

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6349899号
(P6349899)

(45) 発行日 平成30年7月4日(2018.7.4)

(24) 登録日 平成30年6月15日(2018.6.15)

| | | | | | | |
|-------------------------|-----------|-----------|------|--|--|--|
| (51) Int.Cl. | F 1 | | | | | |
| HO4R 3/02 | (2006.01) | HO4R 3/02 | | | | |
| HO4R 3/00 | (2006.01) | HO4R 3/00 | 310 | | | |
| HO4R 1/40 | (2006.01) | HO4R 3/00 | 320 | | | |
| HO4M 3/56 | (2006.01) | HO4R 1/40 | 320A | | | |
| HO4M 1/60 | (2006.01) | HO4M 3/56 | B | | | |
| 請求項の数 4 (全 15 頁) 最終頁に続く | | | | | | |

(21) 出願番号 特願2014-83209 (P2014-83209)
 (22) 出願日 平成26年4月14日(2014.4.14)
 (65) 公開番号 特開2015-204535 (P2015-204535A)
 (43) 公開日 平成27年11月16日(2015.11.16)
 審査請求日 平成29年2月20日(2017.2.20)

(73) 特許権者 000004075
 ヤマハ株式会社
 静岡県浜松市中区中沢町10番1号
 (74) 代理人 100123940
 弁理士 村上 辰一
 (72) 発明者 鶴飼 訓史
 静岡県浜松市中区中沢町10番1号 ヤマ
 ハ株式会社内
 (72) 発明者 山川 高史
 静岡県浜松市中区中沢町10番1号 ヤマ
 ハ株式会社内
 審査官 下林 義明

最終頁に続く

(54) 【発明の名称】 放收音装置

(57) 【特許請求の範囲】

【請求項1】

スピーカと、

前記スピーカに供給される音声信号である放音信号を処理するフィルタと、

指向性を有する複数のマイクと、

各マイクに対応して設けられた複数のエコーキャンセラであって、各々がその対応するマイクの收音信号から前記スピーカによって放音された音声の回帰音信号をキャンセルするものと、

前記エコーキャンセラから取り出された適応フィルタ係数列を統合する第1統合部と、

統合された前記係数列に基づき、前記スピーカおよびマイクが存在する空間の周波数帯域毎の残響時間を推定する残響時間推定部と、

残響時間の長い周波数帯域を抽出し、該周波数帯域のパワーを抑制するフィルタ係数を算出して前記フィルタに設定する演算部と、

を備えた放收音装置。

【請求項2】

前記複数のマイクは、水平の全方向にいずれかのマイクが感度を持つようそれぞれ異なる方向に向けて配置されている請求項1に記載の放收音装置。

【請求項3】

前記複数のマイクおよび前記複数のエコーキャンセラを含む收音器、および、前記スピーカおよび前記フィルタを含む通信器を、それぞれ別体に有する請求項1または請求項2

に記載の放收音装置。

【請求項 4】

前記收音器を複数備え、

前記第 1 統合部は、前記複数の收音器に対応して複数設けられ、

前記残響時間推定部は、各收音器の残響時間を統合する第 2 統合部をさらに有し、

前記演算部は、前記第 2 統合部によって統合された残響時間に基づいて、前記フィルタ係数を算出する

請求項 3 に記載の放收音装置。

【発明の詳細な説明】

【技術分野】

10

【0001】

この発明は、遠隔の音声会議等に用いられる放收音装置に関し、特に放音される音声の残響の抑制に関する。

【背景技術】

【0002】

拠点間をネットワークで接続して音声を送受信する音声会議システムが実用化されている。会議に使用される会議室の音響特性は様々であり、場合によっては非常に長い残響を有する部屋で会議が行われる場合もある。残響が長いとスピーカから放音される音声の明瞭度が低下する。これに対処するため、放音される音声の残響を抑制する装置が提案されている（特許文献 1）。

20

【0003】

特許文献 1 の装置は、参加者 M_a に対応するキー K_{ia} を操作することにより、参加者 M_a からマイクロホン 31 までの空間伝達関数 H_a の逆フィルタ係数 G_a が、ROM 41 から読み出されてデジタルフィルタ 34i に供給され、このデジタルフィルタ 34i が逆フィルタ演算をリアルタイムで行って参加者 M_a の音声信号を逆フィルタリング処理するものである。すなわち、この装置では、各参加者 $M_a \sim M_n$ の席から複数のマイクロホン 31 までの空間伝達関数が予め測定されており、各伝達関数の逆フィルタ係数 $G_a \sim G_n$ が ROM 41 に予め格納されている。

【先行技術文献】

【特許文献】

30

【0004】

【特許文献 1】特開平 09 - 247788 号公報

【発明の開示】

【発明が解決しようとする課題】

【0005】

しかしながら、上記従来の装置では、予めインパルス応答を測定する必要があるため、会議開始直前にテスト信号を再生するか、会議の途中でテスト信号を再生する必要があり、テスト信号の再生中は静かにしている必要があるため、会議の円滑な進行の妨げになった。

【0006】

40

会議参加者が入場する前に行っておくことも考えられるが、参加者が居るときと不在のときとでは会議室の音響特性が変化するため、やはり参加者が揃った時点であるのが好ましい。

【0007】

この発明は、事前にテスト音声を再生しなくても、会議用の收音およびエコーキャンセル機能を利用して残響を抑制できる放收音装置を提供することを目的とする。

【課題を解決するための手段】

【0008】

本発明の放收音装置は、スピーカと、スピーカに供給される音声信号である放音信号を処理するフィルタと、指向性を有する複数のマイクと、複数のエコーキャンセラと、第 1

50

統合部と、残響特性推定部と、演算部と、を備える。複数のエコーキャンセラは、複数のマイクの各々に対応して設けられ、各々がその対応するマイクの收音信号からスピーカによって放音された音声の回帰音信号をキャンセルする。第1統合部は、エコーキャンセラから取り出された適応フィルタ係数列を統合する。残響時間推定部は、統合されたフィルタ係数列に基づき、スピーカおよびマイクが存在する空間の周波数帯域ごとの残響時間を推定する。演算部は、残響時間の長い周波数帯域を抽出し、その周波数帯域のパワーを抑制するフィルタ係数を算出してフィルタに設定する。

【0009】

この発明の放收音装置は、会議などに好適化された指向性マイクを複数備える。各マイクには、スピーカ音のエコーをキャンセルするためのエコーキャンセラが設けられている。エコーキャンセラは、模擬回帰音信号を生成する適応フィルタを備え、スピーカとマイクとの間のインパルス応答を模擬した適応フィルタ係数（推定インパルス応答）を持つ。また、この推定インパルス応答は、放音信号およびマイクの收音信号に基づいて常に更新されている。マイクは指向性マイクであるため、この推定インパルス応答はマイクの指向性の方向から到来する残響成分しか多く含まず、会議室全体の残響特性を完全に現しているとはいえない。しかし、第1統合部が、複数の指向性マイクのパラメータを統合することにより、会議用の指向性マイクでありながら、広い範囲の方向から到来する残響成分を含むインパルス応答を模擬することができる。そして、この統合されたパラメータ（推定インパルス応答）を用いて残響時間を算出し、この残響を抑制するようなフィルタ係数を算出する。これにより、会議室全体の残響特性を正確に再現でき、効果的な残響の抑制ができる。また、推定インパルス応答は、エコーキャンセラが回帰音除去に使用するものがそのまま流用できるので、特別な計算量が必要にならず、また、このためにテスト音声を放音する必要も無い。

【0010】

上記発明において、複数のマイクを、水平の全方向にいずれかのマイクが感度を持つようそれぞれ異なる方向に向けて配置してもよい。マイクをこのように配置することにより、第1統合部で統合されたパラメータをほぼ無指向性のパラメータとすることが可能になる。

【0011】

上記発明において、複数のマイクおよび複数のエコーキャンセラを含む收音器、およびスピーカおよびフィルタを含む通信器を、それぞれ別体としてもよい。これにより、設置の自由度が増すとともに、收音器を複数設けることも可能になる。

【0012】

上記発明において、收音器を複数設けるとともに、第1統合部を、複数の收音器に対応して複数設け、さらに、残響時間推定部に、各收音器の残響時間を統合する第2統合部をさらに設けてもよい。そして、演算部は、第2統合部によって統合された残響時間に基づいてフィルタ係数を算出するようにしてもよい。

【0013】

これにより、装置が設置される部屋の複数の場所に集音器を設置することができ、多人数が参加する会議でも漏れなく発話音声を收音することができる。さらに、部屋の各所の残響時間を算出して第2統合部でこれを統合することができるため、偏りのない平均的な残響時間を得ることが可能になる。

【発明の効果】

【0014】

この発明によれば、会議に好適な指向性マイクを用い、且つ、エコーキャンセラのパラメータ（適応フィルタのフィルタ係数など）を利用して、適切な残響の抑制が可能になる。

【図面の簡単な説明】

【0015】

【図1】この発明の実施形態である音声会議システムの設置形態の一例を示す図

【図 2】会議室における音声の反射の形態を説明する図

【図 3】音声会議システムの収音器のブロック図

【図 4】音声会議システムの収音器のマイクの指向性を示す図

【図 5】収音器のエコーキャンセラのブロック図

【図 6】音声会議システムの通信器のブロック図

【図 7】通信器のパラメータ推定部の機能ブロック図

【図 8】残響抑制のプロセスの流れを示す図

【図 9】パラメータ推定部において現れる信号波形を例示した図

【図 10】通信器の補正特性算出部のゲインテーブルの例を示す図

【図 11】ゲインテーブルの他の例を示す図

10

【図 12】パラメータ推定部に周波数特性補正機能を追加した実施形態を示す図

【図 13】音声会議システムの収音器の他の結線形態を示す図

【図 14】個別のマイクをグルーピングして収音器を構成する例を示す図

【発明を実施するための形態】

【0016】

図面を参照して本発明の実施形態である音声会議システムについて説明する。図 1 は、本発明の実施形態である音声会議システムの設置形態の一例を示す図である。

【0017】

音声会議システム 1 は、会議室 C の会議机 D 上に設置される。音声会議システム 1 は、1 台の通信器 10 および 1 または複数台の収音器 11 を有する（この実施形態では 4 台）。通信器 10 は、スピーカ 26 を有する。収音器 11 は、複数のマイク素子 31 を備える。通信器 10 および収音器 11 は通信ケーブル 12 で相互に接続され、デジタル通信を行う。収音器 11 は、マイク素子 31 で収音した音声信号およびエコーキャンセラ 32（図 3 参照）のフィルタ係数を通信器 11 に送信する。通信器 10 は、ホスト装置であるパーソナルコンピュータ 2 に接続される。パーソナルコンピュータ 2 は、インターネットなどのネットワーク 3 を介して、他の拠点に設置されている他の音声会議システムと通信し、この音声会議システム 1 の通信器 10 から入力された（マイク素子 31 が収音した）音声信号を他の音声会議システムに送信するとともに、他の音声会議システムから受信した音声信号を通信器 10 に入力する。通信器 10 は、他の音声会議システムから送られてきた音声信号をスピーカ 26 から放音する。

20

30

【0018】

図 2 は、会議室 C における音声の反射の形態を説明する図である。スピーカ 26 から放音された音声は、直接会議の参加者 M およびマイク素子 31 に到達するとともに、会議室 C の壁および天井で様々に反射して、参加者 M およびマイク素子 31 に到達する。

【0019】

スピーカ 26 から放音された音声、すなわち、他拠点に設置された他の音声会議システムで収音された音声、マイク素子 31 によって収音され、他の音声会議システムに送信されると、送信した音声に戻って再生されるいわゆるエコーになる。このエコーを防止するため、マイク素子 31 には、スピーカ 26 から放音された音声をキャンセルするエコーキャンセラ 32（図 3 参照）が接続される。また、スピーカ 26 から放音された音声、会議室 C の壁や天井で反射する残響によってこもった音声になり明瞭度が低下するのを改善するため、通信器 10 には、残響抑制用のフィルタ 24（図 6 参照）が設けられている。このフィルタ 24 のフィルタ係数は、エコーキャンセラ 32 の適応フィルタ 35（図 5 参照）のフィルタ係数を利用して算出される。

40

【0020】

エコーキャンセラ 32 および残響抑制用のフィルタ 24 の機能・動作については、図 3 以下の図面を参照して後述する。なお、以下に説明する、通信器 10 および収音器 11 に内蔵の機能部は、電子回路で構成されてもよく、コンピュータなどのプロセッサとプログラムの協働によって実現されてもよい。

【0021】

50

図3は、収音器11のブロック図である。図4は、収音器11の3つのマイク素子31のそれぞれの指向性を示す図である。図5は、収音器11のエコーキャンセラ32のブロック図である。

【0022】

収音器11は、3つのマイク素子31を備える。図1および図4に示すように、収音器11は円盤状の平面形状をしており、その円周上に120度ずつの間隔で3つのマイク素子31が外向き（法線方向）に放射状に設けられている。各マイク素子31は単一指向性マイクであり、マイク素子31が向いている方向を中心にカージオイド形状の収音特性を有する。各マイク素子31は120度の間隔で設けられ、その指向特性は図4に示すような配置になるため、各マイク素子31の収音信号を合成すればほぼ無指向性に近い特性の信号が得られる。なお、マイク素子31は、指向特性がカージオイドのものに限定されない。後方に若干の指向性を持つものであってもよく、双指向性のものであってもよい。

10

【0023】

図3において、各マイク素子31にはそれぞれエコーキャンセラ32が設けられている。エコーキャンセラ32の詳細構成は、図5で説明するが、マイク素子31で収音された音声信号のなかからスピーカ26から放音された音声をキャンセルする。エコーキャンセラ32によって、スピーカ26の回り込み音声をキャンセルされた音声信号が、音声選択部33に入力される。音声選択部33には、3つのマイク素子31で収音された音声信号がそれぞれ入力される。音声選択部33は、入力された3つの音声信号のレベル・持続時間等に基づいて、どのマイク素子31から入力されている音声信号が優勢か、すなわち話者の発話音声信号と推定されるかを推定し、発話音声信号と推定された1つの音声信号を選択する。すなわち、会議における音声信号の収音においては、指向性マイクの特性を活かして、3つのマイク素子31から最適な1つのマイク素子31を選択し、S/N比の良い発話音声を収音している。選択された音声信号は、通信インタフェース34を介して通信器10に送信される。通信器10に複数の収音器11が接続されている場合、通信器10（マイクミキサ22：図6参照）は、各収音器11から受信した音声信号から音声信号のレベル・持続時間・相関度を比較してさらに1つを選択し、または、複数の音声信号をミキシングし、この選択またはミキシングした音声信号を相手システムに送信する。

20

【0024】

次に、図5を参照して、エコーキャンセラ32の構成について説明する。図5は、エコーキャンセラ32のブロック図である。エコーキャンセラ32は、フィルタ係数設定部35Bおよび可変フィルタ35Bを有する適応フィルタ35を有し、さらに加算器37を有している。一般的に適応フィルタとは、所定の最適化アルゴリズムに従って自己の伝達関数（適応フィルタ係数列）を自動適応させるフィルタである。

30

【0025】

フィルタ係数設定部35Bは、会議室Cの音響伝達系（スピーカ26からマイク素子31に至る音響伝搬経路）の伝達関数を推定し、推定した伝達関数のフィルタとなるようなフィルタ係数を可変フィルタ35Aに設定する。

【0026】

可変フィルタ35Aには、スピーカ26から放音される音声信号（放音信号）が入力される。可変フィルタ35Aの伝達関数は、会議室Cの音響伝達系（スピーカ26からマイク素子31に至る音響伝搬経路）が模擬された伝達関数であるため、可変フィルタ35Aでフィルタリングされた放音信号は、スピーカ26から放音され会議室Cを伝搬してマイク素子31に収音された音声信号（回帰音信号）を模擬した音声信号（擬似回帰音信号）である。この擬似回帰音信号は加算器37に入力される。

40

【0027】

また、加算器37には、マイク素子31が収音した音声信号（収音信号）が、入力される。加算器37は、収音信号から模擬回帰音信号を減算して出力する。収音信号には、会議出席者Mの発話音声信号とともにスピーカ26から放音されて回り込んだ回帰音信号も含まれている。加算器37が、収音信号から模擬回帰音信号を減算することにより、収音

50

信号から回帰音を除去する、すなわちエコーをキャンセルすることができる。エコーをキャンセルされた収音信号は、音声選択部 33 に入力されるとともに、参照信号としてフィルタ係数設定部 35 B に入力される。また、フィルタ係数設定部 35 B には、参照信号としてスピーカ 26 から放音される音声信号である放音信号も入力される。フィルタ係数設定部 35 B は、これらの参照信号に基づいてフィルタ係数を継続的に更新する。また、このフィルタ係数の更新は、スピーカ 26 から音声放音されていて、かつ、会議室 C にいる参加者 M が発話をしていない時間区間を自動的に検出し、その時間区間の参照信号を用いて行われる。

【 0 0 2 8 】

ここで、可変フィルタ 35 A は F I R フィルタである。したがって、可変フィルタ 35 A に設定されるフィルタ係数は、フィルタ係数設定部 35 B が、スピーカ 26 からマイク素子 31 に至る音響伝搬経路のインパルス応答を推定して模擬したものである。フィルタ係数設定部 35 B は、このフィルタ係数を推定インパルス応答として、通信インタフェース 34 を介して通信部 10 に送信する。

【 0 0 2 9 】

上述したように、3 つマイク素子 31 がそれぞれ収音した音声信号は、音声選択部 33 により、そのうちの 1 つが選択されて通信器 10 に送信されるが、3 つのマイク素子 31 に対応する 3 つの推定インパルス応答は、3 つとも通信器 10 に送信される。後述するように通信器 10 のパラメータ推定部 23 では、これら 3 つの推定インパルス応答が合成される。3 つの推定インパルス応答は、図 4 に示すように対応するマイク素子 31 の向いている方向から到来する残響成分を含むインパルス応答であるが、3 つを合成することにより、全ての方向から到来する残響成分を含む無指向性マイクで収音した会議室 C の全方向からのインパルス応答を模擬することができる。

【 0 0 3 0 】

図 6 は、通信器 10 のブロック図である。通信器 10 は、パーソナルコンピュータ 2 と通信するための通信インタフェース 21、マイクミキサ 22、パラメータ推定部 23、フィルタ 24、オーディオ回路 25、スピーカ 26 および收音器 11 と通信するための通信インタフェース 27 を有する。通信インタフェース 21 は、パーソナルコンピュータ 2 とデジタル通信を行うためのインタフェースであり、たとえば U S B インタフェースが用いられる。U S B インタフェースを用いた場合、パソコン 2 がホストであり通信器 10 がオーディオデバイスとなる。通信インタフェース 27 は複数設けられており、それぞれケーブル 12 を介して個別の收音器 11 が接続される。通信インタフェース 27 は例えば有線 LAN のインタフェースを用いればよい。

【 0 0 3 1 】

通信器 10 は、通信インタフェース 27 を介して收音器 11 から音声信号（エコーをキャンセルされた収音信号）および 3 つの推定インパルス応答を受信する。受信した音声信号はマイクミキサ 22 に入力される。マイクミキサ 22 には複数の通信インタフェース 27 から、それぞれ別の收音器 11 から受信した複数の音声信号が入力される。マイクミキサ 22 は、複数の收音器 11 から受信した音声信号を選択またはミキシングしてモノラルの音声信号とし、通信インタフェース 21 を介してパーソナルコンピュータ 2 に送信する。パーソナルコンピュータ 2 は、この音声信号をネットワーク 3 を介して他拠点の音声会議システムに送信する。マイクミキサ 22 は、通信器の音声信号のレベル、持続時間または相関度を比較して S / N 比の良い音声信号を相手システムに送信する発話音声として選択すればよい。

【 0 0 3 2 】

また、パーソナルコンピュータ 2 は、他拠点の音声会議システムから音声信号を受信する。この音声信号は、通信インタフェース 21 を介して入力され、スピーカ 26 から放音される放音信号としてフィルタ 24 に入力されるとともに、通信インタフェース 27 を介して各收音器 11 に送信される。

【 0 0 3 3 】

10

20

30

40

50

フィルタ 2 4 は、会議室 C の残響による音声の明瞭度の低下を抑制するようなフィルタ処理を施す。すなわち、放音信号に対して、残響時間の長い周波数帯域のレベルを抑制するような信号処理を行う。特に、低音域の残響が明瞭度の低下の原因になるため、低音域に対しては抑制の度合いを強くする。このようなフィルタ係数は、パラメータ推定部 2 3 によって決定される。フィルタ 2 4 によって残響時間の長い周波数帯域が抑制された放音信号は、オーディオ回路 2 5 に入力される。オーディオ回路 2 5 は、放音信号をアナログの音声信号を変換し、所定のレベルで増幅してスピーカ 2 6 に入力する。スピーカ 2 6 は、この放音信号を音声として会議室 C に放音する。放音された音声は会議の出席者 M に聴取されるとともに、マイク 3 1 によって收音される。

【 0 0 3 4 】

通信インタフェース 2 7 を介して收音器 1 1 に送信された放音信号は、図 5 に示したエコーキャンセラ 3 2 のフィルタ係数設定部 3 5 B に参照信号として入力される。

【 0 0 3 5 】

図 7 はパラメータ推定部 2 3 のブロック図である。また、図 8 は、パラメータ推定部 2 3 を含む音声会議システム 1 で実行される残響抑制処理の手順を示す図である。また、図 9 は、残響抑制処理の手順のなかで現れる信号波形を例示した図である。

【 0 0 3 6 】

図 8 において、指向性のマイク素子 3 1 による收音 (S 1 0 1)、エコーキャンセル処理 (S 1 0 2)、および、適応フィルタ 3 5 からのフィルタ係数 (推定インパルス応答) の取り出し (S 1 0 3) は、收音器 1 1 が行う。收音器 1 1 は、3 つのマイク素子 3 1 に対応して設けられた 3 つのエコーキャンセラ 3 2 のフィルタ係数を推定インパルス応答として通信器 1 0 に対して送信する。

【 0 0 3 7 】

図 7 において、パラメータ推定部 2 3 は、接続している收音器 1 1 ごとのフィルタ係数統合部 4 0、同じく接続している收音器 1 1 ごとの残響時間推定部 4 1、残響時間統合部 4 2、補正特性算出部 4 3、および、フィルタ係数算出部 4 4 を備える。

【 0 0 3 8 】

通信器 1 0 は、各收音器 1 1 から、それぞれ 3 つずつの推定インパルス応答 (フィルタ係数) を受信する。受信した推定インパルス応答は、パラメータ推定部 2 3 に入力される。パラメータ推定部 2 3 では、入力された推定インパルス応答が、各收音器 1 1 ごとに設けられたフィルタ係数統合部 4 0 にそれぞれ入力される。フィルタ係数統合部 4 0 は、入力された 3 つの推定インパルス応答を時間軸を揃えて合成する。この合成は、単純に加算合成してもよいし、各推定インパルス応答の重みづけを変えてもよいし、各インパルス応答の時間ずれを補正しても良い。図 4 に示した 3 方向の推定インパルス応答を合成することにより、1 マイクに関する推定インパルス応答よりも広い方向の範囲から到来する残響成分を含む (理想的には 3 6 0 度無指向性の) インパルス応答を推定することが可能である。この処理が、図 8 における S 1 0 4 の前段統合である。この処理は、接続された (推定インパルス応答が入力された) 收音器 1 1 毎に行われ、各收音器 1 1 の位置におけるインパルス応答が推定される。

【 0 0 3 9 】

フィルタ係数統合部 4 0 で合成された広指向性の推定インパルス応答は、残響時間推定部 4 1 に入力される。残響時間推定部 4 1 では以下の処理が行われる。まず、推定インパルス応答を複数チャンネルのバンド・パス・フィルタに通して帯域を分割する。分割するチャンネルの数や各チャンネルの周波数帯域は任意であるが、例えば、3 1 5 H z ~ 8 0 0 0 H z までを 1 5 のチャンネルに分割するなどの帯域分割を行えばよい。この処理により、各周波数帯域 (チャンネル) の信号成分のインパルス応答が推定される。この処理が図 8 の S 1 0 5 の処理である。この処理も各收音器 1 1 毎に行われる。

【 0 0 4 0 】

残響時間推定部 4 1 は、各周波数帯域の推定インパルス応答に基づき、各周波数帯域の信号の残響時間を求める。残響時間は、一般的に信号レベルが - 6 0 d B (1 0 0 万分の

10

20

30

40

50

1) に減衰するまでの時間を言い、その算出・推定方法は種々あるが、ここではシュレーダー法により残響時間を求めればよい。シュレーダー法は、インパルス応答をシュレーダー積分または後方累積加算によって、図9(A)に例示するようなシュレーダー曲線(残響減衰曲線)を求め、この曲線が-60dBになるまでの時間を求めれば良い。簡略にはシュレーダー曲線の直接音や誤差成分等の含まれない所定区間を取り出してその区間の傾きをこの曲線の傾きとし、この傾きにより0dBから-60dBまで減衰する時間を推定してもよい。この処理が、図8のS106に相当し、各収音器11毎の周波数帯域毎に行われ、各収音器11の位置における周波数帯域毎の残響時間が推定される。

【0041】

そして、複数の残響時間推定部41によって推定された、各収音器11の位置における周波数帯域毎の残響時間は、残響時間統合部42に入力される。残響時間統合部42は、各収音器11の位置における残響時間を周波数帯域毎に合成する。この処理が図8のS107の後段統合処理であり、各周波数帯域毎に行われる。

【0042】

この後段統合処理の合成は、周波数帯域毎に各収音器11の残響時間を加算平均することによって行われるが、平均値から極端に離れている残響時間値(外れ値)は、平均から除外してもよい。また、外れ値が多い収音器11については、部屋の隅など偏った特性が出やすい場所に設置されている可能性があるとして、その収音器11の残響時間は全ての周波数帯域において全面的に後段統合処理から除外するようにしてもよい。この除外の処理は、通信器10の判断で行ってもよく、会議室Cにこのシステムを設置する係員が、収音器11または通信器10に対するマニュアル操作で、特定の収音器11、たとえば上記部屋の隅に設置される収音器11などをこの後段統合の対象から除外するように設定可能にしてもよい。この場合、統合対象から除外された収音器11については、S103以下の処理は不要になり、処理が軽減される。

【0043】

後段統合処理によって求められた各周波数帯域毎の残響時間を周波数軸にプロットすることにより、例えば、図9(B)に示すような会議室C全体の平均した残響特性が求められる。この残響特性に基づき、どの周波数帯域の残響時間が長いかなどの判定が可能になる。

【0044】

残響時間統合部42によって求められた残響特性は、補正特性算出部43に入力される。補正特性算出部43は、入力された残響特性に基づき、スピーカ26から放音された音声が、その音声の残響音によって覆い隠されないように、残響時間の長い周波数帯域を抑制するような周波数領域の補正特性を決定する。補正特性の決定は、周波数帯域ごとに残響時間の閾値を定めておき、残響時間が閾値を超えている周波数帯域を抽出してこの周波数帯域のパワーを抑制する方法や、残響音を抑制する既知のフィルタリング法、周波数帯域ごとのゲインテーブルを用いることで周波数帯域ごとのパワーの抑制量を定める方法などを選択することができる。周波数帯域ごとのゲインテーブルは図10に示すようなものが使用できる。このゲインテーブルは縦軸がゲイン(dB)、横軸が残響時間RT(秒)となっており、周波数帯域ごとのゲイン値が傾きを有する線分で示されている。f1~fnの線分は、上述したバンドパスフィルタにより分割された周波数帯域に対応し、f1が低音域側、fnが高音域側となっている。例えば帯域f3において残響時間が1.0秒であればゲインは-30dBと定められる。このゲインテーブルでは低音域の線分がより急峻な傾きをもつように設定されている。高帯域側のf4では残響時間が1.0秒であればゲインは-24dB付近となる。このように低音域の残響時間が長い場合には、この低音域を高音域の残響時間が長い場合よりも強く抑制するような補正特性が決定される。

【0045】

なお、ゲイン値には下限があり、所定の値(図に示したゲインテーブルでは-30dB)を超える抑制は行われないようにしている。また各周波数帯域において所定の残響時間(f3においては1.0秒)を超えた場合はゲインの下限値を適用するようにしている。

10

20

30

40

50

また、ゲインテーブルは図 1 1 のように複数の線分が収束する収束点が一定の残響時間だけ正方向にずれて設定されているものでもよい。この図の場合は残響時間が 1 . 0 秒以下の場合にはゲインが 0 となる。この処理が図 8 の S 1 0 8 である。決定された補正特性はフィルタ係数算出部 4 4 に入力される。

【 0 0 4 6 】

フィルタ係数算出部 4 4 は、フィルタ 2 4 が補正特性算出部 4 3 で算出された補正特性を持つようにフィルタ特性を決定する。フィルタ 2 4 は、F I R フィルタや I I R フィルタで構成される。フィルタ係数は、フィルタ 2 4 の構成に合わせて、離散時間逆フーリエ変換や、パラメトリックなピークフィルタなどの演算によって算出される。この処理が図 8 の S 1 0 9 である。算出されたフィルタ係数は、フィルタ 2 4 に設定される (S 1 1 0)。このフィルタ 2 4 で放音信号をフィルタリングすることにより、スピーカ 2 6 から放音された音声の残響が抑制され、参加者 M にとって明瞭度の高い音声となる。

10

【 0 0 4 7 】

以上の実施形態では、エコーキャンセラ 3 2 のフィルタ係数を用いて会議室 C の残響特性を推定し、残響の長い周波数帯域を抑制することによって放音された音声の明瞭度が低下しないようにした。さらに、エコーキャンセラ 3 2 のフィルタ係数を用いて会議室 C の周波数特性を推定し、放音された音声フラットな特性で聴取されるように、放音信号の周波数特性を、会議室 C の周波数特性をキャンセルするように補正してもよい。これにより、残響だけでなく会議室 C の周波数特性を原因とする明瞭度の低下も防ぐことができる。

20

【 0 0 4 8 】

図 1 2 に、パラメータ推定部の変形例を示す。この図のパラメータ推定部 2 3 は、残響を抑制する補正特性に加えて、周波数特性を補正する補正特性を決定してフィルタ 2 4 に設定する。図 1 2 において、図 7 と同一構成の部分は同一番号を付して説明を省略する。この図のパラメータ推定部 2 3 は、図 7 のパラメータ推定部 2 3 の構成に加えて、各収音器 1 1 毎の周波数特性推定部 4 5、および、周波数特性統合部 4 6 をさらに備えている。

【 0 0 4 9 】

フィルタ係数統合部 4 0 から出力された各収音器 1 1 毎の広指向性 (無指向性) の推定インパルス応答は、残響時間推定部 4 1 に入力されるとともに、周波数特性推定部 4 5 に入力される。周波数特性推定部 4 5 は、入力されたインパルス応答をフーリエ変換して、その収音器 1 1 の位置における周波数特性を算出する。この周波数特性は周波数特定統合部 4 6 に入力される。周波数特性統合部 4 6 は、各周波数特性推定部 4 5 から入力された各収音器 1 1 における周波数特性を合成して、会議室 C における全体の周波数特性の平均値を算出する。この平均値の算出は、単純に相加平均してもよく、各周波数特性を正規化したのち平均を求めても良い。

30

【 0 0 5 0 】

周波数特性統合部 4 6 によって求められた会議室 C の周波数特性は、補正特性算出部 4 3 に入力される。補正特性算出部 4 3 は、残響時間の長い周波数帯域を抑制するような特性であり、且つ、放音された音声は会議室 C に影響された周波数特性をキャンセルしてフラットな伝達特性を経て聴者に届くような補正特性を算出する。またこの算出方法は、フラットな伝達特性を経て聴者に届くような補正特性ではなく、あらかじめ設定された任意の理想的な伝達特性を経て聴者に届くような補正特性を算出してもよい。算出された補正特性はフィルタ係数算出部 4 4 に入力される。フィルタ係数算出部 4 4 は、フィルタ 2 4 が補正特性算出部 4 3 で算出された補正特性を持つようにフィルタ特性を決定する。算出されたフィルタ係数は、フィルタ 2 4 に設定される。このフィルタ 2 4 で放音信号をフィルタリングすることにより、スピーカ 2 6 から放音された音声はフラットな伝達特性を経たような特性を有するとともにその残響が抑制され、参加者 M にとって明瞭度の高い音声となる。

40

【 0 0 5 1 】

50

上記の実施形態では、前段統合を通信器 10 が行っているが、これを収音器 11 に行わせてもよい。図 8 において、S 101 - S 103 の処理は収音器 11 で行うことが好適である。また、S 107 以下の処理は通信器 10 が行うことが好適である。その間の S 104 - S 106 の処理は、収音器 11、通信器 10 のいずれが行ってもよい。

【0052】

上記の実施形態では、スピーカ 26 を備えた通信器 10 に、マイク素子 31 を備えた収音器 11 が接続される形態の音声会議システムについて説明したが、複数のマイク素子 31 とスピーカ 26 を一体に備えた（通信器 10 のみの）音声会議装置に対しても、本発明は適用可能である。

【0053】

また、通信器 10 と収音器 11 との接続の形態は、有線接続に限定されない。たとえば、無線 LAN や近距離無線通信規格などの無線接続であってもよい。

【0054】

また、収音器 11 の形状やマイク素子 31 の数は、図 1 や図 4 に示したものに限定されない。たとえば、円盤状の筐体の周縁部に 2 つまたは 4 つのマイク素子 31 を等間隔に設けてもよい。この場合、各マイク素子 31 の角度は、180 度または 90 度となる。また、間隔（角度）は等間隔でなくてもよい。会議参加者 M に向ける方向に片寄せてマイク素子 31 を設けても良い。

【0055】

通信器 10 に複数台の収音器 11 を接続する場合、図 13 に示すように、複数の収音器 11 をケーブル 12 でディジーチェーン接続する形態でもよい。この接続形態にすることで、全ケーブル長を節約することができる。通信インタフェース 21 および 34 を LAN インタフェースとすることで、図 1 のスター型の接続および図 13 のディジーチェーン型の接続の両方に対応させることも可能である。

【0056】

また、図 14 に示すように、1 つのマイク素子 31 を内蔵するマイク 51, 52, 53, 54 を複数組み合わせ（グループ化）し、この 1 つのグループ 60 を 1 つの収音器 11 として機能させてもよい。この場合、係員がテーブルマイク 51 のグルーピング情報を予め通信器 10 に設定しておくようにしてもよく、通信器 10 のフロントエンドに信号振分部を設けて、通信器 10 が自らグルーピングを行うようにしてもよい。この場合、信号振分部は、エコーキャンセラの適応フィルタの時間位置や、收音された音声信号の相関度などに基づき、類似した信号を收音しているテーブルマイク同士を同じグループとしてグルーピングするようにしてもよい。

【0057】

図 14 (A) は、複数のテーブルマイク（スタンドマイク）51 を組み合わせるグループ 60 とした例を示している。また、図 14 (B) は、複数のハンドマイク 52 をグループ 60 化した例を示している。ハンドマイク 52 は有線であっても無線であってもよい。この場合はハンドマイクを持った話者が動くため一定距離に存在する複数のハンドマイク 52 を 1 つのグループ 60 として、各グループ 60 ごとに適応フィルタ係数列を加算してもよい。ハンドマイク 52 のグループ 60 が複数存在する場合は、各グループの残響時間を算出して上述した残響特性を求めればよい。複数のハンドマイク 52 が一定距離に存在することは、2 つのハンドマイク 52 の收音遅延差を算出することによる位置検出や、ハンドマイク 52 が発する無線の強度を相互に検出することで判定することができる。

【0058】

また、収音器 11 は、会議机 D 上に載置されるものでなくてもよい。すなわち、図 14 (C) に示すように、天井から吊り下げられる吊りマイク 53 や壁面に設置される壁面マイク 54 などを用いてもよい。勿論、図 1 の収音器 11、テーブルマイク 51、ハンドマイク 52、吊りマイク 53 および壁面マイク 54 が混在していてもよい。

【0059】

また、本実施形態の音声会議システム 1 は、会議以外に用いることも当然可能である。

10

20

30

40

50

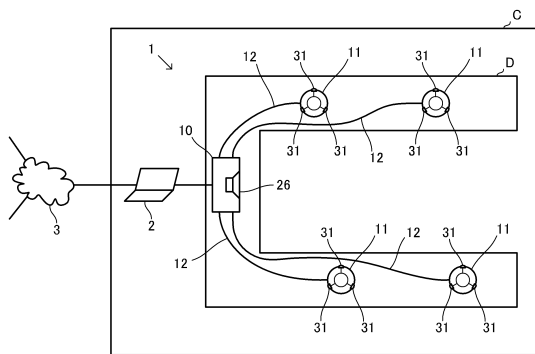
また、使用される場所も会議室に限定されない。

【符号の説明】

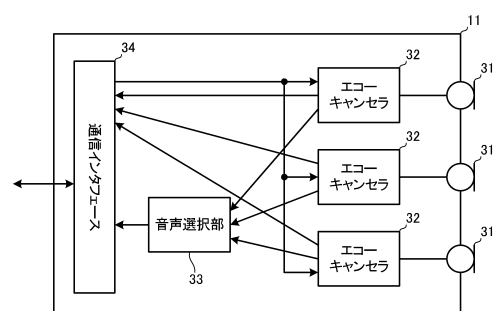
【0060】

- C 会議室
- D 会議机
- M 会議の参加者
- 1 音声会議システム
- 2 パーソナルコンピュータ
- 3 ネットワーク
- 10 通信器
- 11 収音器
- 26 スピーカ
- 31 マイク素子
- 51 テーブルマイク
- 52 ハンドマイク
- 53 吊りマイク
- 54 壁面マイク
- 60 グループ

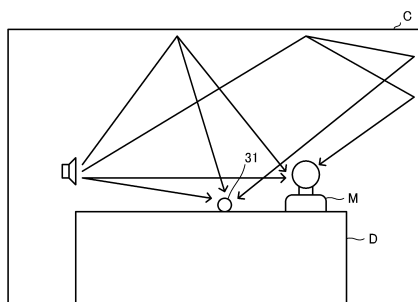
【図1】



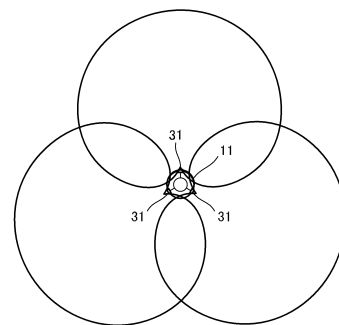
【図3】



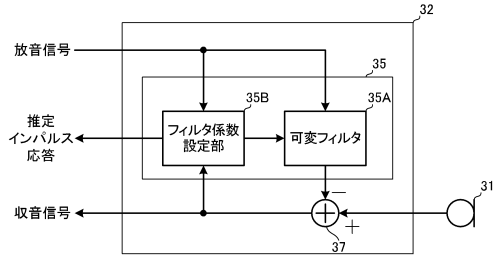
【図2】



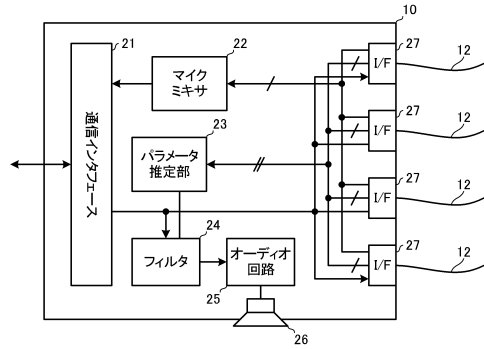
【図4】



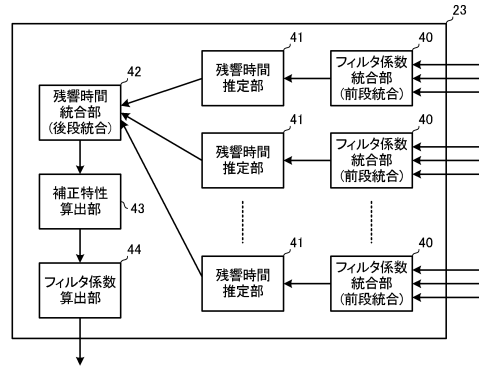
【図5】



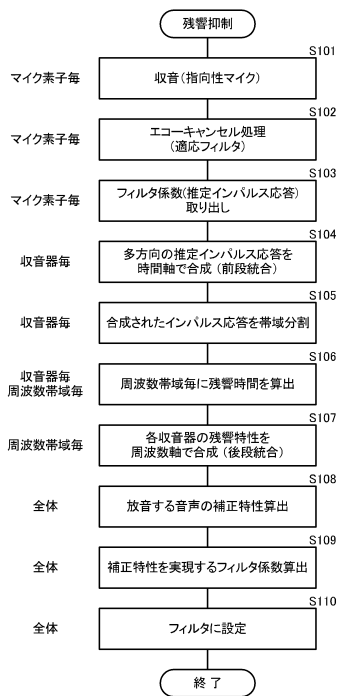
【図6】



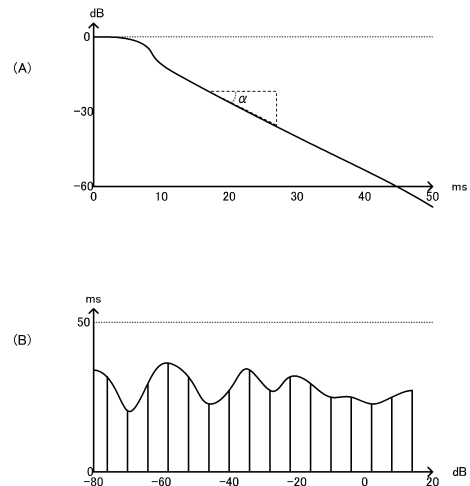
【図7】



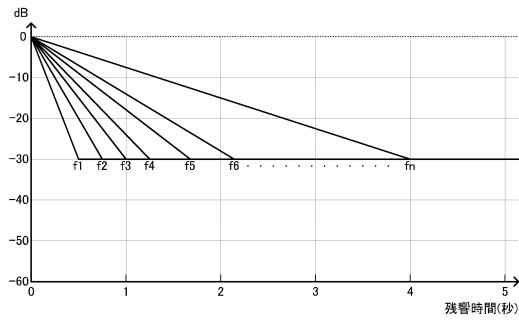
【図8】



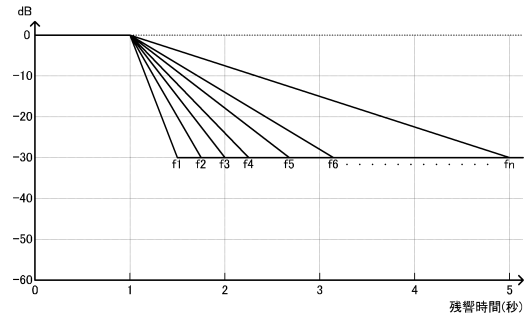
【図9】



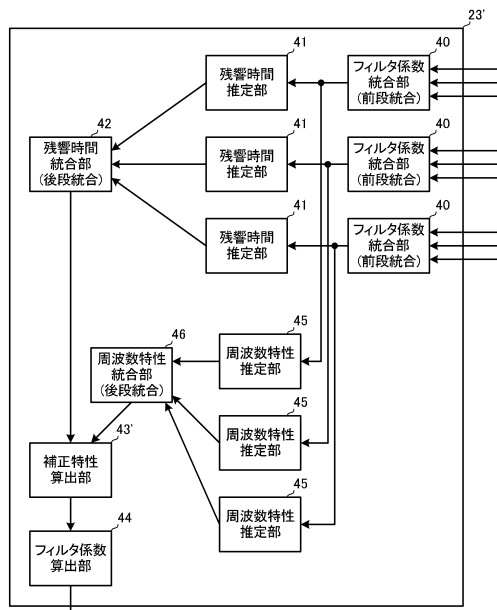
【図10】



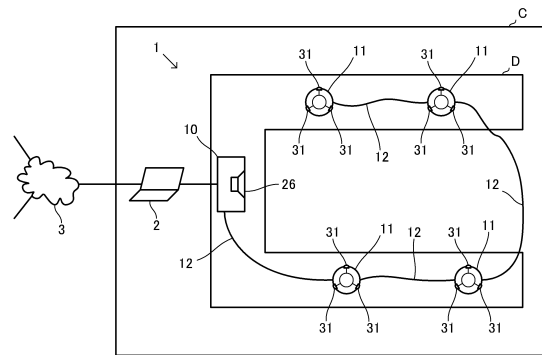
【図11】



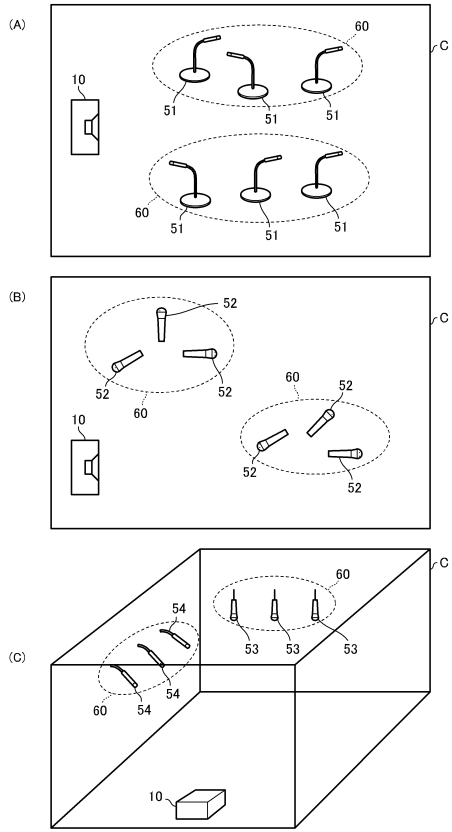
【図12】



【図13】



【 図 14 】



フロントページの続き

(51)Int.Cl. F I
H 0 4 M 1/60 C

(56)参考文献 特開昭64-029094(JP,A)
特開平10-150343(JP,A)

(58)調査した分野(Int.Cl., DB名)

| | | | |
|---------|---------|---|---------|
| H 0 4 R | 3 / 0 0 | - | 3 / 1 4 |
| H 0 4 R | 1 / 4 0 | | |
| H 0 4 M | 1 / 2 4 | - | 1 / 8 2 |
| H 0 4 M | 3 / 5 6 | | |