

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2005-258158

(P2005-258158A)

(43) 公開日 平成17年9月22日(2005.9.22)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
G 1 0 L 21/02	G 1 0 L 9/00	5 D 0 1 5
G 1 0 L 11/00	G 1 0 L 7/02	
G 1 0 L 15/20	G 1 0 L 3/02	3 0 1 D

審査請求 未請求 請求項の数 6 O L (全 12 頁)

(21) 出願番号	特願2004-70786 (P2004-70786)	(71) 出願人	393031586 株式会社国際電気通信基礎技術研究所 京都府相楽郡精華町光台二丁目2番地2
(22) 出願日	平成16年3月12日 (2004.3.12)	(74) 代理人	100099933 弁理士 清水 敏
(出願人による申告) 平成15年度通信・放送機構、研究テーマ「大規模コーパスベース音声対話翻訳技術の研究開発」に関する委託研究、産業活力再生特別措置法第30条の適用を受ける特許出願		(72) 発明者	水町 光徳 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内
		(72) 発明者	中村 哲 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内
		Fターム(参考)	5D015 EE05

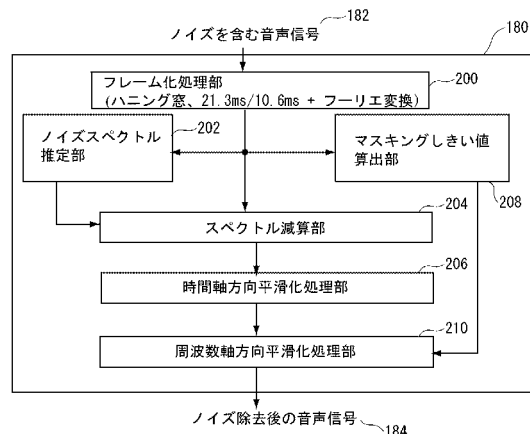
(54) 【発明の名称】 ノイズ除去装置

(57) 【要約】

【課題】 ノイズ除去精度が高く、処理後の信号に生じる歪みが少ないノイズ除去装置を提供する。

【解決手段】 ノイズ除去装置180は、音声信号182をフレーム化し、フレームごとの音響スペクトルを算出するフレーム化処理部200と、各フレームの音響スペクトルに含まれるノイズスペクトルを推定し、各フレームから減算するノイズスペクトル推定部202及びスペクトル減算部204と、フレーム化処理部200の出力する各フレームの音響スペクトルに対する周波数軸方向のマスクしきい値を算出するマスクしきい値算出部208と、スペクトル減算部204から出力されるフレームごとの音響スペクトルに対し、マスクしきい値と処理対象のフレームの直前のフレームの音響スペクトルとを用いて、時間軸方向と周波数軸方向との双方の平滑化を行なう時間軸方向平滑化処理部206及び周波数軸方向平滑化処理部210を含む。

【選択図】 図7



【特許請求の範囲】**【請求項 1】**

入力される音響信号を所定の周期で所定の時間長を有するフレームにフレーム化し、フレームごとの音響スペクトルを算出するためのフレーム化手段と、

前記フレーム化手段により出力される音響スペクトルに基づいて、各フレームの音響スペクトルに含まれるノイズスペクトルを推定し、当該各フレームの音響スペクトルから減算するためのノイズ減算手段と、

前記フレーム化手段の出力する各フレームの音響スペクトルに対する周波数軸方向のマスキングしきい値を算出するためのマスキングしきい値算出手段と、

前記ノイズ減算手段から出力されるフレームごとの音響スペクトルに対し、前記マスキングしきい値と処理対象のフレームに対し時間軸上で所定の関係にあるフレームの音響スペクトルとを用いて、時間軸方向と周波数軸方向との双方の平滑化を行なうことにより平滑化された音声信号を出力するための平滑化処理手段とを含む、ノイズ除去装置。

10

【請求項 2】

前記平滑化処理手段は、

前記ノイズ減算手段から出力されるフレームごとの音響スペクトルに対し、前記所定の関係にあるフレームの音響スペクトルを用いて時間軸方向の平滑化を行なうための時間軸方向平滑化手段と、

前記時間軸方向平滑化手段の出力するフレームごとの音響スペクトルに対し、前記マスキングしきい値算出手段により算出されたマスキングしきい値を用いた周波数軸方向の平滑化を行なって、周波数軸方向に平滑化された音声信号を生成するための周波数軸方向平滑化処理手段とを含む、請求項 1 に記載のノイズ除去装置。

20

【請求項 3】

前記時間軸方向平滑化手段は、

前記ノイズ減算手段の出力する各フレームの音響スペクトルの各周波数成分について、前記所定の関係にあるフレームの音響スペクトル内で対応する周波数成分との値の相違が所定の条件を充足しているか否かを判定するための第 1 の判定手段と、

前記第 1 の判定手段により前記所定の条件を充足していると判定された周波数成分の値を、予め定めた推定方法により推定された値で置換するための手段とを含む、請求項 2 に記載のノイズ除去装置。

30

【請求項 4】

前記第 1 の判定手段は、前記ノイズ減算手段の出力する各フレームの音響スペクトルの各周波数成分と、処理対象のフレームに対し時間軸上で 1 フレーム前のフレームの音響スペクトル内で対応する周波数成分との値の相違が所定のしきい値を超えているか否かを判定するための手段を含む、請求項 3 に記載のノイズ除去装置。

【請求項 5】

前記周波数軸方向平滑化手段は、

前記時間軸方向平滑化手段の出力する各フレームの音響スペクトルの各周波数成分について、その値が前記マスキングしきい値に対し所定の関係を充足しているか否かを判定するための第 2 の判定手段と、

40

前記第 2 の判定手段により前記所定の関係を充足していると判定された周波数成分を所定の値で置換するための置換手段とを含む、請求項 1 ~ 請求項 4 のいずれかに記載のノイズ除去装置。

【請求項 6】

前記第 2 の判定手段は、前記時間軸方向平滑化手段の出力する各フレームの音響スペクトルの各周波数成分について、その値が前記マスキングしきい値を超えているか否かを判定するための手段を含む、請求項 5 に記載のノイズ除去装置。

【発明の詳細な説明】**【技術分野】****【0001】**

50

本発明は、音声信号に付加された音響雑音を除去するためのノイズ除去装置に関し、特に、スペクトルサブトラクション（以下「SS」と呼ぶ。）法を用いたノイズ除去装置に関する。

【背景技術】

【0002】

近年、音響情報を扱う種々の装置の実用化に伴ない、ノイズを除去する技術は必要不可欠となっている。現在、受信系の形態を問わず広く採用されているノイズ除去の手法として、後掲の非特許文献1に記載のSS法がある。

【0003】

非特許文献1に記載のSS法は、ノイズが定常性を有し、かつ音声と無相関であるという仮定に基づいている。すなわち、この方法は、発話の前に観測されたノイズの振幅スペクトル（以下、「ノイズスペクトル」と呼ぶ。）が、発話に含まれるノイズスペクトルと同じであると仮定し、発話から得られた音声信号の振幅スペクトルから、発話直前に観測されたノイズスペクトルを減算することでノイズを除去する。

10

【0004】

この方法は、上記のとおり、ノイズが定常性を有するという仮定に基づいているため、非定常なノイズに対しては十分なノイズ除去を行なうことができない。また、SS法では、独特の残留ノイズ（ミュージカルノイズ）が発生し、ノイズ除去処理後の音声信号の品質が低下する。

【0005】

SS法におけるこれらの問題を解決することを目的として、これまでに種々の改良法が開発されてきている。後掲の非特許文献2には、小規模マイクロホンアレイを用いたSS改良法が開示されている。このSS改良法は、入力音声の受信に小規模マイクロホンアレイを用いる。マイクロホンアレイによる受信信号から到来信号の空間情報を得ることで、時々刻々に変化するノイズスペクトルを推定する。

20

【0006】

【非特許文献1】S. F. ボル、「スペクトルサブトラクションを用いた、音声内の音響ノイズの抑制」、IEEE Trans. ASSP, Vol. 27, No. 2, pp. 113 - 120, 1979年 (S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction.", IEEE Trans. ASSP, Vol. 27, No. 2, pp. 113 - 120 (1979))

30

【非特許文献2】水町他一名、「マイクロホン対を用いたスペクトルサブトラクションによる雑音除去法」電子情報通信学会論文誌(A) Vol. J82-A, No. 4, pp. 503 - 512 (1999)

【発明の開示】

【発明が解決しようとする課題】

【0007】

処理後音声の歪みは、大別すると次の2種のものがある。即ち、第1の種類の歪みは、広い周波数帯域に分散する歪みである。第2の種類の歪みは、特定の時間、及び特定の周波数領域に突発的に発生するスパイク状の歪みである。特に、後者の歪みは処理後音声の品質劣化に大きな影響を及ぼすと考えられている。しかし、非特許文献2に記載の方法をはじめ、現在提案されているいずれのSS改良法も、処理後音声のこのような歪みを完全に除去できるとは言いがたい。

40

【0008】

それゆえに、本発明の目的は、ノイズ除去精度が高く、かつ処理後の信号に生じる歪みが少ないノイズ除去装置を提供することである。

【0009】

本発明の別の目的は、SS処理後音声に存在する不自然なスパイク上の歪みを除去することが可能なノイズ除去装置を提供することである。

50

【課題を解決するための手段】

【0010】

本発明に係るノイズ除去装置は、入力される音響信号を所定の周期で所定の時間長を有するフレームにフレーム化し、フレームごとの音響スペクトルを算出するためのフレーム化手段と、フレーム化手段により出力される音響スペクトルに基づいて、各フレームの音響スペクトルに含まれるノイズスペクトルを推定し、当該各フレームの音響スペクトルから減算するためのノイズ減算手段と、フレーム化手段の出力する各フレームの音響スペクトルに対する周波数軸方向のマスキングしきい値を算出するためのマスキングしきい値算出手段と、ノイズ減算手段から出力されるフレームごとの音響スペクトルに対し、マスキングしきい値と処理対象のフレームに対し時間軸上で所定の関係にあるフレームの音響スペクトルとを用いて、時間軸方向と周波数軸方向との双方の平滑化を行なうことにより平滑化された音声信号を出力するための平滑化処理手段とを含む。

10

【0011】

好ましくは、平滑化処理手段は、ノイズ減算手段から出力されるフレームごとの音響スペクトルに対し、所定の関係にあるフレームの音響スペクトルを用いて時間軸方向の平滑化を行なうための時間軸方向平滑化手段と、時間軸方向平滑化手段の出力するフレームごとの音響スペクトルに対し、マスキングしきい値算出手段により算出されたマスキングしきい値を用いた周波数軸方向の平滑化を行なって、周波数軸方向に平滑化された音声信号を生成するための周波数軸方向平滑化処理手段とを含む。

20

【0012】

より好ましくは、時間軸方向平滑化手段は、ノイズ減算手段の出力する各フレームの音響スペクトルの各周波数成分について、所定の関係にあるフレームの音響スペクトル内で対応する周波数成分との値の相違が所定の条件を充足しているか否かを判定するための第1の判定手段と、第1の判定手段により所定の条件を充足していると判定された周波数成分の値を、予め定めた推定方法により推定された値で置換するための手段とを含む。

【0013】

第1の判定手段は、ノイズ減算手段の出力する各フレームの音響スペクトルの各周波数成分と、処理対象のフレームに対し時間軸上で1フレーム前のフレームの音響スペクトル内で対応する周波数成分との値の相違が所定のしきい値、例えばフレーム長21.3ms、フレーム周期10.6msの場合、30dBを超えているか否かを判定するための手段を含んでもよい。

30

【0014】

また、周波数軸方向平滑化手段は、時間軸方向平滑化手段の出力する各フレームの音響スペクトルの各周波数成分について、その値がマスキングしきい値に対し所定の関係を充足しているか否かを判定するための第2の判定手段と、第2の判定手段により所定の関係を充足していると判定された周波数成分を所定の値で置換するための置換手段とを含んでもよい。

【0015】

好ましくは、第2の判定手段は、時間軸方向平滑化手段の出力する各フレームの音響スペクトルの各周波数成分について、その値がマスキングしきい値を超えているか否かを判定するための手段を含む。

40

【0016】

さらに好ましくは、置換手段は、第2の判定手段により所定の関係を充足していると判定された周波数成分を、判定対象の周波数成分に対応するマスキングしきい値で置換するための手段を含んでもよい。

【発明を実施するための最良の形態】

【0017】

[制約条件]

本発明の一実施の形態に係るノイズ除去装置は、ノイズと音声とが混在する環境下で受音された音声の信号から、SS法を用いてノイズを除去する装置である。しかし、上記の

50

通り、単純にSS法を用いてノイズ除去を行なうと、処理後の音声に歪みが生じる。

【0018】

そこで本実施の形態に係るノイズ除去装置では、まず人間の聴覚特性を考慮した制約条件を導入する。そのためにまず人間聴覚系のマスキング特性を工学的にシミュレートし、連続かつ滑らかに変化するマスキングしきい値を求め、そのマスキングしきい値を用いてSSにより振幅スペクトルに生じた周波数軸方向の不連続性を解消することを試みる。

【0019】

さらに本実施の形態に係るノイズ除去装置では、環境及び時間の関数であり予測困難なノイズ信号ではなく、目的信号である音声に着目した時間軸方向の制約条件を導入する。

【0020】

音声は、時間とともに変化する非定常信号であるが、SS処理後音声の歪みの動特性と比較すると、20ms程度の区間内ではほぼ定常な信号とみなすことができる。音声は時間的に滑らかな信号であるという時間軸方向の制約条件を導入することにより、SSにより音声に生じる突発的なスペクトル歪みの抑圧を試みる。

【0021】

- マスキング特性による制約条件の検討 -

人間の聴覚系のマスキング特性を工学的にシミュレートする方法について図1及び図2を参照して説明する。なお、本実施の形態のノイズ除去装置では、人間の聴覚特性のうち、同時マスキング特性(周波数軸方向のマスキング特性)のみを考慮し、時間軸方向のマスキング特性及びレベル依存性については考慮しない。

【0022】

図1を参照して、ある音声30(マスキングの原因となるという意味でこの音声を「マスカ」と呼ぶ。)が存在するとき、このマスカ30により生じるマスキング範囲を図1の曲線32で示す。この範囲に存在する音は人間には知覚できない。すなわち、マスキング範囲を画する線は、マスキングのしきい値を示す。

【0023】

この曲線32をシミュレートするために、三点ABCで定められる図形34を用いる。図形34の各点A, B, Cの位置は、図1に示すようにマスカ30のレベルとその周波数k(Hz)との関係で定まる。

【0024】

次に、このようなシミュレート方法を用いて、音声信号に対するマスキングしきい値を次のようにして求める。本実施の形態では、パワー・ロー・モデルと呼ばれるモデルを用いる。図2を参照してその概略を説明する。

【0025】

図2に示されるように音声信号の短時間対数スペクトル50が存在しているものとする。この音声信号に対するマスキングしきい値は以下のようにして求められる。まず、このスペクトル50の振幅の最大のピーク(図2ではピーク52)をマスカ候補として求める。このピークに対して、シミュレートされたマスキング領域54を求める。続いて、スペクトル50の周波数を中心とする臨界帯域の外で次に大きなピーク(図2においてはピーク56)を次のマスカ候補として求める。このピークに対して、最初と同様にマスキング領域58を求める。以下同様に、次のマスカ候補としてのピーク60、それに対するマスキング領域62、...、を求める。

【0026】

このようにして求められたマスキング領域54, 58, 62, ...により定まるマスキング領域の和を求めることで、周波数軸方向に連続した、マスキングしきい値を示す(周波数の)関数を考えることができる。

【0027】

すなわち、マスキングしきい値 $Threshold_{total}$ は、次の式で求められる。

【0028】

10

20

30

40

【数 1】

$$Threshold_{total} = \left[\sum_{\omega} threshold^p(\omega) \right]^{\frac{1}{p}}$$

ただし $threshold(\omega)$ は周波数 ω のマスクによるマスクしきい値を表す関数である。p は 0.3 ~ 0.4 程度の値であり、本実施の形態では $p = 0.33$ を用いる。

【0029】

上記のようにして求めたマスクしきい値の時間変化の例を、図 3 に示す。図 3 に示すグラフ 100 のうち、比較的平坦な部分がマスクしきい値を示し、突出して見える部分は音声の短時間スペクトルのうち、マスクしきい値より大きな値の部分を示す。図 3 から分かるように、音声スペクトルのうちしきい値より小さな部分については平滑化される。本実施の形態では、このようにして音声スペクトルの谷を埋め、不連続をなくす。

10

【0030】

- 時間軸方向の制約条件の検討 -

次に、音声時間が時間軸方向に滑らかに変化するという前提による制約条件について検討する。図 4 に、特定のテストデータに対する測定により得られた、各周波数における音声信号の対数振幅スペクトル値の隣接フレーム間での変化量のヒストグラムを示す。

【0031】

図 5 を参照して、対数振幅の相違の概念を説明する。図 5 に示すように、ある時間 t における音響スペクトル 122 の i 番目の周波数成分の対数振幅の値 $X_t(f_i)$ で表すものとする。その 1 フレーム前 (時間 $t-1$) の音響スペクトル 120 の、これに対応する i 番目の周波数成分の対数振幅の値は $X_{t-1}(f_i)$ で表される。したがって、時間軸上で隣接する二つのフレームの対応する周波数成分の値の差 ΔX とすると、 $\Delta X = X_t(f_i) - X_{t-1}(f_i)$ で表される。

20

【0032】

図 4 を参照して容易に分かるように、時間軸上で隣接する二つのフレーム間の振幅の差は 0 の近辺に集中しており、これらの間では音声信号の対数振幅スペクトルは時間的に滑らかに変化することが分かる。

30

【0033】

図 6 に、時間軸上で隣接するフレーム間の対数振幅の差が 10 dB、20 dB、30 dB にそれぞれ入るものの割合を、周波数との関数としてそれぞれ曲線 150、152 及び 154 として示す。図 6 の曲線 154 を参照して分かるように、時間軸上で隣接する二つのフレーム間の振幅の差が 30 dB 以上であることはほとんどなく、30 dB を超える場合は異常であると考えることができる。

【0034】

[構成]

図 7 に、以上の制約条件を考慮した SS によりノイズ除去を試みる、本実施の形態に係るノイズ除去装置の機能的構成をブロック図形式で示す。図 7 を参照して、ノイズ除去装置 180 は、ノイズを含む音声信号 182 を窓長 21.3 ms のハニング窓を用い、時間間隔 10.6 ms でフレーム化しさらにそれらにフーリエ変換を行なってフレームごとの振幅スペクトルを出力するフレーム化処理部 200 と、フレーム化処理部 200 が出力する振幅スペクトルを受け、フレーム毎にノイズスペクトルを推定するノイズスペクトル推定部 202 と、フレーム化処理部 200 から振幅スペクトルを、ノイズスペクトル推定部 202 から推定ノイズスペクトルを、それぞれ受けるように接続され、フレーム化処理部 200 からの振幅スペクトルから推定ノイズスペクトルを減算するスペクトル減算部 204 と、スペクトル減算部 204 の出力に対し、時間軸方向の音声の平滑性による平滑化処理を実行するための時間軸方向平滑化処理部 206 とを含む。

40

【0035】

50

ノイズ除去装置 180 はさらに、フレーム化処理部 200 からフレームごとの振幅スペクトルを受けるように接続され、この振幅スペクトルに対して聴覚特性を考慮したマスキングしきい値を算出するためのマスキングしきい値算出部 208 と、時間軸方向平滑化処理部 206 により時間軸方向の平滑化がされたフレームごとの音声スペクトルについて、マスキングしきい値算出部 208 により算出されたマスキングしきい値を用いた周波数軸方向の平滑化を行ない、音声信号 184 の形で出力するための周波数軸方向平滑化処理部 210 とを含む。

【0036】

本実施の形態におけるノイズスペクトル推定部 202 によるノイズスペクトルの推定処理について説明する。図 8 を参照して、本実施の形態では、ターゲットとなる音源 240 と、ノイズ源 242 とが異なる方向に存在することを仮定し、二つのマイクロホン 250 及び 252 を含む 2 チャンネルのマイクロホンアレイにより、受信位置 254 における、ノイズ源 242 からのノイズのスペクトルを推定する。二つの音源が異なる方向に存在する場合、二つのマイクロホンによる観測信号を用いて、これら音源からの信号を互いに分離できることが知られている。

10

【0037】

すなわち、音源 240 からの音声と、ノイズ源 242 からのノイズとの 2 つの信号到来方向を推定し、音源 240 からの音声を完全に抑制するように 2 チャンネル減算型ビームフォーマを設計し、ノイズスペクトルを推定する。この方法では、信号の到達時間差という空間情報を利用したフィルタリングを行なうため、短時間音声に対するノイズスペクトルをフレーム毎に推定することが可能である。したがってこの方法では、非定常ノイズのスペクトルも推定可能である。

20

【0038】

本実施の形態では、この公知の減算型ビームフォーマを形成するマイクロホンアレイを用いて音源 240 からの音声信号を抑圧し、それによってノイズ源 242 からのノイズのスペクトルを推定する。

【0039】

スペクトル減算部 204 によるノイズスペクトルの減算処理は、通常の SS によるものと同様である。スペクトル減算部 204 の出力を、以下「推定振幅スペクトル」と呼ぶことにする。

30

【0040】

図 9 に、時間軸方向平滑化処理部 206 の実行する処理のフローチャートを示す。図 9 を参照して、時間軸方向平滑化処理部 206 は、推定振幅スペクトルの各周波数成分について、時間軸上で隣接するフレームの同一周波数成分との差を求め、この差が 30 dB を超えるか否かを判定する（ステップ 270）。差が 30 dB を超えない場合には何もしない。差が 30 dB を超える場合、音声の時間軸上の平滑性によりそのような差は異常と考えられる。したがってステップ 272 でこのフレームのこの周波数成分の値を、過去のフレームからの予測値で置換する。

【0041】

この処理により、時間軸上で直前のフレームと比較して大きく違う周波数成分を検出し、より滑らかに変化する値に平滑化することができる。

40

【0042】

図 10 に、周波数軸方向平滑化処理部 210 の実行する処理のフローチャートを示す。図 10 を参照して、周波数軸方向平滑化処理部 210 は、時間軸方向平滑化処理部 206 から時間軸方向の平滑化処理が施されたフレームごとのスペクトルを受け、フレームごとに周波数軸方向の平滑化処理を行なう。

【0043】

具体的には周波数軸方向平滑化処理部 210 は、時間軸方向の平滑化がされた振幅スペクトルの各フレームに含まれる各周波数の対数振幅スペクトル値に対し、その値がマスキングしきい値算出部 208 から与えられるマスキングしきい値より大きいかなんかを判定す

50

る(ステップ280)。各周波数の対数振幅スペクトル値が、対応するマスクングしきい値より大きい場合には何もしない。マスクングしきい値以下の場合には、ステップ282でその部分の値をマスクングしきい値で置換することにより、対数振幅スペクトルの谷を埋める。

【0044】

この処理を各周波数成分に対して実行することにより、周波数軸方向平滑化処理部210はこのフレームに対する周波数軸方向の平滑化処理を行なう。

【0045】

[動作]

本実施の形態に係るノイズ除去装置180は以下のように動作する。

10

【0046】

図7を参照して、ノイズが存在する環境化で受信されたノイズを含む音声からの音声信号182がノイズ除去装置180に与えられる。フレーム化処理部200は、与えられた入力信号を10.6msごとに21.3msの長さのフレームに順次フレーム化する。これらフレームの音声信号に対してフーリエ変換を行なって得られた音声スペクトルは、ノイズスペクトル推定部202及びマスクングしきい値算出部208に与えられる。

【0047】

ノイズスペクトル推定部202は、与えられた各フレームの音声信号スペクトルについて、図8を参照して説明した方法によりノイズスペクトルの推定を行なう。ノイズスペクトル推定部202は、推定されたノイズスペクトルをスペクトル減算部204に与える。

20

【0048】

スペクトル減算部204は、フレーム化処理部200から受けた各フレームの音声信号スペクトルから、ノイズスペクトル推定部202から受けた同一フレームのノイズスペクトルを減算し、推定振幅スペクトルを時間軸方向平滑化処理部206に与える。

【0049】

時間軸方向平滑化処理部206は、一つ前のフレームに関する推定振幅スペクトルを保持しており、図9のフローチャートにより示される処理を処理対象のフレームの各周波数成分に対して実行する。この処理により、直前のフレームに対して30dBを越える相違を持つような周波数成分の値は、過去のフレームから推定された値で置換される。この結果、時間軸方向の平滑化が実現される。

30

【0050】

一方、マスクングしきい値算出部208は、フレーム化処理部200から1フレーム分の音声信号スペクトルが与えられると、与えられたフレームに対するマスクングしきい値を算出する。マスクングしきい値算出部208は、算出したマスクングしきい値を周波数軸方向平滑化処理部210に与える。

【0051】

周波数軸方向平滑化処理部210は、時間軸方向平滑化処理部206の出力と、この出力に対応するフレームのマスクングしきい値とが与えられると、以下のようにして、与えられたフレームにおけるスペクトルの平滑化を行なう。

【0052】

すなわち、周波数軸方向平滑化処理部210は、図10に示すように、与えられた時間軸方向平滑化後のスペクトルの各周波数成分と、当該周波数におけるマスクングしきい値とを比較する(ステップ280)。周波数成分の値がマスクングしきい値以下の場合、当該値をマスクングしきい値で置換する。それ以外の場合には何もしない。マスクングしきい値以下の周波数成分は、本来マスクングされているものでありそれほど重要なものとは考えられない。そこで、このようなスペクトル上の谷に相当する部分をマスクングしきい値で置換して平滑化することにより、最終的な歪みの発生を抑える。周波数軸方向平滑化処理部210は、以上のようにしてノイズが除去され、かつ時間軸方向と周波数軸方向との双方に音声スペクトルを平滑化するようにして音声信号を推定し、音声信号184として出力する。

40

50

【0053】

[実験による検証]

上記した本実施の形態のノイズ除去装置180によるノイズ除去の有効性について、実験により検証を行なった。図11を参照して、この実験では、目的音声となる音源300は正面方向に、ノイズ源は正面から右に30度方向に、それぞれ位置するものとした。ノイズは無相関Gaussianノイズである。マイクロホン314から音源300とノイズ源302とまでの距離は等しい。

【0054】

ノイズスペクトル推定のために使用するマイクロホンアレイのマイクロホン310及び312は、マイクロホン314の左右、等距離に設けた。音源300までの距離は1.0m、マイクロホン314からマイクロホン310及び312までの距離はいずれも0.1mである。

【0055】

図12に、ノイズ付加音声、上記した制約条件を用いずにノイズ除去を行なった音声(ベースライン:BL)、マスキング特性を導入してノイズ除去を行なった音声(BL+周波数軸方向平滑化)、及びこれに加えて時間軸方向平滑化を導入してノイズ除去を行なった音声(BL+周波数軸方向平滑化+時間軸方向平滑化)のノイズ除去後音声の平均スペクトル歪みを示す。

【0056】

図12から、ノイズ環境下において、本実施の形態に係る聴覚特性及び音声生成系に基づく平滑化処理の有効性が確認できる。

【0057】

なお、本実施の形態では、複数マイクロホンを用いた減算型ビームフォーマにより音響信号中のノイズスペクトルの推定を行なったが、これに限らず現在までに提案されているものおよび将来提案されるものも含みどのようなノイズ推定の手法を用いることもできる。

【0058】

また、図9のステップ272で使用される予測値としては、直前の1フレームから予測した値、例えば直前の1フレームの対応周波数成分の値を用いることもできるし、直前の複数フレームでの対応周波数成分の推移から外挿した値を用いてもよい。

【0059】

さらに、図9のステップ270及び図10のステップ280では、各値が所定のしきい値を超えているか否かを判定しているが、判定条件はこれに限らない。例えば各値が所定のしきい値以上か否かを判定してもよい。また、例えば図9のステップ270ではしきい値として30dBという定数を用いているが、このしきい値を、過去のフレームからの予測により求めるようにしてもよい。

【0060】

また、図10のステップ282では、スペクトルの谷部の値をマスキングしきい値で置換しているが、これに限らずマスキングしきい値又は振幅の大きさに基づいて何らかの関数で算出される値で置換するようにしてもよい。

【0061】

今回開示された実施の形態は単に例示であって、本発明が上記した実施の形態のみに制限されるわけではない。本発明の範囲は、発明の詳細な説明の記載を参酌した上で、特許請求の範囲の各請求項によって示され、そこに記載された文言と均等の意味および範囲内でのすべての変更を含む。

【図面の簡単な説明】

【0062】

【図1】マスカによるマスキング領域の工学的シミュレーションを説明するための図である。

【図2】パワー・ロー・モデルによるマスキングしきい値の計算方法を説明するための図

10

20

30

40

50

ラフである。

【図3】マスキングしきい値を適用した音声スペクトルの時間変化を示すグラフである。

【図4】時間軸上で隣接するフレームの対応する周波数成分の変化量のヒストグラムを示す図である。

【図5】時間軸上で隣接するフレームの対応する周波数成分の変化量を説明するための図である。

【図6】時間軸上で隣接するフレームの対応する周波数成分の変化量が所定の大きさ未満のもの率を説明するための図である。

【図7】本発明の一実施の形態に係るノイズ除去装置180の機能的構成を示すブロック図である。

【図8】本発明の一実施の形態のノイズスペクトル推定部202で使用する2チャンネル減算型ビームフォーマの構成を模式的に示す図である。

【図9】時間軸方向平滑化処理部206による処理の過程を示すフローチャートである。

【図10】周波数軸方向平滑化処理部210による処理の過程を示すフローチャートである。

【図11】実験における音源、ノイズ源、受信位置及びマイクロホンの配置を説明するための模式図である。

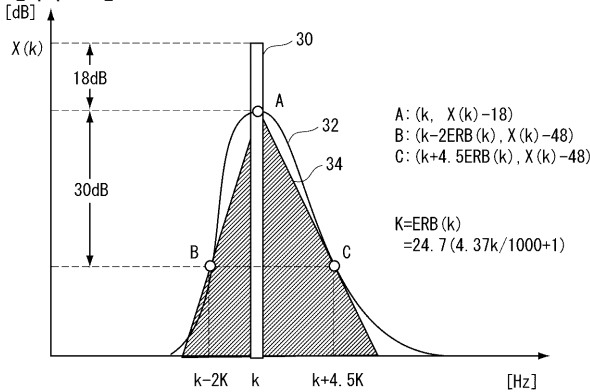
【図12】実験結果を示すグラフである。

【符号の説明】

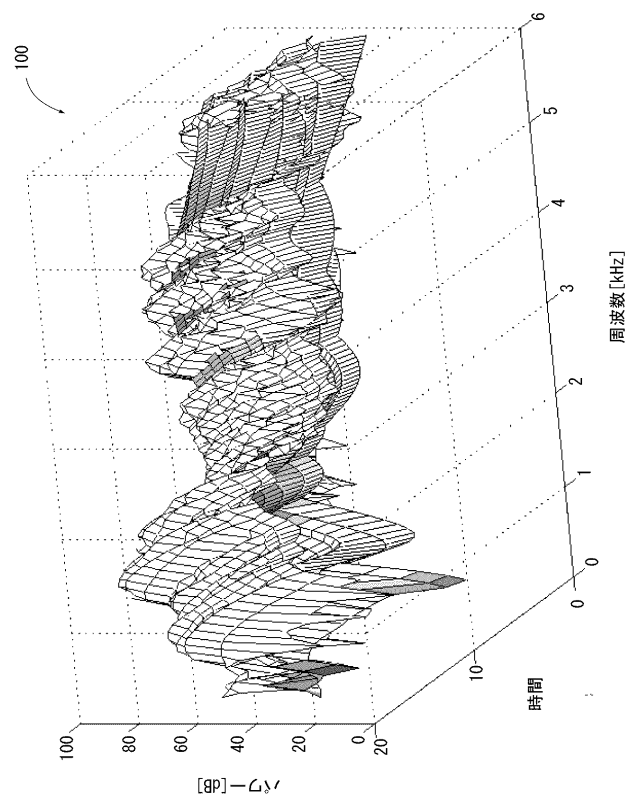
【0063】

180 ノイズ除去装置、182 ノイズを含む音声信号、184 ノイズ除去後の音声信号、200 フレーム化処理部、202 ノイズスペクトル推定部、204 スペクトル減算部、206 時間軸方向平滑化処理部、208 マスキングしきい値算出部、210 周波数軸方向平滑化処理部

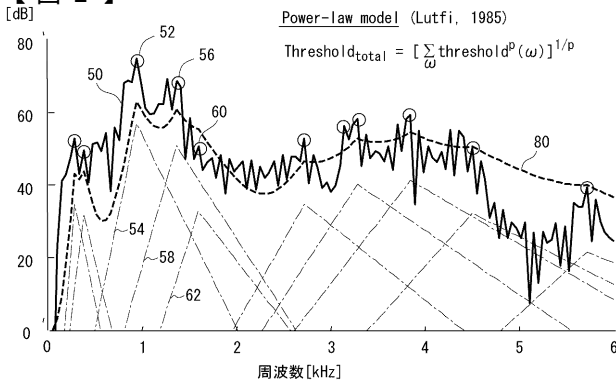
【図1】



【図3】



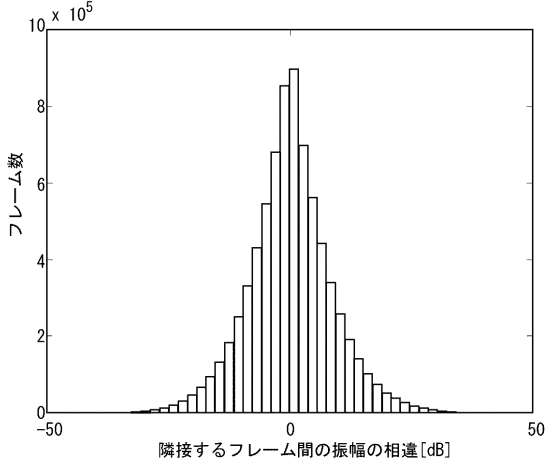
【図2】



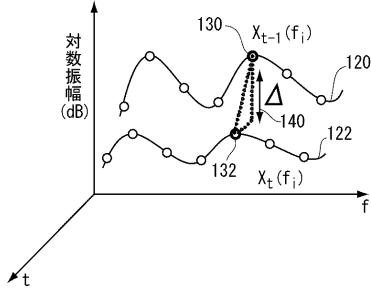
10

20

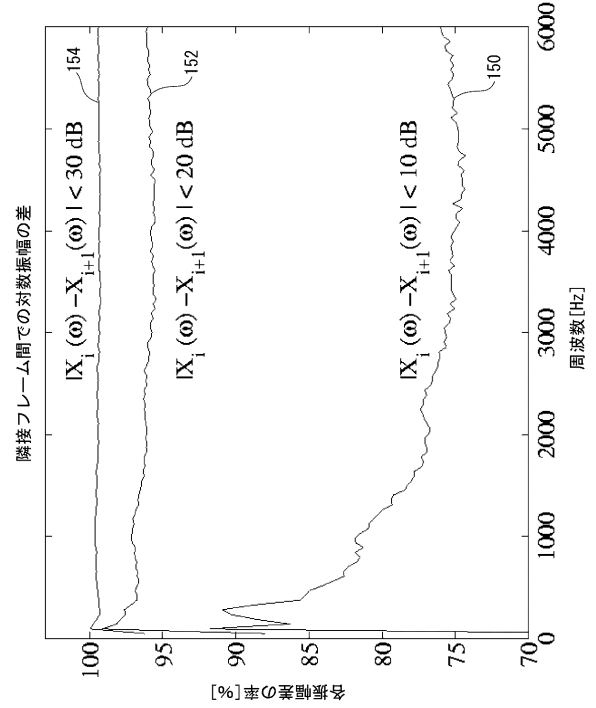
【 図 4 】



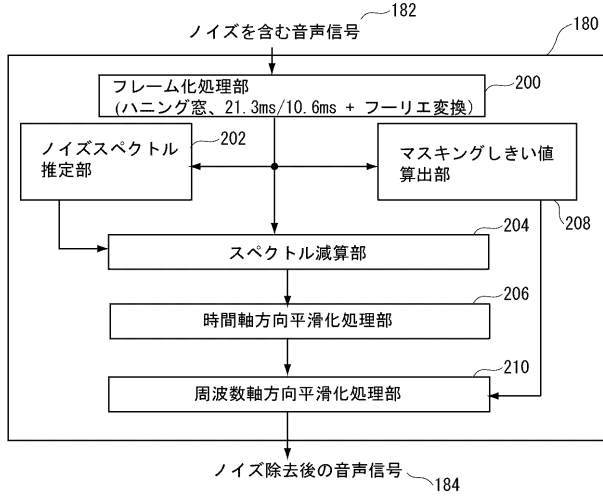
【 図 5 】



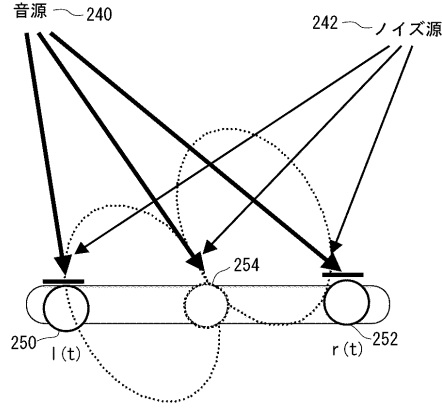
【 図 6 】



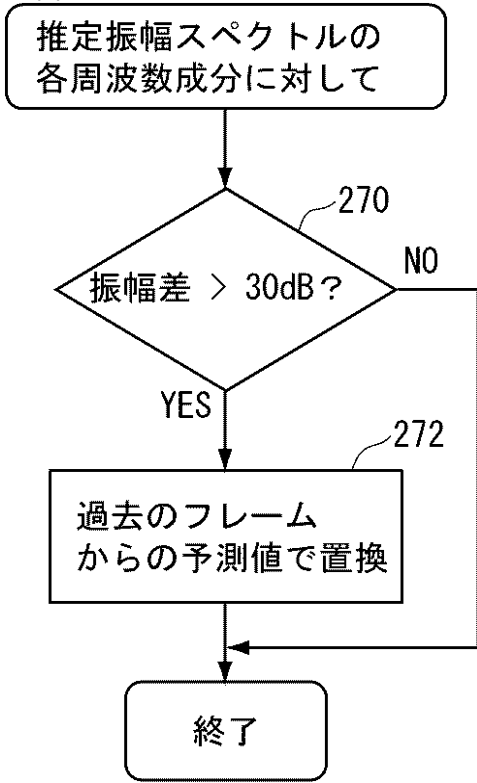
【 図 7 】



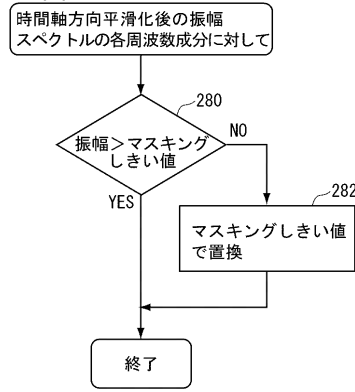
【 図 8 】



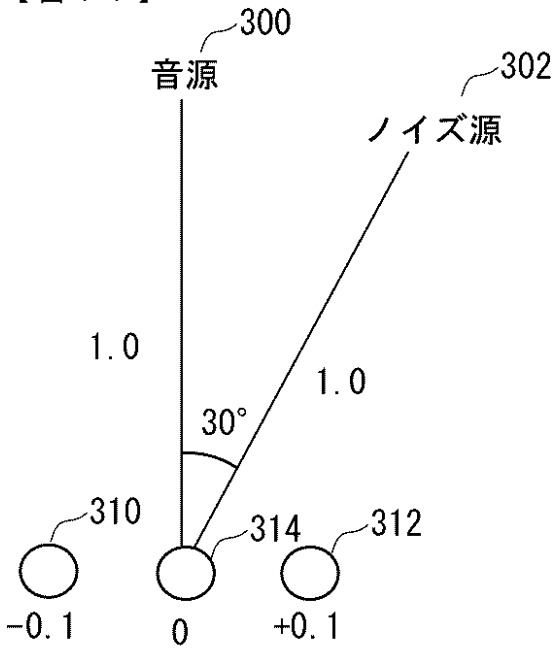
【図9】



【図10】



【図11】



【図12】

