470において、決定部230は、ユーザ入力区間における平均エネルギーを利用して、AGC モジュールの要否を決定することができる。

【0085】

例えば、決定部230は、段階440で決定されたユーザ入力区間における平均エネルギー

20

【0086】

【数20】

$$\overline{Euser_f}$$

が既定のデシベル範囲（一例として、50 dBと60 dB）内の値である場合、ハードウェアAGCが提供されているか、それとも、適切な音量のユーザ入力が入ってくる状況であると判断することができる。この場合、決定部230は、AGC モジュールが必要でないと判断することができる。このように、AGC モジュールが必要でないと判断されれば、決定部230は、AGC モジュールを非活性化するための信号を生成して、活性化部240に送信することができる。この場合、活性化部240は、段階360において、AGC モジュールを非活性化（AGC モジュールが活性化の状態である場合）することができる。

30

【0087】

【数21】

$$\overline{Euser_f}$$

40

が上述したデシベル範囲の外の値である場合、決定部230は、AGC モジュールが必要な状況であると判断することができる。ここで、決定部230は、AGC モジュールを活性化するための信号を生成し、活性化部240は、生成された信号にしたがい、段階350において、AGC モジュールを活性化（AGC モジュールが非活性化の状態である場合）することができる。

【0088】

上述したk、第1デシベル値、第2デシベル値、およびデシベル範囲は、経験的に決定されても良く、電子機器100にインストールされたアプリケーションの目的に応じて決定

50

されても良い。A E C モジュールは、エコーの線形的特性を推定し、推定された線形特性のエコーを取り除くためのモジュールであって良く、N S モジュールは、雑音レベルを推定し、推定された雑音を取り除くためのモジュールであって良い。また、A G C モジュールは、利得 (g a i n) を調節するためのモジュールであって良い。このような A E C モジュール、N S モジュール、および A G C モジュールは、上述したアプリケーションにおいてソフトウェア的に実現されて含まれて良い。

【 0 0 8 9 】

図 5 は、本発明の一実施形態において、ソフトウェアの音質改善機能を活性化する過程の例を示した図である。電子機器 1 0 0 は、上述した入力／出力装置 1 6 0 としてスピーカ 5 1 0 とマイク 5 2 0 を含んでも良く、スピーカ 5 1 0 およびマイク 5 2 0 と接続しても良い。出力信号にしたがって、スピーカ 5 1 0 から音出力されることができる。この際、マイク 5 2 0 には、ユーザの音声のようなニアエンドスピーチ (n e a r - e n d s p e e c h) だけでなく、スピーカ 5 1 0 から出力された音に対するエコー (e c h o) や雑音 (n o i s e) が更に入力されることがある。

【 0 0 9 0 】

電子機器 1 0 0 にインストールされたコンピュータプログラムは、スピーカ 5 1 0 の出力信号 X とマイク 5 2 0 の入力信号 Y の入力を受けて分析を行い、分析結果に基づいてソフトウェアの音質改善機能の要否をリアルタイムで決定することができる。

【 0 0 9 1 】

相関度分析モジュール 5 3 0 と活性化区間決定モジュール 5 4 0 は、決定部 2 3 0 が上述した図 3 の段階 3 3 0 と図 4 の段階 4 1 0 ~ 4 7 0 を実行できるようにするための命令を含む、コンピュータプログラムのコードによって実現されて良い。

【 0 0 9 2 】

相関度分析モジュール 5 3 0 の制御にしたがい、決定部 2 3 0 は、出力信号 X と入力信号 Y の相関度を分析し、エコー区間を決定して良い。また、活性化区間決定モジュール 5 4 0 の制御にしたがい、決定部 2 3 0 は、入力信号 Y を利用し、雑音区間とユーザ入力区間を決定して良い。

【 0 0 9 3 】

決定部 2 3 0 は、コンピュータプログラムの制御にしたがって生成された区間情報 5 5 0 を利用して、区間情報 5 5 0 における平均エネルギーを算出し、算出された平均エネルギーに基づいて上述した A E C モジュール、N S モジュール、および A G C モジュールのうち少なくとも 1 つに対する活性化の要否を決定することができる。活性化部 2 4 0 は、活性化が決定されたモジュールを活性化することにより、ハードウェアの音質改善機能と重複しないように、または、ハードウェアの音質改善機能が実行されている場合でも、より良い音質を得るために、音質改善機能をリアルタイムで活性化または非活性化することができる。

【 0 0 9 4 】

このように、本発明の実施形態によると、マイク入力信号とスピーカ出力信号を分析してリアルタイムでソフトウェアの音質改善機能の要否を判断し、ソフトウェアの音質改善機能を選択的に活性化または非活性化することができる。

【 0 0 9 5 】

上述したシステムまたは装置は、ハードウェア構成要素、ソフトウェア構成要素、又は、ハードウェア構成要素とソフトウェア構成要素の組み合わせにより、実現されて良い。例えば、実施形態で説明された装置および構成要素は、例えば、プロセッサ、コントローラ、A L U (a r i t h m e t i c l o g i c u n i t) 、デジタル信号プロセッサ、マイクロコンピュータ、F P G A (f i e l d p r o g r a m m a b l e g a t e a r r a y) 、P L U (p r o g r a m m a b l e l o g i c u n i t) 、マイクロプロセッサ、又は、命令を実行して応答することができる他の装置のように、1 つ以上の汎用コンピュータ又は特殊な目的のコンピュータを利用して実現されて良い。処理装置は、オペレーティングシステム (O S) や前記 O S 上で実行される 1 つ以上のソフトウェアアプリ

ケーションを実行して良い。また、処理装置は、ソフトウェアの実行に応じて、データにアクセスし、データを格納、操作、処理、及び生成しても良い。理解の便宜のために、1つの処理装置が使用されるとして説明された場合もあるが、本発明の属する技術の分野における通常の知識を有する者にとって、処理装置が複数の処理要素 (Processing Element) および／または複数の種類の処理要素を含み得ることは理解できるであろう。例えば、処理装置は、複数のプロセッサ又は1つのプロセッサと、1つのコントローラを含んで良い。また、並列プロセッサのように、他の処理構成も可能である。

【0096】

ソフトウェアは、コンピュータプログラム、コード、命令、またはこれらのうち1つ以上の組み合わせを含んで良く、思うままに動作するように処理装置を構成したり、独立的または集合的に処理装置に命令したりしてよい。ソフトウェアおよび／またはデータは、処理装置に基づいて解釈されたり、処理装置に命令またはデータを提供したりするために、いかなる種類の機械、コンポーネント、物理装置、仮想装置、コンピュータ格納媒体または装置、または伝送される信号波に永久的または一時的に具現化されてもよい。ソフトウェアは、ネットワークによって接続されたコンピュータシステム上に分散され、分散された状態で格納されても実行されてもよい。ソフトウェアおよびデータは、1つ以上のコンピュータで読み取り可能な記録媒体に格納されてもよい。

【0097】

実施形態に係る音質改善方法は、多様なコンピュータ手段によって実行可能なプログラム命令や、データファイル、データ構造等を単独でまたは組み合わせた形態として実現されて良く、また、該プログラム命令等がコンピュータで読取可能な媒体に記録されても良い。前記プログラム命令等は、実施形態のために特別に設計され構成されたものであっても良く、コンピュータソフトウェアの当業者に公知で使用可能なものであっても良い。コンピュータで読取可能な記録媒体の例には、ハードディスクや、フロッピー（登録商標）ディスク、磁気テープ等の磁気媒体、CD-ROMやDVD等の光記録媒体、フロッピカルディスク (floptical disk) 等の光磁気媒体、ROMや、RAM、フラッシュメモリ等のように、プログラム命令等を格納して実行するように構成されたハードウェア装置が含まれる。プログラム命令等の例には、コンパイラによって生成されるもののような機械語コードだけではなく、インタプリタ等を使用してコンピュータによって実行される高級言語コードが含まれる。上述したハードウェア装置は、実施形態での動作を実行するため、1つ以上のソフトウェアモジュールとして動作するように構成されても良く、その逆に構成されても良い。

【0098】

以上のように、実施形態を、限定された実施例や図面に基づいて説明したが、本発明の属する技術の分野における通常の知識を有する者であれば、上述した記載内容から多様な修正や変更が可能であろう。例えば、説明された技術が、説明された方法とは異なる手順で実行されたり、及び／又は、説明されたシステムや、構造、装置、回路等の構成要素が、説明された方法とは異なる形態で結合又は組合されたり、他の構成要素や均等物に代替又は置換されたりとしても、適切な結果を達成することができる。

【0099】

したがって、異なる実施形態であっても、特許請求の範囲と均等なものであれば、添付される特許請求の範囲に属する。

【符号の説明】

【0100】

100：電子機器

110：プロセッサ

120：バス

130：メモリ

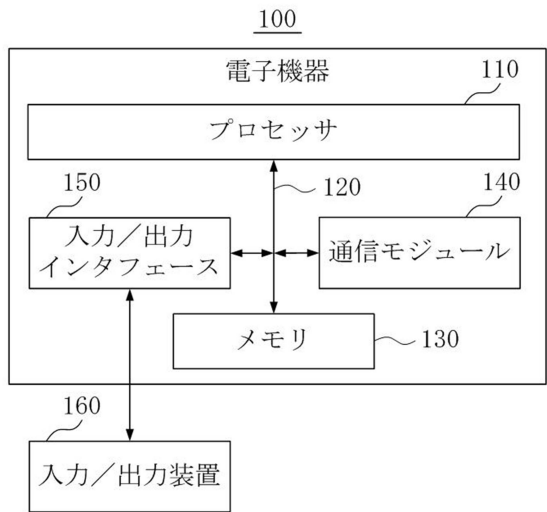
140：通信モジュール

150：入力／出力インタフェース

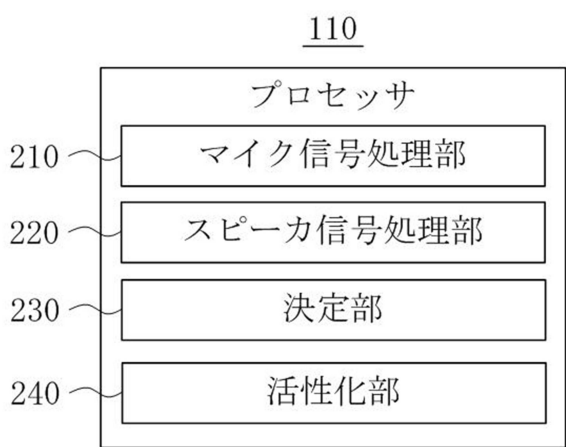
160：入力／出力装置

【図面】

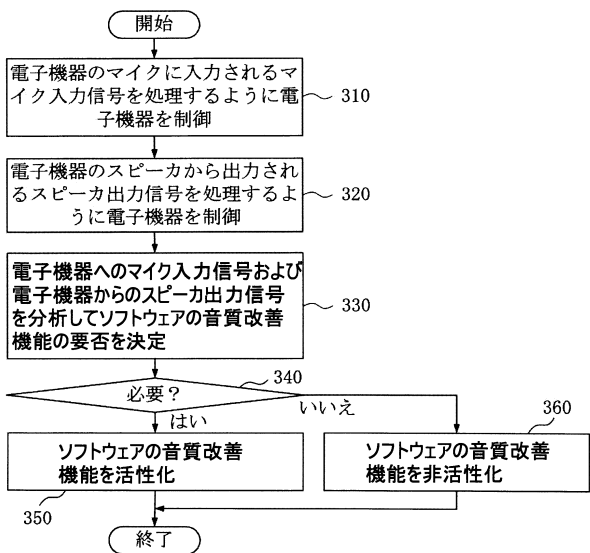
【図1】



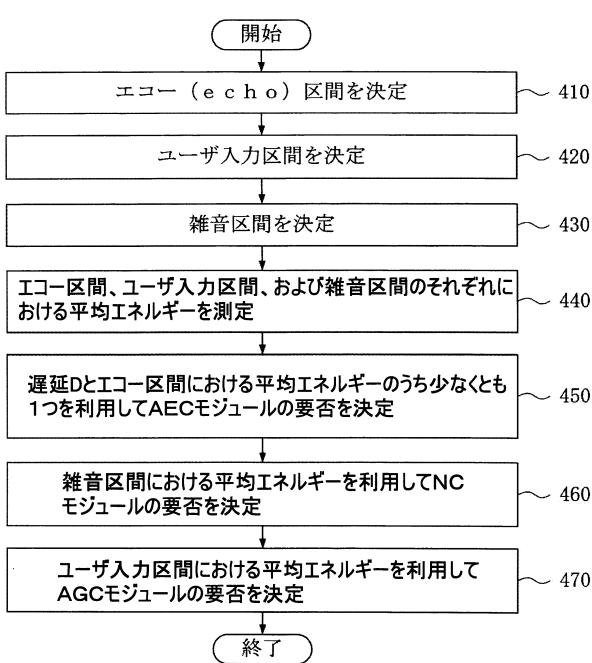
【図2】



【図3】



【図4】



10

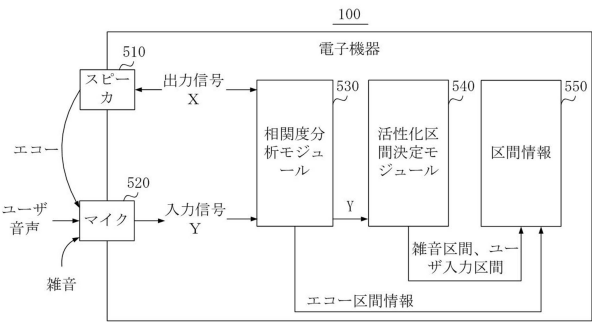
20

30

40

50

【図 5】



10

20

30

40

50

フロントページの続き

- (56)参考文献 特開 2008-211526 (JP, A)
特開 2003-051879 (JP, A)
特表 2016-502779 (JP, A)
米国特許第 06785339 (US, B1)
特表 2009-518663 (JP, A)

- (58)調査した分野 (Int.Cl., DB名)
H04R 3/00-3/04
H04M 1/00-1/23
H04M 1/24-1/82
H04B 1/76-3/60