

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2015-519614

(P2015-519614A)

(43) 公表日 平成27年7月9日(2015.7.9)

(51) Int.Cl.

G 1 O L 21/0264 (2013.01)
G 1 O L 21/0232 (2013.01)

F 1

G 1 O L 21/0264
G 1 O L 21/0232

テーマコード(参考)

Z

審査請求 有 予備審査請求 未請求 (全 30 頁)

(21) 出願番号 特願2015-516415 (P2015-516415)
 (86) (22) 出願日 平成25年4月1日 (2013.4.1)
 (85) 翻訳文提出日 平成26年12月16日 (2014.12.16)
 (86) 國際出願番号 PCT/CN2013/073584
 (87) 國際公開番号 WO2013/189199
 (87) 國際公開日 平成25年12月27日 (2013.12.27)
 (31) 優先権主張番号 201210201879.7
 (32) 優先日 平成24年6月18日 (2012.6.18)
 (33) 優先権主張國 中国(CN)

(71) 出願人 512280079
 ゴーテック インコーポレイテッド
 GOERTEK INC
 中華人民共和国、シャントン 26103
 1、ウェイハン、ハイテック インダス
 トリー ディベロップメント ディストリ
 クト、ドンファン ロード ナンバー26
 8
 (74) 代理人 100129425
 弁理士 小川 譲晃
 (74) 代理人 100098623
 弁理士 奥山 尚一
 (74) 代理人 100087505
 弁理士 西山 春之

最終頁に続く

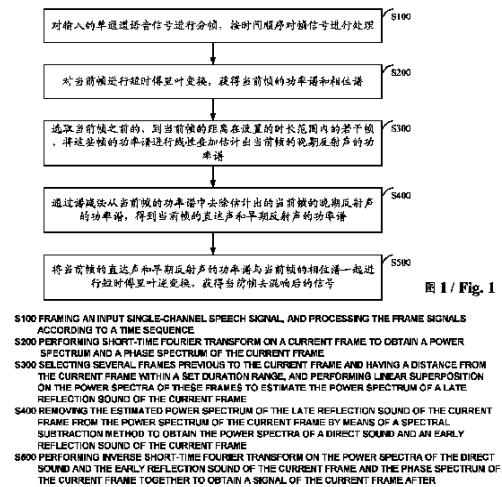
(54) 【発明の名称】シングルチャンネル音声残響除去方法及びその装置

(57) 【要約】

【課題】シングルチャンネル音声残響除去において残響環境の伝達関数又は残響時間を推定し難いという問題を解決する。

【解決手段】入力されたシングルチャンネル音声信号に対してフレーム分割を行い、時間の順に応じてフレーム信号に対して短時間フーリエ変換を行い、現在フレームのパワースペクトル及び位相スペクトルを獲得し、現在フレームに至るまでの距離が設定された持続時間範囲内である数フレームを選んで、これらのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定し、スペクトル減算法によって、現在フレームのパワースペクトルから、前記推定された後期反射音のパワースペクトルを除去して、現在フレームの直接音及び早期反射音のパワースペクトルを獲得し、この獲得したパワースペクトルを現在フレームの位相スペクトルとともに、短時間逆フーリエ変換を行い、現在フレーム残響除去後の信号を獲得する。

【選択図】図1



【特許請求の範囲】**【請求項 1】**

入力されたシングルチャンネル音声信号に対してフレーム分割を行い、時間の順に応じてフレーム信号に対して、

現在フレームに対して短時間フーリエ変換を行い、現在フレームのパワースペクトル及び位相スペクトルを獲得する処理と、

現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内である数フレームを選んで、これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理と、

スペクトル減算法によって、現在フレームのパワースペクトルから、推定された現在フレームの後期反射音のパワースペクトルを除去して、現在フレームの直接音及び早期反射音のパワースペクトルを獲得する処理と、

現在フレームの直接音及び早期反射音のパワースペクトルを現在フレームの位相スペクトルとともに、短時間逆フーリエ変換を行い、現在フレーム残響除去後の信号を獲得する処理と、

を行うことを含むことを特徴とするシングルチャンネル音声残響除去方法。

【請求項 2】

前記後期反射音の減衰特性に基づいて前記持続時間範囲の上限値が設定され、

及び/又は、音声関連特性及び直接音と早期反射音の残響環境下でのインパルス応答分布領域に基づいて、前記持続時間範囲の下限値が設定されることを特徴とする請求項 1 に記載のシングルチャンネル音声残響除去方法。

【請求項 3】

前記持続時間範囲の上限値が 0 . 3 s ~ 0 . 5 s の範囲内の値を選択することを特徴とする請求項 1 に記載のシングルチャンネル音声残響除去方法。

【請求項 4】

前記持続時間範囲の下限値が 50 ms ~ 80 ms の範囲内の値を選択することを特徴とする請求項 1 に記載のシングルチャンネル音声残響除去方法。

【請求項 5】

前記持続時間範囲内である数フレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理は、

自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理、

或は、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理、

或は、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算するとともに、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理、

を含むことを特徴とする請求項 1 に記載のシングルチャンネル音声残響除去方法。

【請求項 6】

入力されたシングルチャンネル音声信号に対してフレーム分割を行い、時間の順に応じてフレーム信号をフーリエ変換ユニットへ出力するためのフレーム分割ユニットと、

受信された現在フレームに対して短時間フーリエ変換処理を行い、現在フレームのパワースペクトル及び位相スペクトルを獲得して、現在フレームのパワースペクトルをスペクトル減算ユニットとスペクトル推定ユニットへ出力し、位相スペクトルを逆フーリエ変換ユニットへ出力するためのフーリエ変換ユニットと、

現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内である数フレームのパワースペクトルを線形重畠加算して、現在フレームの後期反射音のパワースペクトルを推定し、推定された現在フレームの後期反射音のパワースペクトルをスペ

10

20

30

40

50

クトル減算ユニットへ出力するためのスペクトル推定ユニットと、

スペクトル減算法によって、フーリエ変換ユニットより獲得した現在フレームのパワースペクトルから、スペクトル推定ユニットより獲得した現在フレームの後期反射音のパワースペクトルを除去して、現在フレームの直接音及び早期反射音のパワースペクトルを獲得し、現在フレームの直接音及び早期反射音のパワースペクトルを逆フーリエ変換ユニットへ出力するためのスペクトル減算ユニットと、

スペクトル減算ユニットより獲得した現在フレームの直接音及び早期反射音のパワースペクトルを、フーリエ変換ユニットより獲得した現在フレームの位相スペクトルとともに、短時間逆フーリエ変換を行い、現在フレーム残響除去後の信号を出力するための逆フーリエ変換ユニットと、

を含むことを特徴とするシングルチャンネル音声残響除去装置。

【請求項 7】

前記スペクトル推定ユニットは、後期反射音の減衰特性に基づいて前記持続時間範囲の上限値を設定し、及び/又は、音声関連特性及び直接音と早期反射音の残響環境下でのインパルス応答分布領域に基づいて、前記持続時間範囲の下限値を設定するために用いられることを特徴とする請求項 6 に記載のシングルチャンネル音声残響除去装置。

【請求項 8】

前記スペクトル推定ユニットは、持続時間範囲の上限値が 0.3 s ~ 0.5 s の範囲内の値を選択するために用いられることを特徴とする請求項 6 に記載のシングルチャンネル音声残響除去装置。

【請求項 9】

前記スペクトル推定ユニットは、持続時間範囲の下限値が 50 ms ~ 80 ms の範囲内の値を選択するために用いられることを特徴とする請求項 6 に記載のシングルチャンネル音声残響除去装置。

【請求項 10】

前記スペクトル推定ユニットは、

現在フレームの前の、現在フレームに至るまでの距離が前記設定された持続時間範囲内である数フレームに対して、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定するか、

或は、現在フレームの前の、現在フレームに至るまでの距離が前記設定された持続時間範囲内である数フレームに対して、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定するか、

或は、現在フレームの前の、現在フレームに至るまでの距離が前記設定された持続時間範囲内である数フレームに対して、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算するとともに、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する、

ために用いられることを特徴とする請求項 6 に記載のシングルチャンネル音声残響除去装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声強調分野に関し、特に、シングルチャンネル音声残響除去方法及びその装置に関する。

【背景技術】

【0002】

電話会議、スマートテレピネットワーク電話などの音声通信において、発話者がマイクロフォンから遠く離れ、且つ通話環境が相対的に密閉される空間であるため、マイクロフ

10

20

30

40

50

オンが受信した信号は環境残響に影響されやすい。例えば、部屋の中で、音声が壁面、床板、家具などにより複数回の反射を経て、マイクロフォンが受信した信号は直接音と反射音の混合信号となる。この反射音でさえ残響信号である。残響がひどいとき、音声は不明瞭となり、通話品質に影響を与えててしまう。また、残響による干渉は、さらに音声学受信システムの性能を劣化させたり、音声識別システムの性能を低下させたりしてしまう。

【0003】

早期の残響除去方法は、主にデコンヴォルーションを活用して実行していた。このような方法では、事前に残響環境のインパルス応答又は伝達関数を正確に把握する必要がある。残響環境のインパルス応答はある特別な方法又は装置を介して事前に検出することができ、また、その他の方法を介して単独で推定できる。そして、この知っていた残響環境のインパルス応答を活用して、逆フィルタを推定し、残響信号のデコンヴォルーションを実現することによって、残響除去を実現する。このような方法の問題点は、残響環境のインパルス応答を事前に獲得するのがより困難であり、且つ逆フィルタの推定プロセス自体も新しい不安定な要素を引き入れる可能性があるというのである。

10

【0004】

もう1種の残響除去の方法としては、残響環境のインパルス応答を推定する必要がないため、逆フィルタ計算及び逆平滑演算が必要ではなくなり、ブラインド残響除去方法とも称されている。この種の方法は、音声モデルの仮説に基づいて、例えば、残響は、受信された濁音励磁パルスに変化を起こして、その周期性がある程度不明瞭になるようして、引いて音声の明瞭度を影響するのは一般的です。この種の方法は、通常、LPC (Linear Prediction Coding, 線形予測符号) モデルに基づき、音声を生じるモデルが全極型モデルであると仮説して、残響又は他の加法性ノイズはシステム全体に新しい零点を引き入れることで、濁音励磁パルスを干渉するが、全極型フィルタに影響を与えない。残響除去の方法としては、信号のLPC残差を推定して、そして、ピッチ同期クラスタリング基準 (pitch-synchronous clustering criterion) 又は尖度 (Kurtosis) 最大化基準などに基づいて、クリアなパルス励磁配列を推定することにより、残響除去を実現する。この種の方法の問題は、計算の複雑度が非常に高く、且つ残響に対して全零点フィルタの仮説のみに影響を及ぼし、実験分析と一致しないことがあるというのである。

20

【0005】

スペクトル減算法を用いて残響を除去するのは好ましい方法であり、音声信号が直接音、早期反射音及び後期反射音を含み、スペクトル減算法を用いて、後期反射音のパワースペクトルを音声全体のパワースペクトルから除去することで音声品質を向上することができる。しかし、その中で最も主な問題は後期反射音のスペクトルの推定にあたり、即ち、如何により正確な後期反射音のパワースペクトルを獲得して、後期反射音の成分を効果的に除去するとともに音声を損傷しないことができるのかである。シングルチャンネル音声残響除去において、モノパスマイクロフォン信号しか使用できないため、残響環境の伝達関数又は残響時間 (RT60) を推定することが非常に困難である。

30

【発明の概要】

【発明が解決しようとする課題】

【0006】

本発明は、シングルチャンネル音声残響除去においての残響環境の伝達関数又は残響時間を推定しにくいという問題を解決するために、シングルチャンネル音声残響除去方法及びその装置を提供する。

40

【課題を解決するための手段】

【0007】

本発明は、シングルチャンネル音声残響除去方法を開示しており、
入力されたシングルチャンネル音声信号に対してフレーム分割を行い、時間の順に応じてフレーム信号に対して、

現在フレームに対して短時間フーリエ変換を行い、現在フレームのパワースペクトル及び位相スペクトルを獲得する処理と、

50

現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内である数フレームを選んで、これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理と、

スペクトル減算法によって、現在フレームのパワースペクトルから、推定された現在フレームの後期反射音のパワースペクトルを除去して、現在フレームの直接音及び早期反射音のパワースペクトルを獲得する処理と、

現在フレームの直接音及び早期反射音のパワースペクトルを現在フレームの位相スペクトルとともに、短時間逆フーリエ変換を行い、現在フレーム残響除去後の信号を獲得する処理と、を行うことを含む。

【0008】

前記後期反射音の減衰特性に基づいて前記持続時間範囲の上限値を設定することが好ましく、

及び/又は、音声関連特性及び直接音と早期反射音の残響環境下でのインパルス応答分布領域に基づいて、前記持続時間範囲の下限値を設定することが好ましい。

【0009】

前記持続時間範囲の上限値が0.3s～0.5sの範囲内の値を選択することが好ましい。

【0010】

前記持続時間範囲の下限値が50ms～80msの範囲内の値を選択することが好ましい。

【0011】

前記持続時間範囲内である数フレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理は、

自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理、

或は、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理、

或は、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算するとともに、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理、を含むことが好ましい。

【0012】

本発明は、また、シングルチャンネル音声残響除去装置を開示しており、

入力されたシングルチャンネル音声信号に対してフレーム分割を行い、時間の順に応じてフレーム信号をフーリエ変換ユニットへ出力するためのフレーム分割ユニットと、

受信された現在フレームに対して短時間フーリエ変換処理を行い、現在フレームのパワースペクトル及び位相スペクトルを獲得して、現在フレームのパワースペクトルをスペクトル減算ユニットとスペクトル推定ユニットへ出力し、位相スペクトルを逆フーリエ変換ユニットへ出力するためのフーリエ変換ユニットと、

現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内である数フレームのパワースペクトルを線形重畠加算して、現在フレームの後期反射音のパワースペクトルを推定し、そして推定された現在フレームの後期反射音のパワースペクトルをスペクトル減算ユニットへ出力するためのスペクトル推定ユニットと、

スペクトル減算法によって、フーリエ変換ユニットより獲得した現在フレームのパワースペクトルから、スペクトル推定ユニットより獲得した現在フレームの後期反射音のパワースペクトルを除去して、現在フレームの直接音及び早期反射音のパワースペクトルを獲得し、現在フレームの直接音及び早期反射音のパワースペクトルを逆フーリエ変換ユニットへ出力するためのスペクトル減算ユニットと、

スペクトル減算ユニットより獲得した現在フレームの直接音及び早期反射音のパワースペクトルを、フーリエ変換ユニットより獲得した現在フレームの位相スペクトルとともに

10

20

30

40

50

、短時間逆フーリエ変換を行い、現在フレーム残響除去後の信号を出力するための逆フーリエ変換ユニットと、を含む。

【0013】

前記スペクトル推定ユニットは、後期反射音の減衰特性に基づいて前記持続時間範囲の上限値を設定し、及び/又は、音声関連特性及び直接音と早期反射音の残響環境下でのインパルス応答分布領域に基づいて、前記持続時間範囲の下限値を設定するために用いられることが好ましい。

【0014】

前記スペクトル推定ユニットは、持続時間範囲の上限値が0.3s～0.5sの範囲内の値を選択するために用いられることが好ましい。

10

【0015】

前記スペクトル推定ユニットは、持続時間範囲の下限値が50ms～80msの範囲内の値を選択するために用いられることが好ましい。

【0016】

前記スペクトル推定ユニットは、

現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内である数フレームに対して、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定するか、

或は、現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内である数フレームに対して、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定するか、

20

或は、現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内である数フレームに対して、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算するとともに、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する、ために用いられることが好ましい。

【発明の効果】

30

【0017】

本発明の実施例の有益な効果は、現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内である数フレームを選べることによって、これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定することで、残響環境の伝達関数又は残響時間を推定する必要がなく、現在フレームの後期反射音のパワースペクトルを推定することができ、そしてスペクトル減算法を用いて残響除去ができるため、残響除去の操作を簡略化して、より簡単に実現できることと、

音声関連特性及び直接音と早期反射音の残響環境下でのインパルス応答分布領域に基づいて、持続時間範囲の下限値を設定することによって、残響除去で同時に有用な直接音及び早期反射音を保留して、音声の品質を向上できることと、

40

後期反射音の減衰特性に基づいて持続時間範囲の上限値を設定することによって、推定された後期反射音のパワースペクトルの正確性を確保できると同時に重畠加算量を減少できることと、

本発明の実施例が上限値を0.3s～0.5sの範囲内の値と選択しており、該上限値が実験により得たしきい値であり、残響環境が変化するとき、該上限値を調整しなくても、より良い残響除去の効果を得られることと、

本発明の実施例が下限値を50ms～80msの範囲内の値と選択しており、残響環境が変化するとき、該下限値を変えなくても、効果的に直接音及び早期反射音を避けて重畠加算を行うことができ、重畠加算の結果に直接音及び早期反射音がほぼ含まれず、残響除去で同時に有用な直接音及び早期反射音を保留し、より良い音声品質を得られることにある

50

。 前記残響環境の変化は、残響がない消音室から残響が極めてひどい大ホールにまでの変化を含む。

【図面の簡単な説明】

【0018】

【図1】本発明のシングルチャンネル音声残響除去方法の流れの模式図である。

【図2】実際の部屋においてのインパルス応答の模式図である。

【図3A】本発明の実施効果の模式図であり、残響信号の時間ドメインの模式図である。

【図3B】本発明の実施効果の模式図であり、残響除去後信号の時間ドメインの模式図である。

10

【図3C】本発明の実施効果の模式図であり、残響信号及び残響除去後信号のエネルギー包絡曲線である。

【図4】本発明のシングルチャンネル音声残響除去装置の構造図である。

【図5】本発明のシングルチャンネル音声残響除去装置の具体的な実施形態の構造図である。

【発明を実施するための形態】

【0019】

本発明の目的、技術的なソリューション及び利点をより明らかにするために、以下に、図面を参照しながら、本発明の実施形態をさらに詳しく説明する。

【0020】

図1は、本発明のシングルチャンネル音声残響除去方法の流れの模式図である。

ステップS100は、入力されたシングルチャンネル音声信号に対してフレーム分割を行い、時間の順に応じてフレーム信号に対して下記の処理を行う。

【0021】

ステップS200は、現在フレームに対して短時間フーリエ変換を行い、現在フレームのパワースペクトル及び位相スペクトルを獲得する。

【0022】

ステップS300は、現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内である数フレームを選んで、これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する。

前記数フレームは予定数量のフレームであり、持続時間範囲内の全てのフレーム又は該持続時間範囲内の一部のフレームであってもよい。

30

【0023】

ステップS400は、スペクトル減算法によって、現在フレームのパワースペクトルから、推定された現在フレームの後期反射音のパワースペクトルを除去して、現在フレームの直接音及び早期反射音のパワースペクトルを獲得する。

【0024】

ステップS500は、現在フレームの直接音及び早期反射音のパワースペクトルを現在フレームの位相スペクトルとともに、短時間逆フーリエ変換を行い、現在フレーム残響除去後の信号を獲得する。

【0025】

残響環境において、マイクロフォンが採取した信号 $x(t)$ 、即ちシングルチャンネル音声信号は、直接音と反射音との混合であり、下記の残響モデルで表してもよい。

40

$$x(t) = h * s(t) + n(t)$$

そのうち、 $s(t)$ は音源から発信した信号であり、 h は音源位置からマイクロフォン位置まで両点間の部屋のインパルス応答であり、 $*$ は畳み込み演算を表し、 $n(t)$ は残響環境における他の加法性雑音を表す。

【0026】

50

図2に示す実際の部屋のインパルス応答は、直接ピーク hd と、早期反射 he と、後期反射 hl と3つの部分に分けられる。 hd と $s(t)$ との畠み込みは音源より発信した信号は一定の遅延を経てからマイクロフォン端においての再現であり、 $x(t)$ における直接音と対応していると簡単に考えてもよい。早期反射部分のインパルス応答が hd と後の一定持続時間の部分と対応しており、該一定持続時間の終止時間点は50ms～80msの範囲内のある時間点である。一般的に、この部分と $s(t)$ との畠み込みにより生じた早期反射音は直接音の音質を強調及び改善する働きがあると見なされる。後期反射音部分のインパルス応答は、 hd 及び he を除去した後に部屋のインパルス応答が残った長いテーリング部分であり、この部分と $s(t)$ との畠み込みにより生じた反射音は聴覚に影響をもたらす残響成分となる。残響除去算法は主にこの部分の影響を除去するのである。

10

【0027】

従って、残響モデルは次のように表してもよい。

$$x(t) = (hd + he) * s(t) + hl * s(t) + n(t)$$

ここで、 hl 部分が指數減衰モデルの要件を満たしており、以下の数式で近似算出する。

$$hl(t) = b(t)e^{-\frac{3\ln 10}{T_r}t}$$

20

そのうち、 T_r は残響環境の残響時間(RT60)で、 $b(t)$ は零均値ガウス分布ランダム可変量である。

【0028】

以下、後期反射音のパワースペクトルを如何に推定するかを詳しく説明する。

パワースペクトル分析の角度から考えると、信号のパワースペクトル $X(t, f)$ は下記のように表す。

$$X(t, f) = Y(t, f) + R(t, f)$$

そのうち、 $R(t, f)$ は後期反射音のパワースペクトルであるが、 $Y(t, f)$ は直接音及び早期反射音のパワースペクトルであるため、保留される。後期反射音のパワースペクトル $R(t, f)$ を推定した後、スペクトル減算法を用いて $X(t, f)$ から $Y(t, f)$ を推定して、残響除去を実現する。

30

【0029】

残響発生モデルに基づいて分析すると、後期反射音のパワースペクトルはその前の信号パワースペクトルにおけるある成分と線形関係となるが、直接音及び早期反射音のパワースペクトルは人間の音声特性により、丁度、過去の信号パワースペクトルにおけるある成分と線形関係を構成していない。従って、現在フレームの前の、特定した持続時間でのフレームのパワースペクトルに対して線形重畠加算を行うことで、現在フレームの後期反射音のパワースペクトルを推定することができる。そして、スペクトル減算法を介してパワースペクトルから後期反射音のパワースペクトルを除去して、シングルチャンネル音声残響除去を実現できる。

40

【0030】

後期反射音の減衰特性に基づいて前記持続時間範囲の上限値を設置することが好ましい。

スペクトル推定に所用のフレームが多ければ多いほど、推定はより正確になるが、フレームが多すぎると、演算の量が増えてしまう。図2及び hl 部分の指數減衰モデルから分かるように、現在フレームより遠ければ遠いほど離れると、反射音のエネルギーが小さくなり、ある時刻になった後の反射音のエネルギーは見落とされてもよい。従って、後期反射音の減衰特性に基づいて該反射音のパワースペクトルが見落とされる時刻を獲得して、該時刻

50

から現在フレーム時刻までの持続時間を上限値として設定する。これにより、推定された後期反射音のパワースペクトルの正確性を確保できるとともに、重畠加算の量を減らすこともできる。

【0031】

音声関連特性及び直接音と早期反射音の残響環境下でのインパルス応答分布領域に基づいて、前記持続時間範囲の下限値を設定することが好ましい。

図2から分かるように、直接音及び早期反射音のエネルギーは、現在フレームに近づく時間内に集中している。直接音と早期反射音の残響環境下でのインパルス応答分布領域に基づいて、下限値を設定することで、線形重畠加算のとき、直接音及び早期反射音のエネルギーが集中している時間帯を避けて、残響除去の同時に有用な直接音及び早期反射音をより効果的に保留でき、音声品質を向上することができる。

10

【0032】

前記持続時間範囲の下限値は50ms～80msの範囲内の値を選択することが好ましい。

実験によれば、各環境において、下限値を50ms～80msの範囲内の数値とすることが確保できれば、直接音及び早期反射音部分を効果的に迂回して、有効な後期反射音のパワースペクトルをより良く推定することができる。環境の変化が発生した後、下限値を調整しなくても、より良い音声品質を獲得することができる。

20

【0033】

前記持続時間範囲の上限値は0.3s～0.5sの範囲内の数値を選択することが好ましい。

理論上、上限値の設定は、方法の適用の具体的な環境と関係している。本発明に係る後期反射音のパワースペクトル推定において、上限値は理論上部屋のインパルス応答の長さと対応するが、残響発生モデル及び真実な環境のインパルス応答hI部分が指數モデルに基づき減衰するため、現在時刻から遠ければ遠いほど、反射音のエネルギーがより小さくなり、0.5sを超えると、反射音のエネルギーがほぼ見落とされて計上しなくてよい。従って、実際に粗略な上限値さえ用いれば、ほとんどの残響環境に適用することができる。検証したところ、上限値は0.3s～0.5sの範囲内の数値とされるとき、消音室（残響時間が非常に短い）、普通のオフィス部屋環境（残響時間が0.3s～0.5s）乃至大ホール（残響時間>1s）のような多種の残響環境にいずれも優れた適応性を有している。消音室環境の下で、後期反射音がほぼない。

30

【0034】

本発明の方法は線形成分しか推定せず、且つ直接音及び早期反射音のエネルギーが集中している時間帯を避けているため、上限値の設定値が消音室の残響時間よりも長かったとしても、有効な音声成分は除去されることがない。一方、大ホール環境において、上限値の設定値が真実な残響時間よりも小さくなるが、インパルス応答は指數に基づき非常に速く減衰しており、前の0.3s以内の後期反射音成分が後期反射音成分全体のほとんどのエネルギーを占めているため、残響をより効果的に除去することができる。

40

【0035】

具体的な実施形態において、前記これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理は、具体的に、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理を含む。

【0036】

例えば、下記の式でARモデルを用いて現在フレームの後期反射音のパワースペクトルを推定する。

$$R(t, f) = \sum_{j=J_0}^{J_{AR}} \alpha_{j,f} \cdot X(t - j \cdot \Delta t, f)$$

50

そのうち、 $R(t, f)$ は推定された後期反射音のパワースペクトルであり、 J_0 は設定された持続時間範囲内の下限値から得た初期次数であり、 J_{AR} は設定された持続時間範囲内の上限値から得たARモデルの次数であり、 $\alpha_{j,f}$ はARモデル推定パラメータであり、

$$X(t - j \cdot \Delta t, f)$$

は現在フレームより j フレーム分前のフレームのパワースペクトルであり、 t はフレームの間隔である。

【0037】

具体的な実施形態において、前記これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理は、具体的に、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理を含む。

【0038】

例えば、下記の数式で移動平均MAモデルを用いて現在フレームの後期反射音のパワースペクトルを推定する。

$$R(t, f) = \sum_{j=J_0}^{J_{MA}} \beta_{j,f} \cdot Y(t - j \cdot \Delta t, f)$$

そのうち、 $R(t, f)$ は推定された後期反射音のパワースペクトルであり、 J_0 は設定された持続時間範囲内の下限値から得た初期次数であり、 J_{MA} は設定された持続時間範囲内の上限値から得たMAモデルの次数であり、 $\beta_{j,f}$ はMAモデル推定パラメータであり、

$$Y(t - j \cdot \Delta t, f)$$

は現在フレームより j フレーム分前のフレームの直接音及び早期反射音のパワースペクトルであり、 t はフレームの間隔である。

【0039】

具体的な実施形態において、前記これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理は、具体的に、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算するとともに、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理を含む。

【0040】

例えば、下記の数式でARMAモデルを用いて現在フレームの後期反射音のパワースペクトルを推定する。

$$R(t, f) = \sum_{j=J_0}^{J_{AR}} \alpha_{j,f} \cdot X(t - j \cdot \Delta t, f) + \sum_{j=J_0}^{J_{MA}} \beta_{j,f} \cdot Y(t - j \cdot \Delta t, f)$$

そのうち、 $R(t, f)$ は推定された後期反射音のパワースペクトルであり、 J_0 は設定された持続時間範囲内の下限値から得た初期次数であり、 J_{AR} は設定された持続時間範囲内の上限値から得たARモデルの次数であり、 $\alpha_{j,f}$ はARモデル推定パラメータであり、 J_{MA} は設定された持続時間範囲内の上限値から得たMAモデルの次数であり、 $\beta_{j,f}$ はMAモデル推定パラメータであり、

10

20

30

40

50

$Y(t-j\cdot\Delta t, f)$

は現在フレームより j フレーム分前のフレームの直接音及び早期反射音のパワースペクトルであり、

$X(t-j\cdot\Delta t, f)$

は現在フレームより j フレーム分前のフレームのパワースペクトルであり、 t はフレームの間隔である。

10

【0041】

ARモデル、MAモデル、ARMAモデルの解を具体的に求めるにあたって、先行技術に周知算法が存在しており、例えば、Yule-Walker方程式で解を求め、或は、Burgアルゴリズムを用いる。

【0042】

スペクトル減算法を用いた残響除去において、後期反射音のパワースペクトルを推定することが最も重要である。先行技術に取り上げられた後期反射音のパワースペクトル推定は、そもそも上述したAR又はMA又はARMAモデルの1種の特例であると考えられ、また、その他の後期反射音のパワースペクトル推定方法は音声間歇段階において残響環境の残響時間（RT60）を推定し、後期反射音のパワースペクトル推定における重要なパラメータとするのは多い。本発明において、残響時間又は各種の環境に対するインパルス応答を推定する必要がないので、多種の異なる残響環境、及び、発話者が残響環境において移動することなどによる残響インパルス応答又は残響時間が変わった状況に適応することができる。

20

【0043】

具体的な実施形態において、スペクトル減算法を用いて前記フレームのパワースペクトルから残響成分を除去する処理は、具体的に、後期反射音のパワースペクトルに基づいて、スペクトル減算法によって利得函数を求めてきて、利得函数を現在フレームのパワースペクトルと乗算して現在フレームの直接音及び早期反射音のパワースペクトルを得る。

【0044】

後期反射音のパワースペクトル $R(t, f)$ を推定終了後、残響除去された音声信号 $Y(t, f)$ はスペクトル減算法によって得られる。

30

$$Y(t, f) = G(t, f) \cdot X(t, f)$$

そのうち、

$$G(t, f) = \frac{X(t, f) - R(t, f)}{X(t, f)}$$

はスペクトル減算法により得たGain（利得）函数である。

40

【0045】

本発明の実施効果は図3に示す。残響信号（シングルチャンネル音声信号）が会議室から採取され、音源とマイクロフォンとの距離が2mとされ、残響時間が約0.45sとなる。本発明に取り上げられたARモデルに基づき、後期反射音のパワースペクトルを推定し、下限値を80msと設定し、上限値を0.5sと設定する。図示から分かるように、本発明の方法を用いて残響除去した後、残響テーリングが明らかに減衰し、音声の品質が顕著に向上了。

【0046】

図4に示すように、本発明によるシングルチャンネル音声残響除去装置は、以下のユニットを含む。

50

【0047】

フレーム分割ユニット100は、入力されたシングルチャンネル音声信号に対してフレーム分割を行い、時間の順に応じてフレーム信号をフーリエ変換ユニット200に出力するためには用いられる。

【0048】

フーリエ変換ユニット200は、現在フレームに対して短時間フーリエ変換を行い、現在フレームのパワースペクトル及び位相スペクトルを獲得して、スペクトル減算ユニット400及びスペクトル推定ユニット300に現在フレームのパワースペクトルを出力し、逆フーリエ変換ユニット500に位相スペクトルを出力するために用いられる。

【0049】

スペクトル推定ユニット300は、現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内である数フレームを選んで、これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定してから、スペクトル減算ユニット400に出力するために用いられる。

10

【0050】

スペクトル減算ユニット400は、スペクトル減算法によって、フーリエ変換ユニット200より得た現在フレームのパワースペクトルから、スペクトル推定ユニット300より得た現在フレームの後期反射音のパワースペクトルを除去して、現在フレームの直接音及び早期反射音のパワースペクトルを獲得してから、逆フーリエ変換ユニット500に出力するために用いられる。

20

【0051】

逆フーリエ変換ユニット500は、スペクトル減算ユニット400より得た現在フレームの直接音及び早期反射音のパワースペクトルを、フーリエ変換ユニット200より得た現在フレームの位相スペクトルとともに、短時間逆フーリエ変換を行い、現在フレーム残響除去後の信号を出力するために用いられる。

【0052】

前記スペクトル推定ユニット300は、具体的に、後期反射音の減衰特性に基づいて前記持続時間範囲の上限値を設置するために用いられることが好ましい。

スペクトル推定ユニット300は、具体的に、音声関連特性及び直接音と早期反射音の残響環境下でのインパルス応答分布領域に基づいて、前記持続時間範囲の下限値を設定するために用いられることが好ましい。

30

スペクトル推定ユニット300は、具体的に、持続時間範囲の上限値は0.3s～0.5sの範囲内の数値を選択するために用いられることが好ましい。

スペクトル推定ユニット300は、具体的に、持続時間範囲の下限値は50ms～80msの範囲内の値を選択するために用いられることが好ましい。

【0053】

具体的な実施形態の装置は、図5に示すように、前記スペクトル推定ユニット300が、具体的に、現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内である数フレームに対して、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定するために用いられる。

40

【0054】

例えば、下記の式でARモデルを用いて現在フレームの後期反射音のパワースペクトルを推定する。

$$R(t, f) = \sum_{j=J_0}^{J_{AR}} \alpha_{j,f} \cdot X(t - j \cdot \Delta t, f)$$

そのうち、R(t, f)は推定された後期反射音のパワースペクトルであり、J₀は設定された

50

持続時間範囲内の下限値から得た初期次数であり、 J_{AR} は設定された持続時間範囲内の上限値から得たARモデルの次数であり、 $\alpha_{j,f}$ はARモデル推定パラメータであり、

$$X(t-j \cdot \Delta t, f)$$

は現在フレームより j フレーム分前のフレームのパワースペクトルであり、 t はフレームの間隔である。

【0055】

もう1つの具体的な実施形態において、前記スペクトル推定ユニット300は、具体的に、現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内にある数フレームに対して、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定するために用いられる。

【0056】

例えば、下記の数式で移動平均MAモデルを用いて現在フレームの後期反射音のパワースペクトルを推定する。

$$R(t, f) = \sum_{j=J_0}^{J_{MA}} \beta_{j,f} \cdot Y(t-j \cdot \Delta t, f)$$

そのうち、 $R(t, f)$ は推定された後期反射音のパワースペクトルであり、 J_0 は設定された下限値から得た初期次数であり、 J_{MA} は設定された上限値から得たMAモデルの次数であり、 $\beta_{j,f}$ はMAモデル推定パラメータであり、

$$Y(t-j \cdot \Delta t, f)$$

は現在フレームより j フレーム分前のフレームの直接音及び早期反射音のパワースペクトルであり、 t はフレームの間隔である。

【0057】

もう1つの具体的な実施形態において、前記スペクトル推定ユニット300は、具体的に、現在フレームの前の、現在フレームに至るまでの距離が設定された持続時間範囲内にある数フレームに対して、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算するとともに、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定するために用いられる。

【0058】

例えば、下記の数式でARMAモデルを用いて現在フレームの後期反射音のパワースペクトルを推定する。

$$R(t, f) = \sum_{j=J_0}^{J_{AR}} \alpha_{j,f} \cdot X(t-j \cdot \Delta t, f) + \sum_{j=J_0}^{J_{MA}} \beta_{j,f} \cdot Y(t-j \cdot \Delta t, f)$$

そのうち、 $R(t, f)$ は推定された後期反射音のパワースペクトルであり、 J_0 は設定された下限値から得た初期次数であり、 J_{AR} は設定された上限値から得たARモデルの次数であり、 $\alpha_{j,f}$ はARモデル推定パラメータであり、 J_{MA} は設定された上限値から得たMAモデルの次数であり、 $\beta_{j,f}$ はMAモデル推定パラメータであり、

$$Y(t-j \cdot \Delta t, f)$$

10

20

30

40

50

は現在フレームより j フレーム分前のフレームの直接音及び早期反射音のパワースペクトルであり、

$$X(t - j \cdot \Delta t, f)$$

は現在フレームより j フレーム分前のフレームのパワースペクトルであり、 t はフレームの間隔である。

【 0 0 5 9 】

ARモデル、MAモデル、ARMAモデルの解を具体的に求めるにあたって、先行技術に周知算法が存在しており、例えば、Yule-Walker方程式で解を求め、或は、Burgアルゴリズムを用いる。 10

【 0 0 6 0 】

前記スペクトル減算ユニット400は、具体的に、後期反射音のパワースペクトルに基づいて、スペクトル減算法によって利得函数を求めてきて、利得函数を現在フレームのパワースペクトルと乗算して現在フレームの直接音及び早期反射音のパワースペクトルを得るために用いられる。

【 0 0 6 1 】

後期反射音のパワースペクトル $R(t, f)$ を推定終了後、残響除去された音声信号 $Y(t, f)$ はスペクトル減算法によって得られる。 20

$$Y(t, f) = G(t, f) \cdot X(t, f)$$

そのうち、

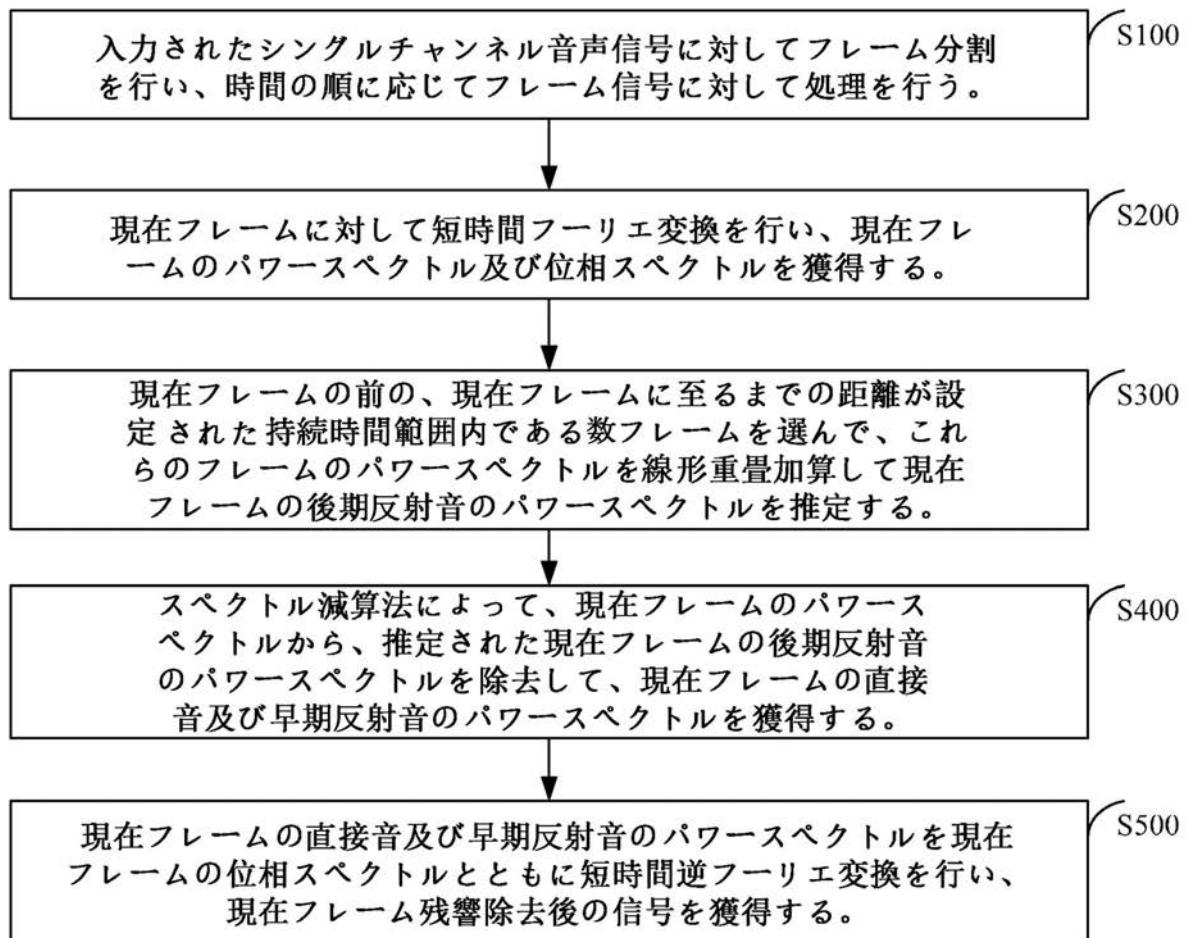
$$G(t, f) = \frac{X(t, f) - R(t, f)}{X(t, f)}$$

はスペクトル減算法により得たGain(利得)函数である。

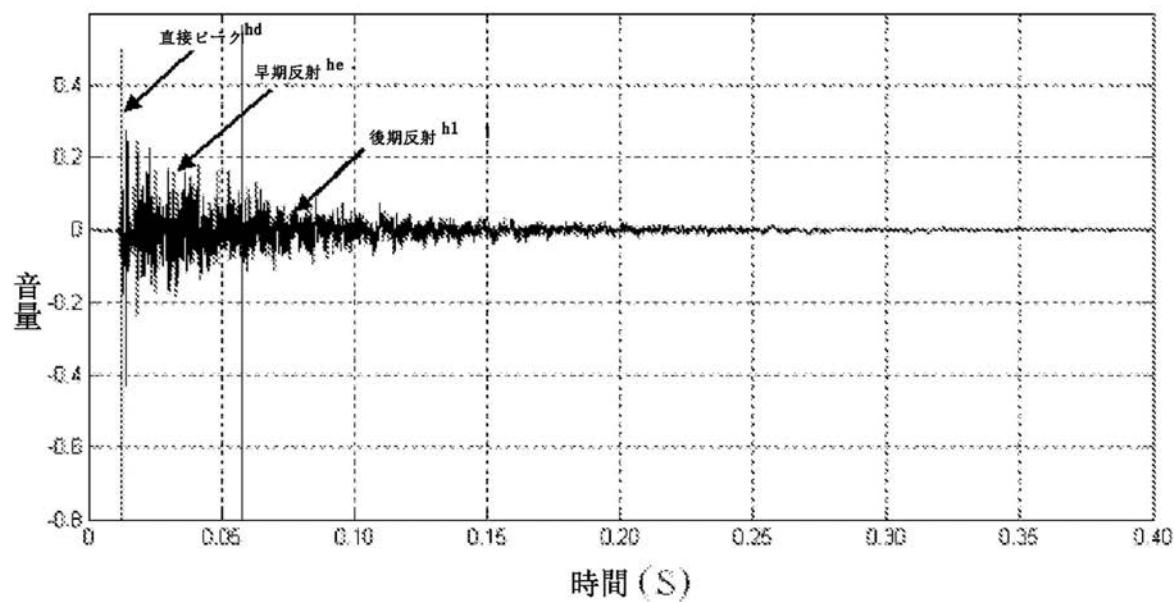
【 0 0 6 2 】

上述したのは、あくまでも本発明の好ましい実施例であり、本発明の保護範囲を限定するためのものではない。本発明の精神及び原則内になされたあらゆる変更、均等置換、改良等は、いずれも本発明の保護範囲内に含まれるものとする。 30

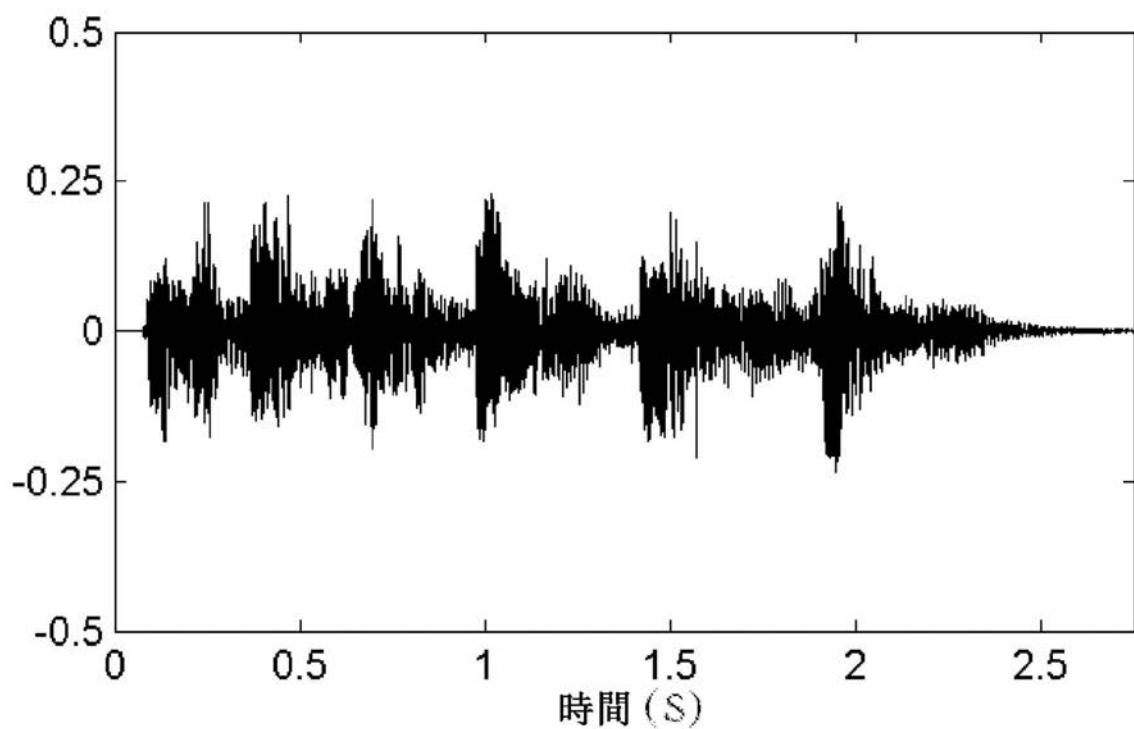
【図1】



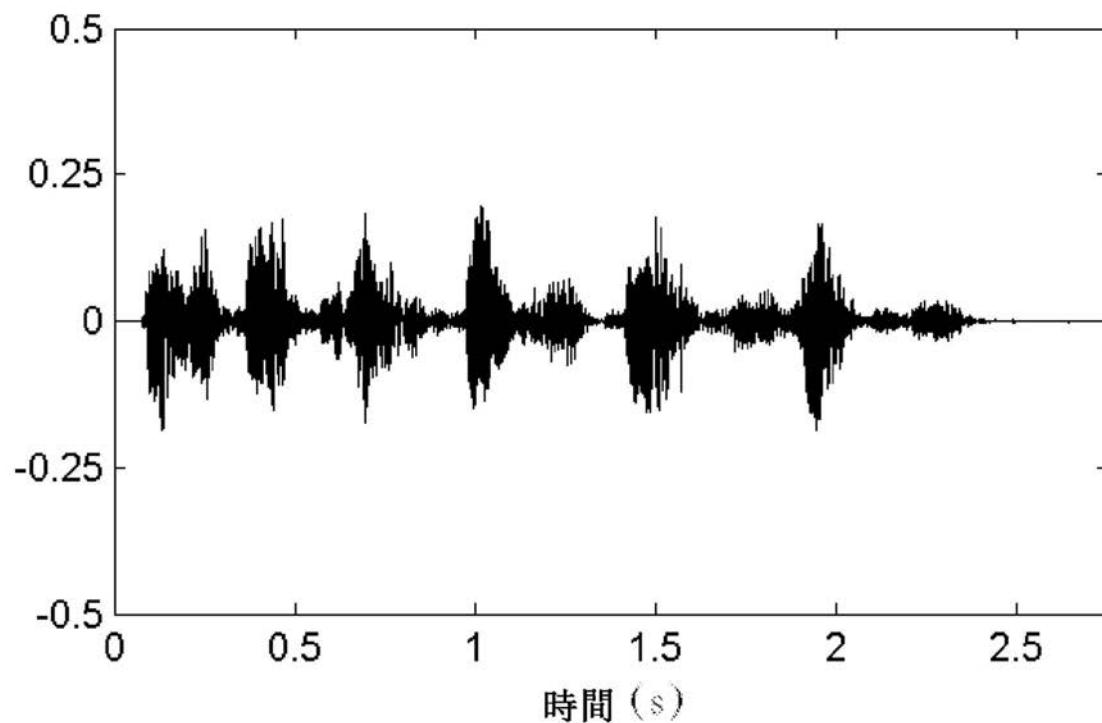
【図2】



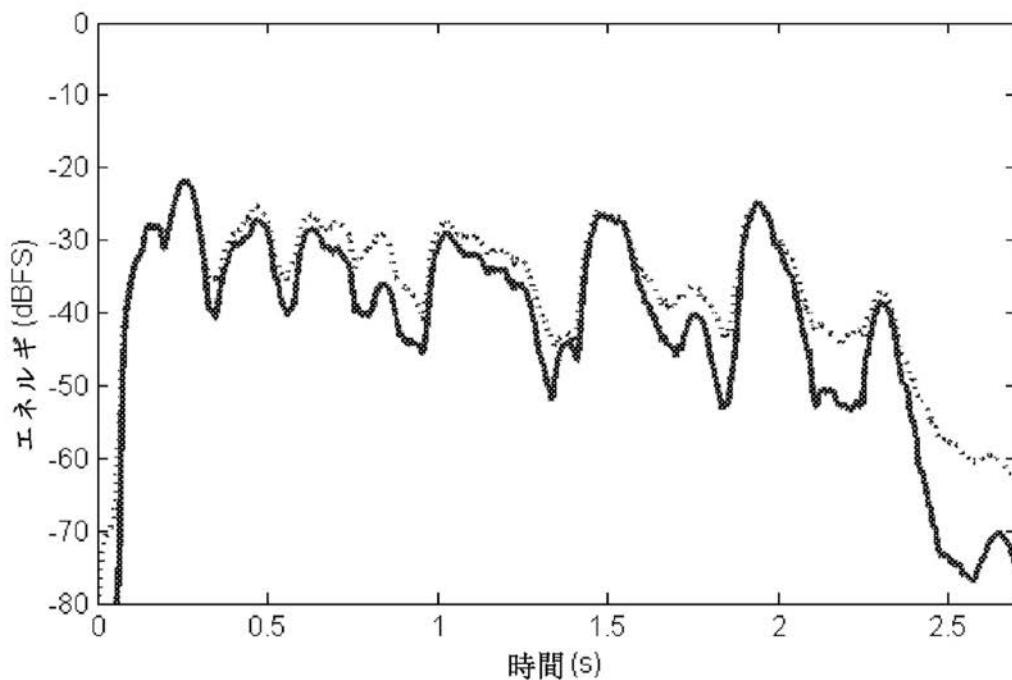
【図3A】



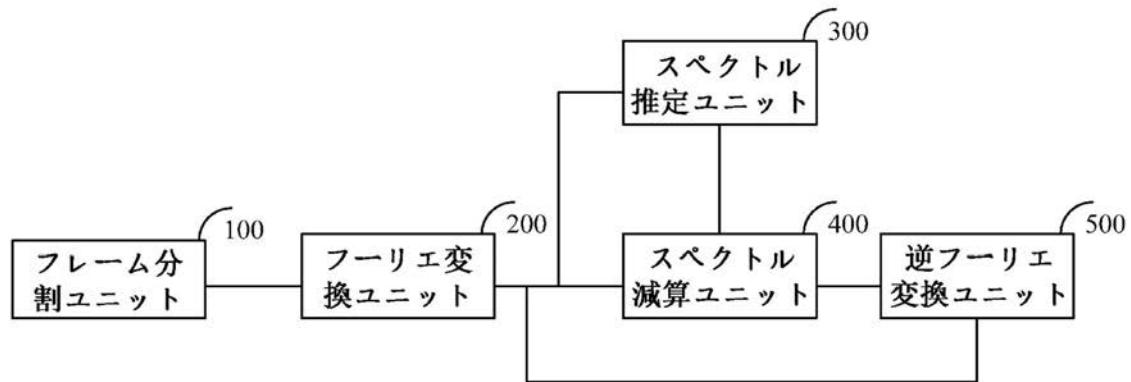
【図 3 B】



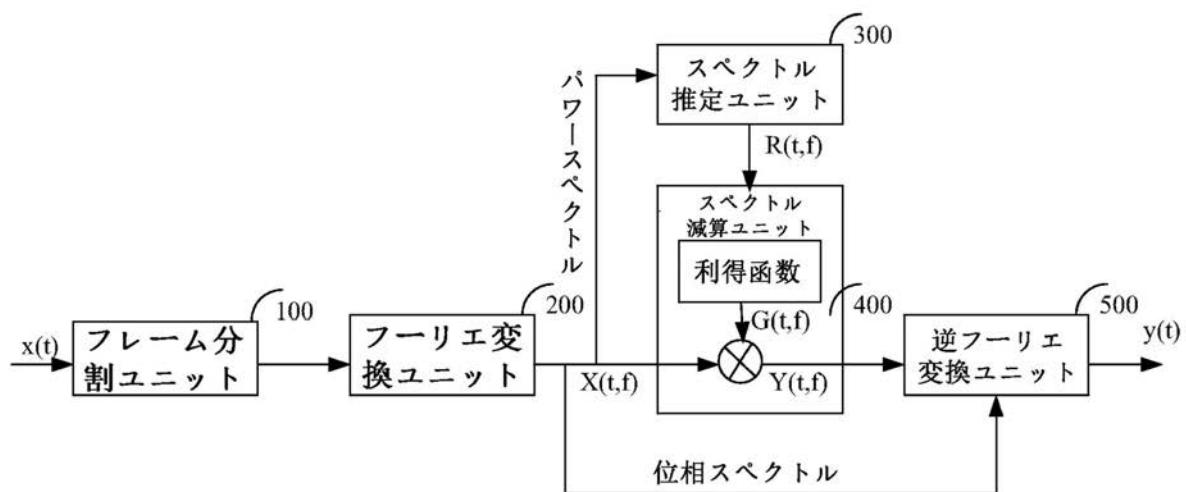
【図 3 C】



【図4】



【図5】



【手続補正書】

【提出日】平成27年1月30日(2015.1.30)

【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】

入力されたシングルチャンネル音声信号に対してフレーム分割を行い、時間の順に応じてフレーム信号に対して、

現在フレームに対して短時間フーリエ変換を行い、現在フレームのパワースペクトル及び位相スペクトルを獲得する処理と、

現在フレームの前の、現在フレームに至るまでの距離が設定の持続時間範囲内である数フレームを選んで、これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理と、

スペクトル減算法によって、現在フレームのパワースペクトルから、推定された現在フレームの後期反射音のパワースペクトルを除去して、現在フレームの直接音及び早期反射音のパワースペクトルを獲得する処理と、

現在フレームの直接音及び早期反射音のパワースペクトルを現在フレームの位相スペクトルとともに、短時間逆フーリエ変換を行い、現在フレーム残響除去後の信号を獲得する処理と、

を行うことを含むことを特徴とするシングルチャンネル音声残響除去方法。

【請求項 2】

前記後期反射音の減衰特性に基づいて前記持続時間範囲の上限値が設定され、

及び/又は、音声関連特性及び直接音と早期反射音の残響環境下でのインパルス応答分布領域に基づいて、前記持続時間範囲の下限値が設定されることを特徴とする請求項1に記載のシングルチャンネル音声残響除去方法。

【請求項 3】

前記持続時間範囲の上限値が0.3s～0.5sの範囲内の値を選択することを特徴とする請求項1に記載のシングルチャンネル音声残響除去方法。

【請求項 4】

前記持続時間範囲の下限値が50ms～80msの範囲内の値を選択することを特徴とする請求項1に記載のシングルチャンネル音声残響除去方法。

【請求項 5】

前記の、これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理は、具体的に、

自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理、

或は、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理、

或は、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算するとともに、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理、

を行うことを特徴とする請求項1に記載のシングルチャンネル音声残響除去方法。

【請求項 6】

入力されたシングルチャンネル音声信号に対してフレーム分割を行い、時間の順に応じてフレーム信号をフーリエ変換ユニットへ出力するためのフレーム分割ユニットと、

受信された現在フレームに対して短時間フーリエ変換処理を行い、現在フレームのパワースペクトル及び位相スペクトルを獲得して、現在フレームのパワースペクトルをスペクトル減算ユニットとスペクトル推定ユニットへ出力し、位相スペクトルを逆フーリエ変換ユニットへ出力するためのフーリエ変換ユニットと、

現在フレームの前の、現在フレームに至るまでの距離が設定の持続時間範囲内である数フレームのパワースペクトルを線形重畠加算して、現在フレームの後期反射音のパワースペクトルを推定し、推定された現在フレームの後期反射音のパワースペクトルをスペクトル減算ユニットへ出力するためのスペクトル推定ユニットと、

スペクトル減算法によって、フーリエ変換ユニットより獲得した現在フレームのパワースペクトルから、スペクトル推定ユニットより獲得した現在フレームの後期反射音のパワースペクトルを除去して、現在フレームの直接音及び早期反射音のパワースペクトルを獲得し、現在フレームの直接音及び早期反射音のパワースペクトルを逆フーリエ変換ユニットへ出力するためのスペクトル減算ユニットと、

スペクトル減算ユニットより獲得した現在フレームの直接音及び早期反射音のパワースペクトルを、フーリエ変換ユニットより獲得した現在フレームの位相スペクトルとともに、短時間逆フーリエ変換を行い、現在フレーム残響除去後の信号を出力するための逆フーリエ変換ユニットと、

を含むことを特徴とするシングルチャンネル音声残響除去装置。

【請求項 7】

前記スペクトル推定ユニットは、具体的に、後期反射音の減衰特性に基づいて前記持続時間範囲の上限値を設定し、及び/又は、音声関連特性及び直接音と早期反射音の残響環境下でのインパルス応答分布領域に基づいて、前記持続時間範囲の下限値を設定するために用いられるることを特徴とする請求項 6 に記載のシングルチャンネル音声残響除去装置。

【請求項 8】

前記スペクトル推定ユニットは、具体的に、持続時間範囲の上限値が 0 . 3 s ~ 0 . 5 s の範囲内の値を選択するために用いられるることを特徴とする請求項 6 に記載のシングルチャンネル音声残響除去装置。

【請求項 9】

前記スペクトル推定ユニットは、具体的に、持続時間範囲の下限値が 5 0 ms ~ 8 0 ms の範囲内の値を選択するために用いられるることを特徴とする請求項 6 に記載のシングルチャンネル音声残響除去装置。

【請求項 10】

前記スペクトル推定ユニットは、具体的に、
現在フレームの前の、現在フレームに至るまでの距離が前記設定の持続時間範囲内である数フレームに対して、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定するか、
或は、現在フレームの前の、現在フレームに至るまでの距離が前記設定の持続時間範囲内である数フレームに対して、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定するか、

或は、現在フレームの前の、現在フレームに至るまでの距離が前記設定の持続時間範囲内である数フレームに対して、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算するとともに、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する、
ために用いられるることを特徴とする請求項 6 に記載のシングルチャンネル音声残響除去装置。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0 0 0 7

【補正方法】変更

【補正の内容】

【0 0 0 7】

本発明は、シングルチャンネル音声残響除去方法を開示しており、
入力されたシングルチャンネル音声信号に対してフレーム分割を行い、時間の順に応じてフレーム信号に対して、

現在フレームに対して短時間フーリエ変換を行い、現在フレームのパワースペクトル及

び位相スペクトルを獲得する処理と、

現在フレームの前の、現在フレームに至るまでの距離が設定の持続時間範囲内である数フレームを選んで、これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理と、

スペクトル減算法によって、現在フレームのパワースペクトルから、推定された現在フレームの後期反射音のパワースペクトルを除去して、現在フレームの直接音及び早期反射音のパワースペクトルを獲得する処理と、

現在フレームの直接音及び早期反射音のパワースペクトルを現在フレームの位相スペクトルとともに、短時間逆フーリエ変換を行い、現在フレーム残響除去後の信号を獲得する処理と、を行うことを含む。

【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】0011

【補正方法】変更

【補正の内容】

【0011】

前記の、これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理は、具体的に、

自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理、

或は、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理、

或は、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算するとともに、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する処理、を含むことが好ましい。

【手続補正4】

【補正対象書類名】明細書

【補正対象項目名】0012

【補正方法】変更

【補正の内容】

【0012】

本発明は、また、シングルチャンネル音声残響除去装置を開示しており、

入力されたシングルチャンネル音声信号に対してフレーム分割を行い、時間の順に応じてフレーム信号をフーリエ変換ユニットへ出力するためのフレーム分割ユニットと、

受信された現在フレームに対して短時間フーリエ変換処理を行い、現在フレームのパワースペクトル及び位相スペクトルを獲得して、現在フレームのパワースペクトルをスペクトル減算ユニットとスペクトル推定ユニットへ出力し、位相スペクトルを逆フーリエ変換ユニットへ出力するためのフーリエ変換ユニットと、

現在フレームの前の、現在フレームに至るまでの距離が設定の持続時間範囲内である数フレームのパワースペクトルを線形重畠加算して、現在フレームの後期反射音のパワースペクトルを推定し、そして推定された現在フレームの後期反射音のパワースペクトルをスペクトル減算ユニットへ出力するためのスペクトル推定ユニットと、

スペクトル減算法によって、フーリエ変換ユニットより獲得した現在フレームのパワースペクトルから、スペクトル推定ユニットより獲得した現在フレームの後期反射音のパワースペクトルを除去して、現在フレームの直接音及び早期反射音のパワースペクトルを獲得し、現在フレームの直接音及び早期反射音のパワースペクトルを逆フーリエ変換ユニットへ出力するためのスペクトル減算ユニットと、

スペクトル減算ユニットより獲得した現在フレームの直接音及び早期反射音のパワース

ペクトルを、フーリエ変換ユニットより獲得した現在フレームの位相スペクトルとともに、短時間逆フーリエ変換を行い、現在フレーム残響除去後の信号を出力するための逆フーリエ変換ユニットと、を含む。

【手続補正5】

【補正対象書類名】明細書

【補正対象項目名】0013

【補正方法】変更

【補正の内容】

【0013】

前記スペクトル推定ユニットは、具体的に、後期反射音の減衰特性に基づいて前記持続時間範囲の上限値を設定し、及び/又は、音声関連特性及び直接音と早期反射音の残響環境下でのインパルス応答分布領域に基づいて、前記持続時間範囲の下限値を設定するために用いられることが好ましい。

【手続補正6】

【補正対象書類名】明細書

【補正対象項目名】0014

【補正方法】変更

【補正の内容】

【0014】

前記スペクトル推定ユニットは、具体的に、持続時間範囲の上限値が0.3s～0.5sの範囲内の値を選択するために用いられることが好ましい。

【手続補正7】

【補正対象書類名】明細書

【補正対象項目名】0015

【補正方法】変更

【補正の内容】

【0015】

前記スペクトル推定ユニットは、具体的に、持続時間範囲の下限値が50ms～80msの範囲内の値を選択するために用いられることが好ましい。

【手続補正8】

【補正対象書類名】明細書

【補正対象項目名】0016

【補正方法】変更

【補正の内容】

【0016】

前記スペクトル推定ユニットは、具体的に、

現在フレームの前の、現在フレームに至るまでの距離が設定の持続時間範囲内である数フレームに対して、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定するか、

或は、現在フレームの前の、現在フレームに至るまでの距離が設定の持続時間範囲内である数フレームに対して、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定するか、

或は、現在フレームの前の、現在フレームに至るまでの距離が設定の持続時間範囲内である数フレームに対して、自己回帰ARモデルを用いて、これらのフレームのパワースペクトルにおける全ての成分を線形重畠加算するとともに、移動平均MAモデルを用いて、これらのフレームのパワースペクトルにおける直接音及び早期反射音の成分を線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定する、

ために用いられることが好ましい。

【手続補正9】**【補正対象書類名】明細書****【補正対象項目名】0017****【補正方法】変更****【補正の内容】****【0017】**

本発明の実施例の有益な効果は、現在フレームの前の、現在フレームに至るまでの距離が設定の持続時間範囲内である数フレームを選べることによって、これらのフレームのパワースペクトルを線形重畠加算して現在フレームの後期反射音のパワースペクトルを推定することで、残響環境の伝達関数又は残響時間を推定する必要がなく、現在フレームの後期反射音のパワースペクトルを推定することができ、そしてスペクトル減算法を用いて残響除去ができるため、残響除去の操作を簡略化して、より簡単に実現できることと、

音声関連特性及び直接音と早期反射音の残響環境下でのインパルス応答分布領域に基づいて、持続時間範囲の下限値を設定することによって、残響除去で同時に有用な直接音及び早期反射音を保留して、音声の品質を向上できることと、

後期反射音の減衰特性に基づいて持続時間範囲の上限値を設定することによって、推定された後期反射音のパワースペクトルの正確性を確保できると同時に重畠加算量を減少できることと、

本発明の実施例が上限値を0.3s～0.5sの範囲内の値と選択しており、該上限値が実験により得たしきい値であり、残響環境が変化するとき、該上限値を調整しなくても、より良い残響除去の効果を得られることと、

本発明の実施例が下限値を50ms～80msの範囲内の値と選択しており、残響環境が変化するとき、該下限値を変えなくても、効果的に直接音及び早期反射音を避けて重畠加算を行うことができ、重畠加算の結果に直接音及び早期反射音がほぼ含まれず、残響除去で同時に有用な直接音及び早期反射音を保留し、より良い音声品質を得られることにある。

前記残響環境の変化は、残響がない消音室から残響が極めてひどい大ホールにまでの変化を含む。

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/CN2013/073584																		
A. CLASSIFICATION OF SUBJECT MATTER G10L 21/0208 (2013.01) i According to International Patent Classification (IPC) or to both national classification and IPC																				
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC: G10L; G10; H04B; H04M; H04R																				
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched																				
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) CNABS, CNKI, CNTXT, VEN, USTXT, EPTXT, WOTXT, TWTXT: reverberat+, reverberant+, dereverberat+, de, cancel+, remov+, restrain+, suppress+, reduc+, lessen+, decreas+, attenuat+, eliminat+, echo??, late, latter, reflect+																				
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">Category*</th> <th style="text-align: left;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="text-align: left;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 102750956 A (GOERTEK INC.) 24 October 2012 (24.10.2012) claims 1 to 10, description, paragraphs [0007] to [0127] and figures 1 to 5</td> <td>1-10</td> </tr> <tr> <td>X</td> <td>CN 101385386 A (NIPPON TELEGRAPH & TELEPHONE CORP) 11 March 2009 (11.03.2009) description, page 8 to page 42 and figures 1-23C</td> <td>1-10</td> </tr> <tr> <td>X</td> <td>US 8160262 B2 (NUANCE COMMUNICATIONS INC.) 17 April 2012 (17.04.2012) description, column 3, line 42 to column 13, line 41 and figures 3 to 7</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>US 2008292108 A1 (BUCK M. et al.) 27 November 2008 (27.11.2008) the whole document</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>US 2008059157 A1 (FUKUDA T. et al.) 06 March 2008 (06.03.2008) the whole document</td> <td>1-10</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	PX	CN 102750956 A (GOERTEK INC.) 24 October 2012 (24.10.2012) claims 1 to 10, description, paragraphs [0007] to [0127] and figures 1 to 5	1-10	X	CN 101385386 A (NIPPON TELEGRAPH & TELEPHONE CORP) 11 March 2009 (11.03.2009) description, page 8 to page 42 and figures 1-23C	1-10	X	US 8160262 B2 (NUANCE COMMUNICATIONS INC.) 17 April 2012 (17.04.2012) description, column 3, line 42 to column 13, line 41 and figures 3 to 7	1-10	A	US 2008292108 A1 (BUCK M. et al.) 27 November 2008 (27.11.2008) the whole document	1-10	A	US 2008059157 A1 (FUKUDA T. et al.) 06 March 2008 (06.03.2008) the whole document	1-10
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.																		
PX	CN 102750956 A (GOERTEK INC.) 24 October 2012 (24.10.2012) claims 1 to 10, description, paragraphs [0007] to [0127] and figures 1 to 5	1-10																		
X	CN 101385386 A (NIPPON TELEGRAPH & TELEPHONE CORP) 11 March 2009 (11.03.2009) description, page 8 to page 42 and figures 1-23C	1-10																		
X	US 8160262 B2 (NUANCE COMMUNICATIONS INC.) 17 April 2012 (17.04.2012) description, column 3, line 42 to column 13, line 41 and figures 3 to 7	1-10																		
A	US 2008292108 A1 (BUCK M. et al.) 27 November 2008 (27.11.2008) the whole document	1-10																		
A	US 2008059157 A1 (FUKUDA T. et al.) 06 March 2008 (06.03.2008) the whole document	1-10																		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.																				
<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&" document member of the same patent family</p>																				
Date of the actual completion of the international search 28 June 2013 (28.06.2013)		Date of mailing of the international search report 18 July 2013 (18.07.2013)																		
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451		Authorized officer YANG, Shilin Telephone No. (86-10) 62085717																		

INTERNATIONAL SEARCH REPORT		International application No. PCT/CN2013/073584
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 1989550 A (KONINK PHILIPS ELECTRONICS N.V.) 27 June 2007 (27.06.2007) the whole document	1-10
A	CN 101454825 A (HARMAN INTERNATIONAL INDUSTRIES INC.) 10 June 2009 (10.06.2009) the whole document	1-10
A	CN 101315772 A (SHANGHAI JIAOTONG UNIVERSITY) 03 December 2008 (03.12.2008) the whole document	1-10

Form PCT/ISA/210 (continuation of second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2013/073584

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102750956 A	24.10.2012	None	
CN 101385386 A	11.03.2009	WO 2007100137 A1 JPWO 2007100137 SX US 2009248403 A1 JP 4774100 B2 CN 101385386 B US 8271277 B2 EP 1993320 A1	07.09.2007 23.07.2009 01.10.2009 14.09.2011 09.05.2012 18.09.2012 19.11.2008
US 8160262 B2	17.04.2012	EP 2058804 A1 US 2009117948 A1	13.05.2009 07.05.2009
US 2008292108 A1	27.11.2008	EP 1885154 A1	06.02.2008
US 2008059157 A1	06.03.2008	JP 2008058900 A US 7590526 B2 JP 4107613 B2	13.03.2008 15.09.2009 25.06.2008
CN 1989550 A	27.06.2007	KR 20070036777 A US 2008300869 A1 JP 2008507720 A KR 1149591 B1 JP 5042823 B2 US 8116471 B2 WO 2006011104 A1 IN 200700280 P4 CN 1989550 B EP 1774517 A1	03.04.2007 04.12.2008 13.03.2008 29.05.2012 03.10.2012 14.02.2012 02.02.2006 24.08.2007 13.10.2010 18.04.2007
CN 101454825 A	10.06.2009	WO 2008034221 A1 JP 4964943 B2 JP 2009531722 A EP 2064699 A1	27.03.2008 04.07.2012 03.09.2009 03.06.2009
CN 101315772 A	03.12.2008	None	

国际检索报告	国际申请号 PCT/CN2013/073584	
A. 主题的分类		
G10L 21/0208 (2013.01) i 按照国际专利分类(IPC)或者同时按照国家分类和 IPC 两种分类		
B. 检索领域		
检索的最低限度文献(标明分类系统和分类号) IPC: G10L, G10, H04B, H04M, H04R		
包含在检索领域中的除最低限度文献以外的检索文献		
在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) CNABS, CNKI, CNTXT: 混响, 回响, 回声, 反射, 漫射, 后期, 晚期, 后, 除, 去, 消, 减, 抑制 VEN, USTXT, EPTXT, WOTXT, TWTXT: reverberat+, reverberant+, dereverberat+, de, cancel+, remov+, restrain+, suppress+, reduc+, lessen+, decreas+, attenuat+, eliminat+, echo??, late, latter, reflect+		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 102750956 A (歌尔声学股份有限公司) 24.10 月 2012 (24.10.2012), 权利要求 1-10, 说明书第[0007]段-第[0127]段, 附图 1-5	1-10
X	CN 101385386 A (日本电信电话株式会社) 11.3 月 2009 (11.03.2009), 说明书第 8 页-第 42 页, 附图 1-23C	1-10
X	US 8160262 B2 (NUANCE COMMUNICATIONS INC.) 17.4 月 2012 (17.04.2012), 说明书第 3 栏第 42 行-第 13 栏第 41 行, 附图 3-7	1-10
A	US 2008292108 A1 (BUCK M. 等) 27.11 月 2008 (27.11.2008), 全文	1-10
A	US 2008059157 A1 (FUKUDA T. 等) 06.3 月 2008 (06.03.2008), 全文	1-10
A	CN 1989550 A (皇家飞利浦电子股份有限公司) 27.6 月 2007 (27.06.2007), 全文	1-10
<input checked="" type="checkbox"/> 其余文件在 C 栏的续页中列出。		<input checked="" type="checkbox"/> 见同族专利附件。
<p>* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 </p>		<p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件 </p>
国际检索实际完成的日期 28.6 月 2013 (28.06.2013)	国际检索报告邮寄日期 18.7 月 2013 (18.07.2013)	
ISA/CN 的名称和邮寄地址: 中华人民共和国国家知识产权局 中国北京市海淀区蓟门桥西土城路 6 号 100088 传真号: (86-10)62019451	受权官员 杨士林 电话号码: (86-10) 62085717	

国际检索报告		国际申请号 PCT/CN2013/073584
C(续). 相关文件		
类 型	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 101454825 A (哈曼国际工业有限公司) 10.6 月 2009 (10.06.2009), 全文	1-10
A	CN 101315772 A (上海交通大学) 03.12 月 2008 (03.12.2008), 全文	1-10

PCT/ISA/210 表(第 2 页续页) (2009 年 7 月)

国际检索报告 关于同族专利的信息		国际申请号 PCT/CN2013/073584	
检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN 102750956 A	24.10.2012	无	
CN 101385386 A	11.03.2009	WO 2007100137 A1 JP WO2007100137S X US 2009248403 A1 JP 4774100 B2 CN 101385386 B US 8271277 B2 EP 1993320 A1	07.09.2007 23.07.2009 01.10.2009 14.09.2011 09.05.2012 18.09.2012 19.11.2008
US 8160262 B2	17.04.2012	EP 2058804 A1 US 2009117948 A1	13.05.2009 07.05.2009
US 2008292108 A1	27.11.2008	EP 1885154 A1	06.02.2008
US 2008059157 A1	06.03.2008	JP 2008058900 A US 7590526 B2 JP 4107613 B2	13.03.2008 15.09.2009 25.06.2008
CN 1989550 A	27.06.2007	KR 20070036777 A US 2008300869 A1 JP 2008507720 A KR 1149591 B1 JP 5042823 B2 US 8116471 B2 WO 2006011104 A1 IN 200700280 P4 CN 1989550 B EP 1774517 A1	03.04.2007 04.12.2008 13.03.2008 29.05.2012 03.10.2012 14.02.2012 02.02.2006 24.08.2007 13.10.2010 18.04.2007
CN 101454825 A	10.06.2009	WO 2008034221 A1 JP 4964943 B2 JP 2009531722 A EP 2064699 A1	27.03.2008 04.07.2012 03.09.2009 03.06.2009
CN 101315772 A	03.12.2008	无	

PCT/ISA/210 表(同族专利附件) (2009 年 7 月)

フロントページの続き

(81)指定国 AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,RW,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,R,S,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,IL,IN,IS,JP,KE,KG,KM,KN,KP,KR,KZ,LA,LC,LK,LR,LS,LT,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC

(74)代理人 100167025

弁理士 池本 理絵

(74)代理人 100168642

弁理士 關谷 充司

(74)代理人 100096769

弁理士 有原 幸一

(74)代理人 100107319

弁理士 松島 鉄男

(74)代理人 100114591

弁理士 河村 英文

(72)発明者 ルー, シャシャ

中華人民共和国、シャントン 261031、ウェイハン シティー、ハイ - テック インダストリー ディストリクト、ドンファン ロード #268

(72)発明者 ウー, シャオチエ

中華人民共和国、シャントン 261031、ウェイハン シティー、ハイ - テック インダストリー ディストリクト、ドンファン ロード #268

(72)発明者 リー, ボー

中華人民共和国、シャントン 261031、ウェイハン シティー、ハイ - テック インダストリー ディストリクト、ドンファン ロード #268