

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4520732号
(P4520732)

(45) 発行日 平成22年8月11日(2010.8.11)

(24) 登録日 平成22年5月28日(2010.5.28)

(51) Int.Cl.

F I

G 1 O L 21/02 (2006.01)

G 1 O L 21/02 I O I B

請求項の数 6 (全 23 頁)

(21) 出願番号	特願2003-404595 (P2003-404595)	(73) 特許権者	000005223
(22) 出願日	平成15年12月3日 (2003.12.3)		富士通株式会社
(65) 公開番号	特開2005-165021 (P2005-165021A)		神奈川県川崎市中原区上小田中4丁目1番1号
(43) 公開日	平成17年6月23日 (2005.6.23)	(74) 代理人	100074099
審査請求日	平成18年8月22日 (2006.8.22)		弁理士 大菅 義之
		(74) 代理人	100067987
			弁理士 久木元 彰
		(72) 発明者	遠藤 香緒里
			神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
		(72) 発明者	大谷 猛
			神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

最終頁に続く

(54) 【発明の名称】 雑音低減装置、および低減方法

(57) 【特許請求の範囲】

【請求項 1】

入力音声信号の1フレーム分を周波数分析して周波数領域の信号に変換して出力する分析部と、

現在処理対象となっているフレームにおける周波数帯域毎にスペクトルパワーの分布を求め、該周波数帯域毎に、上位一定割合のスペクトルパワーの平均値を算出する音声情報推定手段と、

前記周波数帯域毎の前記平均値と前記スペクトルパワーとの差に基づいて、前記フレームに対する抑圧ゲインを算出する抑圧ゲイン算出手段と、

前記分析部が出力した周波数領域の信号を前記抑圧ゲイン算出手段が算出した抑圧ゲインに基づいて抑圧する抑圧手段と、

前記抑圧手段により抑圧された周波数領域の信号を時間領域の信号に合成する合成手段と、

を備えることを特徴とする雑音低減装置。

【請求項 2】

更に、前記現在処理対象となっているフレームに対する前記周波数帯域毎のスペクトル雑音を求める雑音推定手段を備え、

前記抑圧ゲイン算出手段は、前記周波数帯域毎の前記平均値と前記スペクトルパワーとの差、および前記周波数帯域毎の前記平均値と前記スペクトル雑音との差に基づいて、前記フレームに対する抑圧ゲインを算出する、

10

20

ことを特徴とする請求項 1 記載の雑音低減装置。

【請求項 3】

更に、前記抑圧ゲイン算出手段は、

前記入力音声信号のスペクトルパワーの分布の下限を、前記平均値から純粋な音声パワーの存在幅を引くことによって算出し、

前記周波数帯域毎のスペクトル雑音を N とし、雑音判定区間におけるパワーの幅を表す値を λ としたときの両者の和 $(N + \lambda)$ を前記周波数帯域毎に求め、前記現在処理対象となっているフレームを含める或る一定の過去からの監視期間における複数の音声フレーム信号において、各フレームの前記周波数帯域毎のスペクトルパワーが該周波数帯域毎に求めた該和 $(N + \lambda)$ と前記算出した前記入力音声信号のスペクトルパワー分布の下限との間に
10 あるか否かによって、それぞれのフレームに非定常雑音が含まれるかどうかを判定し、前記監視期間に非定常雑音を検出された頻度を前記周波数帯域毎に算出し、

前記フレームに対する抑圧ゲインを算出する際に、前記頻度の値が大きいほど抑圧ゲインを小さくする関数を前記周波数帯域毎に用いて算出することを特徴とする請求項 2 記載の雑音低減装置。

【請求項 4】

入力音声信号の 1 フレーム分を周波数分析して周波数領域の信号に変換して出力する分析部を用いて雑音を低減する雑音低減方法であって、

現在処理対象となっているフレームにおける周波数帯域毎のスペクトルパワーの分布を求め、
20

該周波数帯域毎に、上位一定割合のスペクトルパワーの平均値を算出し、

前記周波数帯域毎の前記平均値と前記スペクトルパワーとの差に基づいて、前記フレームに対する抑圧ゲインを算出し、

前記分析部が出力した周波数領域の信号を、前記抑圧ゲインに基づいて抑圧し、

前記抑圧ゲインに基づいて抑圧された周波数領域の信号を時間領域の信号に合成する、
ことを特徴とする雑音低減方法。

【請求項 5】

請求項 4 に記載の雑音低減方法において、

更に、前記現在処理対象となっているフレームに対する前記周波数帯域毎のスペクトル雑音を求め、
30

前記抑圧ゲインの算出では、前記周波数帯域毎の前記平均値と前記スペクトルパワーとの差、および前記周波数帯域毎の前記平均値と前記スペクトル雑音との差に基づいて、前記フレームに対する抑圧ゲインを算出する、

ことを特徴とする雑音低減方法。

【請求項 6】

請求項 5 に記載の雑音低減方法において、更に、

前記入力音声信号のスペクトルパワーの分布の下限を、前記平均値から純粋な音声パワーの存在幅を引くことによって算出し、

前記周波数帯域毎のスペクトル雑音を N とし、雑音判定区間におけるパワーの幅を表す値を λ としたときの両者の和 $(N + \lambda)$ を前記周波数帯域毎に求め、前記現在処理対象となっているフレームを含める或る一定の過去からの監視期間における複数の音声フレーム信号において、各フレームの前記周波数帯域毎のスペクトルパワーが該周波数帯域毎に求めた該和 $(N + \lambda)$ と前記算出した前記入力音声信号のスペクトルパワー分布の下限との間に
40 あるか否かによって、それぞれのフレームに非定常雑音が含まれるかどうかを判定し、前記監視期間に非定常雑音を検出された頻度を前記周波数帯域毎に算出し、

前記フレームに対する抑圧ゲインを算出する際に、前記頻度の値が大きいほど抑圧ゲインを小さくする関数を前記周波数帯域毎に用いて算出することを特徴とする雑音低減方法

°
【発明の詳細な説明】

【技術分野】

10

20

30

40

50

【0001】

本発明は、環境雑音などの雑音が重畳した音声信号から雑音成分を低減させる方式に係り、さらに詳しくは、例えば携帯電話システムやIP電話システムなどで、マイクから入力される非音声の環境雑音が重畳した音声信号から雑音成分を低減させ、信号対雑音比(SNR)を向上させて、通話品質を高めるための雑音低減装置、および低減方法に関する。

【背景技術】

【0002】

近年、携帯電話などのデジタル移動通信システムが急速に普及している。このような通信においては、周囲の環境雑音が大きい状態で通信が行われることが多く、音声信号に含まれる雑音成分を抑圧することが重要である。

10

【0003】

このような雑音抑圧の技術として、例えば時間軸上の入力信号を周波数軸上の信号(振幅スペクトルと位相スペクトル)に変換し、非音声区間の信号によって推定される背景雑音から抑圧ゲインを求めて、振幅スペクトルを抑圧し、位相スペクトルと抑圧された振幅スペクトルを時間軸上の信号に戻すことで雑音を除去する技術(図20)がある。

【0004】

このような従来技術の問題点について次の4つの文献を例として説明する。

【非特許文献1】S. F. Boll, "Suppression of Acoustic Noise in Speech Using Spectral Subtraction", IEEE Transaction on Acoustics, Speech, and Signal Processing, ASSP-33, vol. 27, pp. 113-120, (1979)

20

【特許文献1】特許第3269969号 背景雑音消去装置

【特許文献2】特許第3437264号 雑音抑圧装置

【特許文献3】特開2002-73066号 雑音抑圧装置、および雑音抑圧方法

【0005】

非特許文献1では、入力の振幅スペクトルから推定雑音の振幅スペクトルを減算して、抑圧された振幅スペクトルを求めるスペクトルサブトラクションが提案されている。

【0006】

30

特許文献1では、入力信号を周波数軸上の信号に変換して、入力信号と推定雑音から算出される信号対雑音比(SNR)に基いて、抑圧ゲインの算出が行われている。抑圧ゲインの算出法としては、あらかじめ経験的にSNRと抑圧ゲインの関係式を定めておく方法が用いられている。

【0007】

特許文献2では、推定非音声区間のパワーが小さい場合には、抑圧度合いを小さくして、小さいパワーの音声区間が抑圧されることによる劣化を防ぎ、また非音声区間のパワーが大きい場合には、抑圧度合いを大きくして非音声区間をより大きく抑圧することによって、より適切に非音声区間の雑音を抑圧する技術が開示されている。

【0008】

40

特許文献3では、有音判定された区間の平滑化スペクトルパワーから音声信号のパワーを求め、無音判定された区間の平滑化スペクトルパワーから無音信号のパワーを求め、これらからSNRの算出が行われ、SNRの高い信号部分に対してはより強い雑音抑圧を行い、SNRの低い信号部分に対しては抑圧による歪の生じる部分に抑圧の制限をかける技術が開示されている。

【0009】

しかしながらこれらの従来技術において、背景雑音の推定を誤った場合には適切な抑圧ゲインを求めることができず、雑音抑圧した結果の音声信号が劣化してしまうという問題点があった。例えば、背景雑音にバブル雑音(人の声が含まれるような背景雑音)が多く含まれる場合には、バブル雑音の区間が非音声区間として判定されず、バブル雑音以外の

50

定常な雑音区間で推定雑音が算出されることになる。定常雑音のパワーがバブル雑音のパワーよりも小さい場合には、バブル雑音の区間では推定雑音が過小評価され、抑圧不足が起こり、十分な抑圧ができないという問題点があった。

【 0 0 1 0 】

また例えば特許文献 2 では、推定音声区間のパワーは、長区間での短区間パワーの最大値として推定されており、音声パワーの分布が考慮されていない。人の声の特性や話し方によって変化する音声パワーの分布を考慮しない場合には、適切な抑圧係数を必ずしも算出できないという問題点がある。例えば音声パワーの分布が広い場合には、音声パワーの最大値が大きくても小さいパワーの音声が存在するため、抑圧を強めてしまうと音声が悪化する場合がある。

10

【 発明の開示 】

【 発明が解決しようとする課題 】

【 0 0 1 1 】

このように従来の技術においては、入力音声信号から雑音成分を除いた純粋な音声パワーの検出やその分布の推定などが行われていないため、背景雑音の推定を誤った場合には、適切な抑圧ゲインの算出ができないという問題点があった。

【 0 0 1 2 】

本発明の課題は上述の問題点に鑑み、入力音声信号に含まれる純粋な音声のパワーに関する情報を推定し、音声パワーの分布や音声パワーの存在範囲に基いて抑圧ゲインを算出することにより、様々な背景雑音が存在する場合にも適切な雑音抑圧を行うことができる雑音低減装置、および低減方法を提供することである。

20

【 課題を解決するための手段 】

【 0 0 1 3 】

図 1 は本発明の雑音低減装置の原理構成ブロック図である。同図は、入力音声信号を周波数分析して周波数領域の信号に変換する分析部 2 と、該周波数領域の信号を抑圧する抑圧部 3 と、抑圧された周波数領域の信号を用いて、抑圧された時間領域の信号を合成して出力する合成部 4 とを備える雑音低減装置 1 の原理構成ブロック図である。

【 0 0 1 4 】

本発明の雑音低減装置 1 は、さらに少なくとも音声情報推定手段 5、および抑圧ゲイン算出手段 6 を備える。音声情報推定手段 5 は、分析部 2 の出力する周波数領域の信号、例えばスペクトル振幅を用いて信号の抑圧ゲイン算出の基礎となる情報であって、少なくとも入力音声信号の中で雑音成分を除く純粋音声成分に対応すべき情報を音声情報として推定するものであり、抑圧ゲイン算出手段 6 は、音声情報推定手段 5 と分析部 2 の出力とに対応して、抑圧ゲインを算出し、抑圧部 3 に与えるものである。

30

【 0 0 1 5 】

発明の実施の形態においては、音声情報推定手段 5 が前述の純粋音声成分のパワーを推定することもでき、また入力された過去の複数の音声信号フレームに対する純粋音声の各周波数におけるパワー分布において、パワーの大きい方から積算されるサンプルの数が全サンプル数の一定割合となるパワーの平均値を推定することもできる。

【 0 0 1 6 】

40

この場合抑圧ゲイン算出手段 6 は、現在処理対象となっているフレーム k の周波数インデックス i に対応する前述のパワー平均値 $P_{MAX} k i$ と、フレーム k に対応するスペクトルパワー $P k i$ との差に基いて、フレーム k に対する抑圧ゲインを算出することもできる。

【 0 0 1 7 】

また実施の形態においては、音声情報推定手段 5 が前記抑圧ゲイン算出の基礎となる情報として、純粋音声成分に対応すべき情報としての純粋音声のパワー分布の推定値に加えて、入力音声信号としての雑音重畳音声信号のパワー分布を算出し、その算出結果を抑圧ゲイン算出手段 6 に与えることもできる。

【 0 0 1 8 】

50

この場合、音声情報推定手段 5 が過去に入力された複数の音声フレーム信号に対する純粋音声の各周波数におけるパワー分布において、パワーの大きい方から積算されるサンプルの数が全サンプル数のそれぞれ一定割合となるパワーの 2 つの平均値を用いて、純粋音声のパワー分布に相当する確率密度関数を推定することもでき、抑圧ゲイン算出手段 6 が音声情報推定手段 5 の出力としての純粋音声パワーの分布と、雑音重畳音声信号のパワー分布とのそれぞれについて、パワーの大きい方から積算されるサンプルの数が、全サンプル数の一定の割合となるようにパワー分布を複数の区間に分割し、その複数の各区間におけるパワーの平均値に基づいて抑圧ゲインを求めることもできる。

【0019】

また本発明の雑音低減装置は、前述の分析部 2、抑圧部 3、合成部 4、および音声情報推定手段 5 に加えて、さらに入力音声信号内の雑音成分のスペクトルを推定する雑音推定手段を備え、抑圧ゲイン算出手段が雑音推定手段、音声情報推定手段、および分析部の出力に対応して抑圧ゲインを算出し、抑圧部 3 に与える。

【0020】

この雑音低減装置においては、前述と同様に音声情報推定手段 5 が、前述の純粋音声成分のパワーを推定することもでき、また前述の過去複数の音声フレームに対する純粋音声パワーの分布において、パワーの大きい方から積算されるサンプルの数が全サンプル数の一定割合となるパワーの平均値を推定することもできる。

【0021】

この場合抑圧ゲイン算出手段 6 は、前述のパワー平均値 P_{MAXki} と、雑音推定手段の出力としての現フレームに対するスペクトル雑音 N_{ki} と、現フレームのスペクトルパワー P_{ki} の入力に対して、 P_{MAXki} と P_{ki} との差、および P_{MAXki} と N_{ki} との差に基づいて抑圧ゲインを算出することもできる。

【0022】

あるいは抑圧ゲイン算出手段 6 が、純粋音声パワーの下限を推定し、その推定結果を用いて、現フレームを含み、過去に入力された複数の音声フレーム信号において、非定常雑音を検出された頻度 H_{ki} を算出し、前述の P_{MAXki} 、 N_{ki} 、および P_{ki} の入力に対して、 P_{MAXki} と P_{ki} との差、 P_{MAXki} と N_{ki} との差、および H_{ki} に基づいて、抑圧ゲインを算出することもできる。

【0023】

次に本発明の雑音低減方法は、前述の分析部、抑圧部、および合成部を用いて雑音を低減するものであり、分析部の出力を用いて、信号の抑圧ゲイン算出の基礎となる情報であって、少なくとも入力音声信号の内で雑音を除く純粋音声成分に対応すべき情報を音声情報として推定し、その推定結果と前記分析部の出力とに対応して抑圧ゲインを算出し、前記抑圧部に与える方法が用いられる。

【0024】

また実施の形態における雑音低減方法として、前述の音声情報を推定し、入力音声信号の内の雑音成分のスペクトルを推定し、推定された音声情報、雑音スペクトル、および分析部の出力に対応して抑圧ゲインを算出し、抑圧部に与える方法が用いられる。

【0025】

実施の形態においては、これらの 2 つの方法にそれぞれ対応して、計算機にこの雑音低減方法を実行させるプログラムを用いることも、またそのようなプログラムが格納された可搬型記憶媒体を用いることもできる。

【発明の効果】

【0026】

本発明によれば、雑音の推定を行うことなく純粋音声のパワー情報を推定し、その分布や存在範囲に基づいて、抑圧ゲインを算出するために、雑音推定性能の影響を受けずに適切な音声抑圧を行うことができ、高品質な音声信号を得ることができる。さらに純粋音声のパワー分布に加えて、雑音重畳音声のパワー分布を用いて抑圧ゲインを算出することもでき、音声区間に重畳された雑音パワーの影響を考慮した抑圧ゲインを算出できるため、非

10

20

30

40

50

定常な雑音が重畳した場合にも、雑音区間で推定された雑音推定値を用いる従来の手法に比べてより正確に抑圧ゲインを求めることが可能となる。

【 0 0 2 7 】

また本発明によれば、純粹音声のパワー情報の推定値に加えて、さらに雑音を推定し、その結果を用いて抑圧ゲインを算出するため、純粹音声のパワー分布やその存在範囲と、推定された雑音パワーに基いて抑圧ゲインを算出することができ、非定常な雑音が重畳した場合でも、単に雑音区間で算出される雑音推定値を用いる従来手法に比べてより正確に抑圧ゲインを求めることができる。さらに非定常雑音の頻度を用いて抑圧ゲインを算出することも可能となり、より正確に雑音を抑圧でき、例えば移動体通信における通信品質の向上に寄与するところが多い。

10

【発明を実施するための最良の形態】

【 0 0 2 8 】

図 2 は、本発明の第 1 の実施形態における音声信号の雑音低減装置の構成ブロック図である。同図において、分析部 11 はフレームごとの入力信号、すなわち雑音が重畳された音声信号の入力に対して、ハミング窓などの時間窓を掛けた上で、高速フーリエ変換 F F T などを用いて入力フレームの分析を行い、スペクトル振幅（＝振幅スペクトル）とスペクトル位相（＝位相スペクトル）の算出を行う。この F F T や入力信号に対する窓掛けなどについては次の文献に詳細に説明されている。

【非特許文献 2】辻井、鎌田「ディジタル信号処理シリーズ第 1 巻 ディジタル信号処理」 94～120 頁、昭晃堂

20

【非特許文献 3】C u r t i s R o a d s 著、青柳他訳・監修「コンピュータ音楽」 452～457 頁、東京電機大学出版局

【 0 0 2 9 】

分析部 11 の出力としてのスペクトル振幅は、音声推定部 12、抑圧ゲイン算出部 14、および抑圧部 15 に与えられる。音声推定部 12 は入力信号のスペクトル振幅を用いて雑音が重畳された入力音声信号のうちで雑音を除く成分、すなわち純粹音声信号に対応すべき情報であって、抑圧ゲインの算出のために使用される音声情報を推定するものである。この第 1 の実施形態においては、図 20 で説明したような雑音推定を行って抑圧ゲインを算出する代わりに、純粹な音声信号に対応する音声情報を推定して、抑圧ゲインの算出を行うところに基本的な特徴がある。

30

【 0 0 3 0 】

スペクトルパワー記憶部 13 は、例えば過去の 100 フレームに対応して、スペクトルパワーの値を記憶し、後述するように音声推定部 12、および抑圧ゲイン算出部 14 に与えるものである。

【 0 0 3 1 】

抑圧ゲイン算出部 14 は、音声推定部 12 の出力としての音声情報と入力信号のスペクトル振幅とを用いて、スペクトル振幅を調整するための抑圧ゲインを算出し、抑圧部 15 は算出された抑圧ゲインの値と入力信号のスペクトル振幅とを用いて、抑圧されたスペクトル振幅を算出し、合成部 16 に与える。

【 0 0 3 2 】

40

合成部 16 は、抑圧されたスペクトル振幅と分析部 11 が出力するスペクトル位相とを用いて、逆高速フーリエ変換 I F F T によって周波数軸上の信号を時間軸上の信号に変換し、オーバーラップ加算によって、直前フレームの時間軸上の抑圧処理音声と重ね合わせ、抑圧された出力音声信号として出力する。雑音低減装置 10 の動作は以上であるが、合成部 16 の出力信号は、例えば音声符号化部 17 に与えられ、符号化結果が送信部 18 によって送信されるという形式で、音声通信方式に適用することが可能となる。

【 0 0 3 3 】

なお合成部 16 によって、時間軸上に変換された信号と直前フレームの時間軸上の抑圧処理音声とがオーバーラップ加算によって重ね合わせられる理由は、一般に F F T における窓掛けによって窓の外側の信号が小さくなることを補正するためであり、一般的に公知

50

の従来技術として行われるものである。

【0034】

図3は、図2の雑音低減装置による雑音低減処理の全体的なフローチャートである。同図においてステップS1で入力信号が1フレーム分入力され、ステップS2でハミング窓などの時間窓掛けが行われた上でFFT分析が行われ、スペクトル分析結果としてのスペクトル振幅 SA_{ki} とスペクトル位相 SP_{ki} が得られる。ここで、 k はフレームのインデックスであり、 i は周波数（帯域）のインデックスである。

【0035】

続いてステップS3で音声情報の推定が行われる。ここでは入力信号のスペクトル振幅 SA_{ki} を用いて抑圧ゲイン算出の基礎となる情報としての音声情報が算出されるが、その詳細については後述する。ステップS4で音声情報算出結果から抑圧ゲイン G_{ki} が算出され、ステップS5で抑圧された振幅スペクトル SA'_{ki} が次の(1)式を用いて算出される。

【0036】

$$SA'_{ki} = SA_{ki} \cdot G_{ki} \quad 0 \leq i < N \quad \dots \quad (1)$$

【0037】

ステップS6で抑圧された振幅スペクトル SA'_{ki} とスペクトル位相 SP_{ki} とを用いて、IFFTが行われ、オーバーラップ加算がされて音声合成される。ステップS7で入力フレームの全てに対する処理が終了したか否かが判定され、まだ終了していない場合にはステップS1以降の処理が繰り返され、全てのフレームに対する処理が終了したと判定された時点で処理を終了する。

【0038】

図4は、図3のステップS2のスペクトル分析の詳細処理フローチャートである。同図において処理が開始されると、まずステップS11で入力信号 x_{kt} に対して窓関数 H_t を使って、次の(2)式によって窓掛け信号 w_{kt} が得られる。

【0039】

$$w_{kt} = H_t \cdot x_{kt} \quad t = 0, \dots, 2N-1 \quad \dots \quad (2)$$

【0040】

次にステップS12で、窓掛け信号に対するFFT処理が行われ、結果として実数部 XR_{ki} 、虚数部 XI_{ki} が得られる。続いてステップS13で次の(3)式によってスペクトル振幅 SA_{ki} が求められる。

【0041】

$$SA_{ki} = (XR_{ki}^2 + XI_{ki}^2)^{1/2} \quad 0 \leq i < N \quad \dots \quad (3)$$

【0042】

さらにステップS14で、次の(4)式によってスペクトル位相 SP_{ki} が算出されて処理を終了する。

【0043】

$$SP_{ki} = \tan^{-1}(XI_{ki} / XR_{ki}) \quad 0 \leq i < N \quad \dots \quad (4)$$

【0044】

なお、これらの式において $2N$ はFFTを行う点数、例えば128や256であり、窓関数 H_t は、例えばハミング窓である。

【0045】

図5は、図3の音声情報算出処理（ステップS3）の1つの実施例としての、純粹音声のパワー分布において、パワーの大きい方から、積算されるサンプルの数が全サンプル数の一定割合となるパワーの平均値を音声情報として推定する処理の詳細フローチャートである。同図において処理が開始されると、まずステップS16で現在処理対象となっている現フレームのスペクトルパワー P_{ki} が次の(5)式によって算出される。すなわち、 k フレームにおける周波数（帯域） i 毎にスペクトル振幅の2乗が求められ、スペクトルパワーとして算出される。

【0046】

10

20

30

40

50

$$P_{ki} = S A_{ki}^2 \quad 0 \leq i < N \quad \dots \quad (5)$$

【0047】

続いてステップS17で任意の監視期間、例えば現フレームを含め監視期間100フレームに対応して、算出されたスペクトルパワーを用いて周波数(帯域)インデックス*i*毎にスペクトルパワーの分布を求め、例えば上位10%となるスペクトルパワー、すなわち10個のスペクトルパワーの値が抽出される。そしてステップS18で、この例えば上位10%、すなわち上位一定割合のスペクトルパワーの平均値 $P_{MAX\ ki}$ の算出が行われ、これを音声推定部12が出力する音声情報として処理を終了する。

【0048】

図6は、図3の抑圧ゲイン算出処理(ステップS4)の詳細フローチャートである。同図において処理が開始されると、まずステップS20で抑圧ゲイン G_{ki} を決定する関数*f*における引数としての d_{ki} が、次の(6)式によって算出される。

【0049】

$$d_{ki} = P_{MAX\ ki} - P_{ki} \quad 0 \leq i < N \quad \dots \quad (6)$$

続いてステップS21で抑圧ゲイン G_{ki} の算出が、次の(7)式を用いて行われ、処理を終了する。

【0050】

$$G_{ki} = f(d_{ki}) \quad 0 \leq i < N \quad \dots \quad (7)$$

【0051】

図7は、この抑圧ゲイン算出関数*f*の例を示す。この関数*f*は、音声パワーの分布の位置に対応して抑圧ゲインを決めるものであり、例えばあらかじめ音声の抑圧と雑音低減効果のバランスから経験的に求められるものである。図7では、この関数*f*の引数 d_{ki} が小さいほど抑圧ゲイン G_{ki} を大きくして実際の抑圧を小さくし、 d_{ki} が大きくなるほど抑圧ゲインを小さくして実際の抑圧を大きくする形式となっている。

【0052】

図8は、図7の抑圧ゲイン算出関数*f*の引数 d_{ki} が小さい範囲で、抑圧ゲイン G_{ki} を大きくする理由の説明図である。一般的に入力音声信号は、雑音重畳信号であり、純粋音声の成分と雑音成分とが含まれるが、純粋音声成分のパワーが平均的に雑音パワーより大きい場合には、雑音重畳入力信号のパワーが大きい区間では、入力信号パワーによって純粋音声パワーを近似することができる。従って現フレームの入力信号パワー P_{ki} と、例えば100フレームに対応して求められた音声パワーの上位一定割合、例えば10%以内のパワー平均値 $P_{MAX\ ki}$ との差が小さい場合には、雑音重畳音声信号に含まれる純粋音声パワーが大きく、雑音成分の影響は小さいと考えられるため、抑圧ゲインを大きくする、すなわち、実際の抑圧を小さくすることが適切となる。また実際の入力信号、すなわち雑音重畳音声信号ではなく、あらかじめ純粋音声パワーの存在幅を実験的に算出しておいたり、分布を仮定するなどの方法で、図8に点線で示す純粋な音声パワーの分布を推定することができ、その分布から求められる $P_{MAX\ ki}$ と現フレームの P_{ki} との差から、 d_{ki} を算出することもできる。

【0053】

次に図3のステップS3における音声情報算出処理の異なる実施例と、それに対応するステップS4における抑圧ゲイン算出処理について図9 - 図11を用いて説明する。図9は、音声情報算出処理の異なる実施例のフローチャートである。同図において処理が開始されると、まずステップS23で前述の(3)式によって求められたスペクトル振幅 $S A_{ki}$ が入力され、ステップS24で(5)式によってスペクトルパワー P_{ki} が周波数(帯域)*i*毎に算出される。

【0054】

続いてステップS25で、図5におけると同様にして雑音重畳音声信号のスペクトルパワーのうちで大きさがそれぞれ上位一定割合となるスペクトルパワーの2つの平均値 $P_{MAX\ 1\ ki}$ 、および $P_{MAX\ 2\ ki}$ の算出が行われる。すなわち前述のように、例えば100フレームに対応して、周波数のインデックスが*i*のスペクトルパワーの大きさのうち

10

20

30

40

50

で、上位 $\times 1\%$ (ガウス分布の $a_1 \cdot \sigma$ の位置に相当) のパワーの平均値となるように P_{MAX1ki} を算出し、 P_{MAX2ki} を上位 $\times 2\%$ (ガウス分布の $a_2 \cdot \sigma$ の位置に相当) のパワーの平均値となるように算出する。ここで例えば a_1 の方が a_2 より大きいものとし、また σ は標準偏差を示す。

【 0 0 5 5 】

続いてステップ S 2 6 で、周波数のインデックス i 毎の純粋音声パワーの分布がガウス分布と仮定され、ガウス分布の標準偏差が次の (8) 式によって算出される。

【 0 0 5 6 】

$$\sigma_{ki} = (P_{MAX1ki} - P_{MAX2ki}) / (a_1 - a_2) \quad 0 \leq i < N \quad \dots (8) \quad 10$$

続いてステップ S 2 7 で、ガウス分布の平均 m が次の (9) 式によって算出される。

【 0 0 5 7 】

$$m_{ki} = P_{MAX1ki} - a_1 \cdot \sigma_{ki} \quad 0 \leq i < N \quad \dots (9)$$

【 0 0 5 8 】

これによって、純粋音声パワーに対する標準偏差と平均を用いて、音声パワーの確率密度関数を次の (1 0) 式によって求めることができる。なおここで x は、純粋音声パワーを示す。

【 0 0 5 9 】

$$P_{1ki}(x) = \{ 1 / (2\pi)^{1/2} \} \exp [- (x - m_{ki})^2 / 2\sigma_{ki}^2] \quad 0 \leq i < N \quad \dots (10) \quad 20$$

【 0 0 6 0 】

この例では純粋音声のパワー分布をガウス分布と仮定したが、純粋音声パワーのヒストグラムを算出することにより、確率密度関数を求めることもできる。

【 0 0 6 1 】

続いて図 9 のステップ S 2 8 で、雑音重畳入力信号のスペクトルパワーを監視して、そのヒストグラム $P_{2ki}(x)$ が作成され、ステップ S 2 9 で音声情報として純粋音声パワーの確率密度関数 $P_{1ki}(x)$ と雑音重畳音声パワーのヒストグラム $P_{2ki}(x)$ が出力されて処理を終了する。

【 0 0 6 2 】

ここでステップ S 2 5 における P_{MAX1ki} と P_{MAX2ki} の算出の具体例についてさらに説明する。前述の a_1 の値を 3、 a_2 の値を 2 として、 P_{MAX1ki} を上位から 0.3%、 P_{MAX2ki} を上位から 4.6% の位置のパワーの値となるように算出する。

【 0 0 6 3 】

すなわち P_{MAX1ki} の算出においては、例えば過去 1000 フレームのスペクトルパワーを大きさ順に並べ、大きいものから 6 つ選ぶ、すなわち上位 0.6% のものを選び、選んだスペクトルパワーの平均値を求めることで、また P_{MAX2ki} の算出では、過去 1000 フレームスペクトルパワーの大きいものから 92 個、すなわち上位 9.2% のものを選び、選んだスペクトルパワーの平均を求めることによって算出することができる。

【 0 0 6 4 】

図 10 は、図 9 の音声情報算出処理に対応する抑圧ゲイン算出処理の詳細フローチャートである。同図において処理が開始されると、図 9 の処理によって出力された純粋音声パワーの確率密度関数 $P_{1ki}(x)$ と雑音重畳音声信号のヒストグラム $P_{2ki}(x)$ がステップ S 3 1 で入力され、ステップ S 3 2 で (純粋) 音声パワーと雑音重畳音声パワーの分布において上位から $\eta\%$ 毎に分布を区分し、区分区間毎にパワーの平均値を算出する処理が行われる。

【 0 0 6 5 】

図 11 は、この処理の説明図である。例えば、雑音重畳音声パワーの分布において、過去の 100 フレームを用いて上位から 10% 毎のパワーの平均値を算出する場合を例とし

10

20

30

40

50

て説明する。純粋音声パワーについては、あらかじめ雑音が混入しない音声信号を用いて同様に算出することができる。

【 0 0 6 6 】

まず過去 1 0 0 フレームの雑音重畳音声パワーをその大きさ順に並べ、上位から 1 0 個毎の雑音重畳音声パワーの平均値 V_{2n} の算出を行う。すなわち上位 1 位から 1 0 個の雑音重畳音声パワーの平均値を V_{2_1} 、上位 1 1 位から 1 0 個の雑音重畳音声パワーの平均値を V_{2_2} 、上位 9 1 位から 1 0 個の雑音重畳音声パワーの平均値 $V_{2_{10}}$ として求める。純粋な音声パワーについても、例えば上位から n 番目の区間に対応して音声パワーの平均値 V_{1n} が求められる。

【 0 0 6 7 】

10

図 1 0 のステップ S 3 3 で、区間毎の抑圧ゲイン G_{ikn} の算出が行われる。この処理では純粋な音声パワーの分布と雑音重畳音声パワーの分布において、同じ区間 n においては、その区間の（純粋な）音声パワーに雑音が重畳して雑音重畳音声パワーになったものとし、次の（ 1 1 ）、（ 1 2 ）式を用いて、雑音重畳音声パワーの n 番目の区間に対する平均値 V_{2n} に対する抑圧ゲインは（ 1 3 ）式によって与えられるものとする。

【 0 0 6 8 】

$$V_{1n} = 10 \log_{10} (\text{音声パワー}) \quad \dots \quad (11)$$

$$V_{2n} = 10 \log_{10} (\text{音声パワー} + \text{雑音パワー}) \quad \dots \quad (12)$$

【 0 0 6 9 】

【数 1】

20

$$G_{ikn} = \left(10 \frac{V_{2n} - V_{1n}}{10} \right)^{\frac{1}{2}} \quad \dots \quad (13)$$

【 0 0 7 0 】

ステップ S 3 3 で求められた抑圧ゲイン G_{ikn} は区間毎に求められた離散的な値であり、実際の雑音重畳音声パワー信号 x の関数として抑圧ゲインを算出するために、ステップ S 3 4 で G_{ikn} を次の（ 1 4 ）式を用いて補間し、抑圧ゲイン関数の算出が行われる。

【 0 0 7 1 】

30

【数 2】

$$G_{ik}(x) = \frac{G_{ikn} - G_{ik}(n-1)}{V_{2n} - V_{2(n-1)}} \{x - V_{2(n-1)}\} \quad \dots \quad (14)$$

ここで $V_{2(n-1)}$ は、 $n-1$ 番目の区間の V_2 の値を示す。

【 0 0 7 2 】

続いてステップ S 3 5 で、現在のフレームの雑音重畳音声パワー x の値を用いて抑圧ゲイン $G_{ik}(x)$ の値が算出され、ステップ S 3 6 でその値が出力されて処理を終了する。

40

【 0 0 7 3 】

続いて本発明の第 2 の実施形態について説明する。図 1 2 は、第 2 の実施形態における雑音低減装置の構成ブロック図である。同図を、図 2 の第 1 の実施形態における構成と比較すると雑音推定部 1 9 が追加され、抑圧ゲイン算出部 1 4 が音声推定部 1 2 の出力する音声情報に加えて、雑音推定部 1 9 の出力としての推定雑音を用いて抑圧ゲインを算出する点だけが異なっている。雑音推定部 1 9 は、分析部 1 1 の出力するスペクトル振幅を用いて、入力信号に含まれるスペクトル雑音（＝雑音スペクトル）を推定するものであるが、スペクトル振幅に代わって、時間軸上の入力信号を用いて雑音を推定することも可能である。

【 0 0 7 4 】

50

図１３は、第２の実施形態における雑音低減処理の全体フローチャートである。同図を第１の実施形態における図３と比較すると、ステップＳ５３でスペクトル雑音の推定が行われ、ステップＳ５４でその推定結果に対応して音声情報が算出され、ステップＳ５５で抑圧ゲインの算出が行われる点だけが異なっている。

【００７５】

図１４は、図１３のステップＳ５３におけるスペクトル雑音推定処理の詳細フローチャートである。同図において処理が開始されると、まずステップＳ６１で前述の（５）式を用いて、スペクトルパワー P_{ki} が算出され、ステップＳ６２で音声区間と雑音区間との判定処理が行われる。この判定には公知の従来技術を用いることができ、例えば長期間のフレームパワーの平均と現フレームのパワーとの差を監視する方法や、相関係数を算出する

10

【００７６】

ステップＳ６３で、雑音区間であるか否かが判定され、雑音区間でない場合には、そのフレームに対する処理を終了し、雑音区間である場合にはステップＳ６４で推定スペクトル雑音 N_{ki} の更新処理が行われる。

【００７７】

この更新処理では、現在のフレーム（雑音フレーム）のスペクトルパワー（雑音スペクトルパワー）と、それまでに算出された過去の雑音スペクトルパワーに、それぞれの寄与率を乗算して、雑音スペクトルパワーの更新が行われる。これによってフレーム毎のパワー変動の高周波成分を取り除くことができる。ここでは、次の（１５）式を用いて推定スペクトル雑音の更新が行われる。ここで ξ は、前述の寄与率に対応する定数である。

20

【００７８】

$$N_{ki} = \xi \cdot P_{ki} + (1 - \xi) N_{(k-1)i} \quad 0 \leq i < N \quad \dots \quad (15)$$

ここで $N_{(k-1)i}$ は、 $k-1$ 番目のフレームの i 番目の帯域の雑音スペクトラムパワーを示す。

【００７９】

図１５は、図１３のステップＳ５５における抑圧ゲイン算出処理の詳細フローチャートである。なおステップＳ５４における音声情報算出処理は第１の実施形態における、例えば図５におけると同様に行われるものとする。

30

【００８０】

図１５において処理が開始されると、まずステップＳ６６で周波数（帯域）毎の現在のフレームのパワー P_{ki} と、雑音重畳音声信号のスペクトルパワーの内、大きさが上位一定割合となるスペクトルパワーの平均値 P_{MAXki} 、すなわち音声推定部１２の出力する音声情報と、雑音推定スペクトル N_{ki} 、すなわち雑音推定部１９の出力とが入力され、ステップＳ６７で次の（１６）式を用いて d_{1ki} が、ステップＳ６８で（１７）式を用いて d_{2ki} が算出され、ステップＳ６９で抑圧ゲイン G_{ki} が次の（１８）式を用いて算出され、ステップＳ７０で算出された抑圧ゲインが出力されて処理を終了する。

【００８１】

$$d_{1ki} = P_{MAXki} - P_{ki} \quad 0 \leq i < N \quad \dots \quad (16)$$

40

$$d_{2ki} = P_{MAXki} - N_{ki} \quad 0 \leq i < N \quad \dots \quad (17)$$

$$G_{ki} = g(d_{1ki}, d_{2ki}) \quad 0 \leq i < N \quad \dots \quad (18)$$

【００８２】

図１６は、抑圧ゲイン G_{ki} を算出するための、（１８）式によって与えられる関数 g の引数としての d_{1ki} と d_{2ki} との説明図である。同図において雑音重畳音声パワーの上位一定割合のパワースペクトルの平均値 P_{MAXki} と現在のフレームパワー P_{ki} との差 d_{1ki} は、現在のフレームに含まれる純粋な音声パワーの大きさに相当し、 P_{MAXki} と定常雑音の推定スペクトルのパワー N_{ki} との差 d_{2ki} は、雑音重畳音声パワーの分布と定常雑音パワーの分布との間の距離に対応する。ここで定常雑音パワーの分布に対してはそのピークの位置が用いられているのに対して、雑音重畳音声パワーの分布

50

に対してはピークの位置が用いられていないが、ここでは d_{2ki} を 2 つのパワーの分布の距離を表すものとして定義する。

【0083】

本実施形態では、 d_{1ki} と d_{2ki} の 2 つの値を用いることによって、純粋な音声パワー情報と雑音パワー情報の両方を考慮して抑圧ゲインを決めることにする。すなわち d_{1ki} の値が大きいほど、純粋な音声パワーが小さくなるために抑圧ゲインも小さくするように、また d_{2ki} が大きいほど、雑音重畳音声パワーの分布と定常雑音パワーの分布が離れ、含まれる雑音パワーが小さくなるため抑圧ゲインが大きくなるように、例えば次の (19) 式を用いて抑圧ゲイン G_{ki} を与える関数 g を設定するものとする。

【0084】

$$g(d_{1ki}, d_{2ki}) = \tau - \kappa \cdot d_{1ki} + \mu \cdot d_{2ki} \quad 0 \leq i < N \quad \dots (19)$$

ここで、 τ 、 κ 、 μ は正の値をとる係数

【0085】

図 17 は、第 2 の実施形態における抑圧ゲイン算出処理の異なる実施例のフローチャートである。同図において処理が開始されると、まずステップ S72 で、図 15 のステップ S66 におけると同様に P_{ki} 、 P_{MAXki} 、および N_{ki} の入力が行われ、ステップ S73 で d_{1ki} 、ステップ S74 で d_{2ki} が前述と同様に算出され、ステップ S75 で純粋な音声パワーの下限 P_{MINki} の算出処理が行われる。

【0086】

図 18 は、この抑圧ゲイン算出処理の説明図である。同図において純粋な音声パワーの分布の下限の位置が P_{MINki} の値として次の (20) 式によって推定される。

【0087】

$$P_{MINki} = P_{MAXki} - \psi_{ki} \quad 0 \leq i < N \quad \dots (20)$$

【0088】

この (20) 式においては、入力レベルが一定であれば、純粋な音声パワーの存在幅 (最大パワーと最小パワーの差) ψ_{ki} が一定であることを仮定している。その存在幅の値は、あらかじめ純粋な音声パワーの分布から調べることもでき、あるいは純粋な音声パワーの分布をガウス分布と仮定し、入力信号のパワーを観測して求められる標準偏差 σ にある定数を乗算して算出することも可能である。

【0089】

続いて図 17 のステップ S76 で、非定常雑音の頻度 H_{ki} の算出処理が行われる。この処理では、図 18 の定常雑音パワーの分布の位置を示す N_{ki} と、雑音判定区間におけるパワーの幅を表す値としての λ との和が求められ、現在のフレームに対応する P_{ki} がこの $N_{ki} + \lambda$ と純粋な音声パワーの分布における下限 P_{MINki} の間にあるか否かによって、各フレームに非定常雑音が含まれるかどうか頻度として調べられる。すなわちパルス雑音のような非定常雑音があるか否かが、それぞれのフレームについて調べられ、入力されるフレームに対応して次の (21) 式、または (22) 式を用いて、頻度 H_{ki} の更新が行われる。

【0090】

$$H_{ki} = [\{ H_{(k-1)i} \cdot (N_{ki} + \lambda - P_{MINki}) / (P_{ki} - P_{MINki}) \} + 1] / k \quad \dots (21)$$

$$H_{ki} = \{ H_{(k-1)i} \cdot (P_{ki} - P_{MINki}) / (P_{ki} - P_{MINki}) \} / k \quad \dots (22)$$

ここで $H_{(k-1)i}$ は前フレームに対する頻度、 $0 \leq i < N$

【0091】

すなわち $N_{ki} + \lambda$ は雑音の上限パワーを表し、その値と純粋な音声パワーの分布の下限の値 P_{MINki} との間に P_{ki} が存在するフレームが、今までの入力フレームの総数に対して何割となるかによって、非定常雑音の頻度 H_{ki} が算出される。

【0092】

10

20

30

40

50

続いて図１７のステップＳ７７で、次の（２３）式を用いて抑圧ゲイン G_{ki} が算出され、ステップＳ７８で、その抑圧ゲインが出力されて処理が終了される。

【００９３】

$$G_{ki} = h(d_{1ki}, d_{2ki}, H_{ki}) \quad 0 \leq i < N \quad \dots \quad (23)$$

【００９４】

この抑圧ゲイン G_{ki} を算出するための（２３）式における関数 h は、例えば次の（２４）式によって決定される。

【００９５】

$$h(d_{1ki}, d_{2ki}, H_{ki}) = \tau - \kappa \cdot d_{1ki} + \mu \cdot d_{2ki} - \nu \cdot H_{ki} \quad 0 \leq i < N \quad \dots \quad (24) \quad 10$$

ここで τ 、 κ 、 μ 、 ν は正の値をとる係数

【００９６】

図１８において図１６と同様に、 d_{1ki} が大きいほど純粋な音声のパワーは小さくなるため、抑圧ゲインを小さくするように、また d_{2ki} が大きいほど雑音パワーが小さくなるため、抑圧ゲインを大きくするように関数 h の設定が行われるが、それに加えて非定常雑音の頻度 H_{ki} が大きいほど非定常雑音が多く存在するため、抑圧ゲインを小さくするように h の設定が行われる。

【００９７】

以上において本発明の雑音低減装置、および低減方法についてその詳細を説明したが、この雑音低減装置は当然プロセッサや一般的なコンピュータシステムとして構成することが可能である。図１９はそのようなコンピュータシステム、すなわちハードウェア環境の構成ブロック図である。 20

【００９８】

図１９においてコンピュータシステムは中央処理装置（ＣＰＵ）２０、リードオンリメモリ（ＲＯＭ）２１、ランダムアクセスメモリ（ＲＡＭ）２２、通信インタフェース２３、記憶装置２４、入出力装置２５、可搬型記憶媒体の読取り装置２６、およびこれらの全てが接続されたバス２７によって構成されている。

【００９９】

記憶装置２４としては、ハードディスク、磁気ディスクなど様々な形式の記憶装置を使用することができ、このような記憶装置２４、またはＲＯＭ２１に図３～図６、図９、図１０、図１３～図１５、および図１７のフローチャートに示されたプログラムなどが格納され、そのようなプログラムがＣＰＵ２０によって実行されることにより、本実施形態における純粋音声に関する情報の推定、その情報に対応する雑音の抑圧などが可能となる。 30

【０１００】

このようなプログラムは、プログラム提供者２８側からネットワーク２９、および通信インタフェース２３を介して、例えば記憶装置２４に格納されることも、また市販され、流通している可搬型記憶媒体３０に格納され、読取り装置２６にセットされて、ＣＰＵ２０によって実行されることも可能である。可搬型記憶媒体３０としてはＣＤ－ＲＯＭ、フレキシブルディスク、光ディスク、光磁気ディスクなど様々な形式の記憶媒体を使用することができ、このような記憶媒体に格納されたプログラムが読取り装置２６によって読み取られることにより、本実施形態におけるパルス雑音を含む各種雑音の抑圧などが可能となる。 40

【０１０１】

（付記１） 入力音声信号を周波数分析して周波数領域の信号に変換する分析部と、該周波数領域の信号を抑圧する抑圧部と、該抑圧された周波数領域の信号を用いて抑圧された時間領域の信号を合成して出力する合成部とを備える雑音低減装置において、

該分析部の出力を用いて、信号の抑圧ゲイン算出の基礎となる情報であって、少なくとも入力音声信号の中で雑音成分を除く純粋音声成分に対応すべき情報を音声情報として推定する音声情報推定手段と、 50

該音声情報推定手段と前記分析部との出力に対応して、前記抑圧ゲインを算出して前記抑圧部に与える抑圧ゲイン算出手段とを備えることを特徴とする雑音低減装置。

【0102】

（付記2） 前記音声情報推定手段が、前記雑音成分を除く純粋音声成分のパワーを推定することを特徴とする付記1記載の雑音低減装置。

（付記3） 前記音声情報推定手段が、過去に入力された複数の音声フレームに対する前記純粋音声の各周波数におけるパワー分布において、パワーの大きい方から積算されるサンプルの数が全サンプル数の一定割合となるパワーの平均値を推定することを特徴とする付記1記載の雑音低減装置。

【0103】

（付記4） 前記抑圧ゲイン算出手段が、現在処理対象となっているフレームkの周波数インデックスiに対応する前記パワー平均値 P_{MAXki} と、該フレームkに対応するスペクトルパワー P_{ki} との差に基いて、該フレームkに対する抑圧ゲインを算出することを特徴とする付記3記載の雑音低減装置。

【0104】

（付記5） 前記音声情報推定手段が、前記抑圧ゲイン算出の基礎となる情報として、前記純粋音声成分に対応すべき情報としての純粋音声のパワー分布に加えて、前記入力音声信号としての雑音重畳音声信号のパワー分布を算出して前記抑圧ゲイン算出手段に与えることを特徴とする付記1記載の雑音低減装置。

【0105】

（付記6） 前記音声情報推定手段が、過去に入力された複数の音声フレーム信号に対する純粋音声の各周波数におけるパワー分布において、パワーの大きい方から積算されるサンプルの数が全サンプル数のそれぞれ一定割合となるパワーの2つの平均値を用いて、前記純粋音声のパワー分布に相当する確率密度関数を推定することを特徴とする付記5記載の雑音低減装置。

【0106】

（付記7） 前記抑圧ゲイン算出手段が、前記音声情報推定手段の出力としての純粋音声パワーの分布と、雑音重畳音声信号のパワー分布とのそれぞれについて、パワーの大きい方から積算されるサンプルの数が全サンプル数の一定割合となるようにパワー分布を複数の区間に分割し、該複数の各区間におけるパワーの平均値に基いて前記抑圧ゲインを求めることを特徴とする付記5記載の雑音低減装置。

【0107】

（付記8） 入力音声信号を周波数分析して周波数領域の信号に変換する分析部と、該周波数領域の信号を抑圧する抑圧部と、該抑圧された周波数領域の信号を用いて抑圧された時間領域の信号を合成して出力する合成部とを備える雑音低減装置において、

前記入力音声信号の内の雑音成分のスペクトルを推定する雑音推定手段と、

前記分析部の出力を用いて、信号の抑圧ゲイン算出の基礎となる情報であって、少なくとも入力音声信号の内で雑音を除く純粋音声成分に対応すべき情報を音声情報として推定する音声情報推定手段と、

該雑音推定手段、音声情報推定手段、および前記分析部の出力に対応して、前記抑圧ゲインを算出して前記抑圧部に与える抑圧ゲイン算出手段とを備えることを特徴とする雑音低減装置。

【0108】

（付記9） 前記音声情報推定手段が、前記雑音成分を除く純粋音声成分のパワーを推定することを特徴とする付記8記載の雑音低減装置。

（付記10） 前記音声情報推定手段が、過去に入力された複数の音声フレームに対する前記純粋音声の各周波数におけるパワー分布においてパワーの大きい方から積算されるサンプルの数が全サンプル数の一定割合となるパワーの平均値を推定すること特徴とする付記8記載の雑音低減装置。

【0109】

10

20

30

40

50

(付記 1 1) 前記抑圧ゲイン算出手段が、現在処理対象となっているフレーム k の周波数インデックス i に対応する前記パワー平均値 $P_{MAX\ k\ i}$ 、前記雑音推定手段の出力としての現フレームに対するスペクトル雑音 $N_{k\ i}$ 、現フレームのスペクトルパワー $P_{k\ i}$ の入力に対して、 $P_{MAX\ k\ i}$ と $P_{k\ i}$ との差、および $P_{MAX\ k\ i}$ と $N_{k\ i}$ との差に基いて抑圧ゲインを算出することを特徴とする付記 1 0 記載の雑音低減装置。

【0 1 1 0】

(付記 1 2) 前記抑圧ゲイン算出手段が、純粹音声パワーの下限を推定し、該推定結果を用いて、現フレームを含み過去に入力された複数の音声フレーム信号において非定常雑音を検出された頻度を算出し、現在処理対象となっているフレーム k の周波数インデックス i に対応する前記パワー平均値 $P_{MAX\ k\ i}$ 、該フレーム k に対応するスペクトルパワー $P_{k\ i}$ 、および雑音推定手段の出力としての現フレームに対するスペクトル雑音 $N_{k\ i}$ の入力に対して、 $P_{MAX\ k\ i}$ と $P_{k\ i}$ との差、 $P_{MAX\ k\ i}$ と $N_{k\ i}$ との差、および算出された頻度に基づいて抑圧ゲインを算出することを特徴とする付記 1 0 記載の雑音低減装置。

10

【0 1 1 1】

(付記 1 3) 入力音声信号を周波数分析して周波数領域の信号に変換する分析部と、該周波数領域の信号を抑圧する抑圧部と、抑圧された周波数領域の信号を用いて抑圧された時間領域の信号を合成して出力する合成部とを用いて雑音を低減する雑音低減方法において、

該分析部の出力を用いて、信号の抑圧ゲイン算出の基礎となる情報であって、少なくとも入力音声信号の中で雑音成分を除く純粹音声成分に対応すべき情報を音声情報として推定し、

20

該推定された音声情報と前記分析部の出力とに対応して抑圧ゲインを算出して前記抑圧部に与えることを特徴とする雑音低減方法。

【0 1 1 2】

(付記 1 4) 入力音声信号を周波数分析して周波数領域の信号に変換する分析部と、該周波数領域の信号を抑圧する抑圧部と、抑圧された周波数領域の信号を用いて抑圧された時間領域の信号を合成して出力する合成部とを用いて雑音を低減する雑音低減方法において、

前記入力音声信号のうちの雑音成分のスペクトルを推定し、

30

該分析部の出力を用いて、信号の抑圧ゲイン算出の基礎となる情報であって、少なくとも入力音声信号の中で雑音成分を除く純粹音声成分に対応すべき情報を音声情報として推定し、

該推定された雑音成分スペクトル、音声情報、および前記分析部の出力に対応して抑圧ゲインを算出して前記抑圧部に与えることを特徴とする雑音低減方法。

【0 1 1 3】

(付記 1 5) 入力音声信号を周波数分析して周波数領域の信号に変換する分析手順と、該周波数領域の信号を抑圧する抑圧手順と、抑圧された周波数領域の信号を用いて抑圧された時間領域の信号を合成して出力する合成手順を実行して雑音を低減する計算機によって使用されるプログラムにおいて、

40

該分析手順の処理結果を用いて、信号の抑圧ゲイン算出の基礎となる情報であって、少なくとも入力音声信号の中で雑音成分を除く純粹音声成分に対応すべき情報を音声情報として推定する手順と、

該推定された音声情報と前記分析部の出力とに対応して抑圧ゲインを算出して前記抑圧手順に与える手順とを計算機に実行させるためのプログラム。

【0 1 1 4】

(付記 1 6) 入力音声信号を周波数分析して周波数領域の信号に変換する分析手順と、該周波数領域の信号を抑圧する抑圧手順と、抑圧された周波数領域の信号を用いて抑圧された時間領域の信号を合成して出力する合成手順を実行して雑音を低減する計算機によって使用されるプログラムにおいて、

50

前記入力音声信号の内の雑音成分のスペクトルを推定する手順と、

該分析手順の処理結果を用いて、信号の抑圧ゲイン算出の基礎となる情報であって、少なくとも入力音声信号の内で雑音成分を除く純粋音声成分に対応すべき情報を音声情報として推定する手順と、

該推定された雑音成分スペクトル、音声情報、および前記分析部の出力に対応して抑圧ゲインを算出して前記抑圧手順に与える手順とを計算機に実行させるためのプログラム。

【0115】

(付記17) 入力音声信号を周波数分析して周波数領域の信号に変換する分析ステップと、該周波数領域の信号を抑圧する抑圧ステップと、抑圧された周波数領域の信号を用いて抑圧された時間領域の信号を合成して出力する合成ステップとを実行して雑音を低減する計算機によって使用される記憶媒体において、

10

該分析ステップの処理結果を用いて、信号の抑圧ゲイン算出の基礎となる情報であって、少なくとも入力音声信号の内で雑音成分を除く純粋音声成分に対応すべき情報を音声情報として推定するステップと、

該推定された音声情報と前記分析ステップの処理結果とに対応して抑圧ゲインを算出して前記抑圧ステップに与えるステップとを計算機に実行させるプログラムを格納した計算機読出し可能可搬型記憶媒体。

【0116】

(付記18) 入力音声信号を周波数分析して周波数領域の信号に変換する分析ステップと、該周波数領域の信号を抑圧する抑圧ステップと、抑圧された周波数領域の信号を用いて抑圧された時間領域の信号を合成して出力する合成ステップとを実行して雑音を低減する計算機によって使用される記憶媒体において、

20

前記入力音声信号の内の雑音成分のスペクトルを推定するステップと、

該分析ステップの処理結果を用いて、信号の抑圧ゲイン算出の基礎となる情報であって、少なくとも入力音声信号の内で雑音成分を除く純粋音声成分に対応すべき情報を音声情報として推定するステップと、

該推定された雑音成分スペクトル、音声情報、および前記分析ステップの処理結果に対応して抑圧ゲインを算出して前記抑圧ステップに与えるステップとを計算機に実行させるプログラムを格納した計算機読出し可能可搬型記憶媒体。

【産業上の利用可能性】

30

【0117】

本発明は、携帯電話やIP電話など、音声による通信サービスを提供する通信産業は当然のこととして、雑音重畳音声信号から雑音成分を低減させることが必要となるすべての産業において利用可能である。

【図面の簡単な説明】

【0118】

【図1】本発明の雑音低減装置の原理構成ブロック図である。

【図2】本発明の第1の実施形態における雑音低減装置の構成例を示す図である。

【図3】第1の実施形態における雑音低減処理の全体フローチャートである。

【図4】スペクトル分析処理の詳細フローチャートである。

40

【図5】音声情報推定処理の詳細フローチャートである。

【図6】抑圧ゲイン算出処理の詳細フローチャートである。

【図7】抑圧ゲイン算出関数の例を示す図である。

【図8】図7の抑圧ゲイン算出関数の例を説明するための音声パワー分布の説明図である。

【図9】音声情報推定処理の他の実施例のフローチャートである。

【図10】図9の音声情報推定処理に対応する抑圧ゲイン算出処理のフローチャートである。

【図11】図10における抑圧ゲイン算出処理を説明するための音声パワー分布の説明図である。

50

【図 1 2】第 2 の実施形態における雑音低減装置の構成を示すブロック図である。
【図 1 3】第 2 の実施形態における雑音低減処理の全体フローチャートである。
【図 1 4】第 2 の実施形態における雑音推定処理の詳細フローチャートである。
【図 1 5】第 2 の実施形態における抑圧ゲイン算出処理の詳細フローチャートである。
【図 1 6】図 1 5 の抑圧ゲイン算出処理を説明するためのパワー分布の説明図である。
【図 1 7】抑圧ゲイン算出処理の異なる実施例の詳細フローチャートである。
【図 1 8】図 1 7 の抑圧ゲイン算出処理におけるパワー分布の説明図である。
【図 1 9】本発明を実現するためのプログラムのコンピュータへのローディングを説明する図である。

【図 2 0】雑音低減装置の従来例の構成を示すブロック図である。

10

【符号の説明】

【 0 1 1 9 】

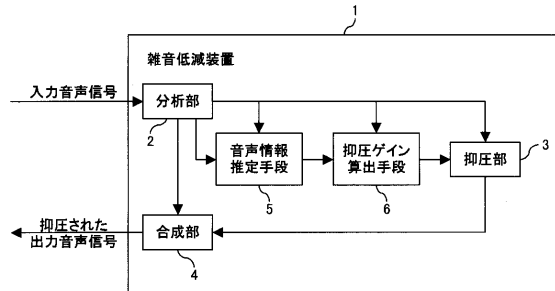
- 1、1 0 雑音低減装置
- 2、1 1 分析部
- 3、1 5 抑圧部
- 4、1 6 合成部
- 5 音声情報推定手段
- 6 抑圧ゲイン算出手段
- 1 2 音声推定部
- 1 3 スペクトルパワー記憶部
- 1 4 抑圧ゲイン算出部
- 1 7 音声符号化部
- 1 8 送信部
- 1 9 雑音推定部
- 2 0 C P U
- 2 1 R O M
- 2 2 R A M
- 2 3 通信インタフェース
- 2 4 記憶装置
- 2 5 入出力装置
- 2 6 読取り装置
- 2 7 バス
- 2 8 プログラム提供者
- 2 9 ネットワーク
- 3 0 可搬型記憶媒体

20

30

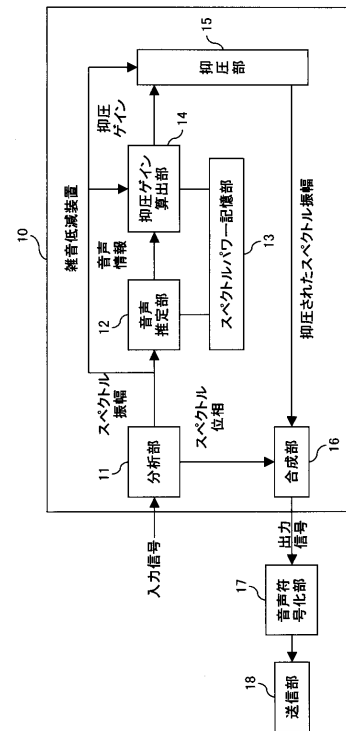
【図 1】

本発明の雑音低減装置の原理構成ブロック図



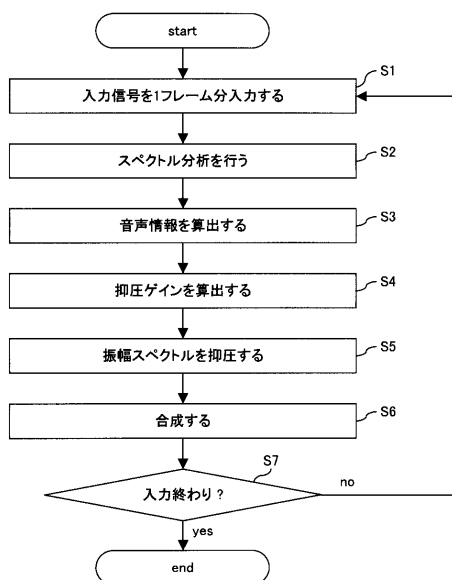
【図 2】

本発明の第1の実施形態における雑音低減装置の構成例を示す図



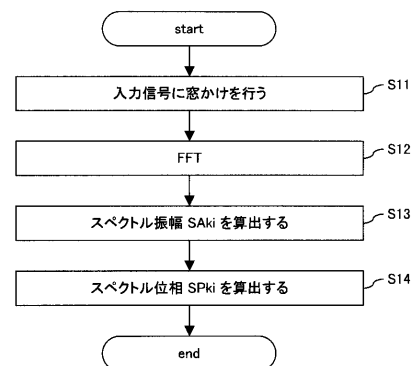
【図 3】

第1の実施形態における雑音低減処理の全体フローチャート



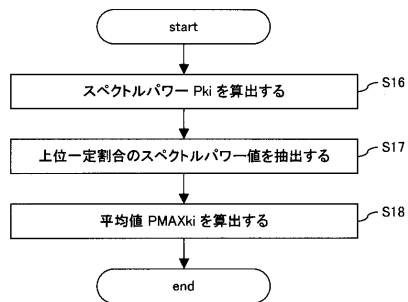
【図 4】

スペクトル分析処理の詳細フローチャート



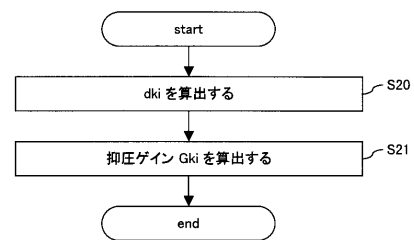
【図 5】

音声情報推定処理の詳細フローチャート



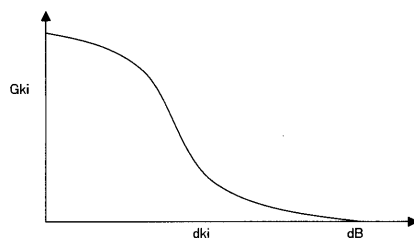
【図 6】

抑圧ゲイン算出処理の詳細フローチャート



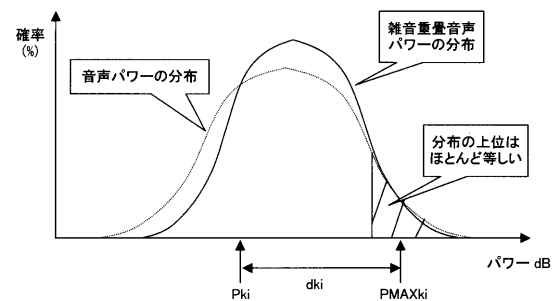
【図 7】

抑圧ゲイン算出関数の例を示す図



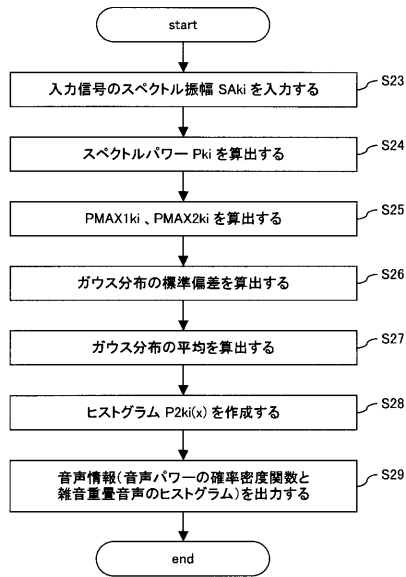
【図 8】

図7の抑圧ゲイン算出関数の例を説明するための音声パワー分布の説明図



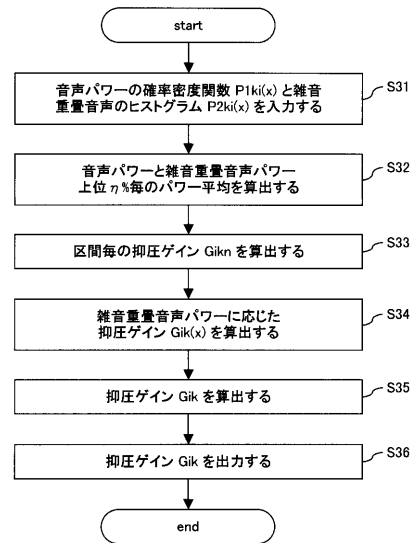
【図 9】

音声情報推定処理の他の
実施例のフローチャート



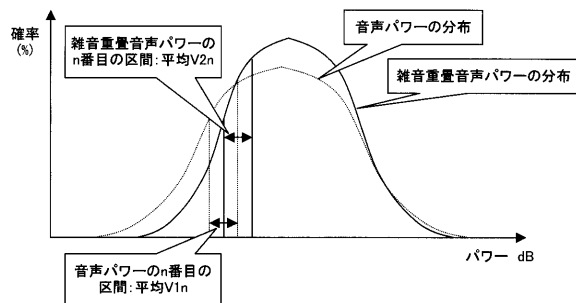
【図 10】

図9の音声情報推定処理に対応する
抑圧ゲイン算出処理のフローチャート



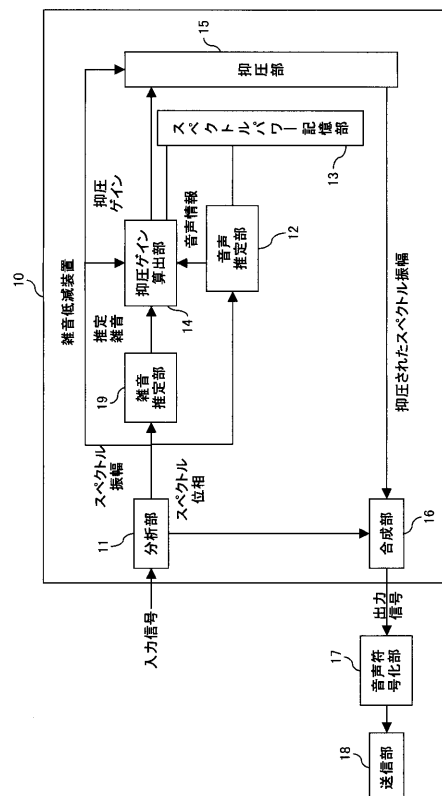
【図 11】

図10における抑圧ゲイン算出処理を
説明するための音声パワー分布の説明図

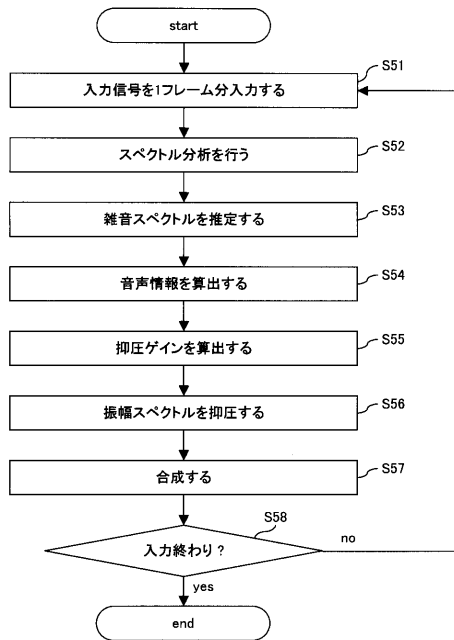


【図 12】

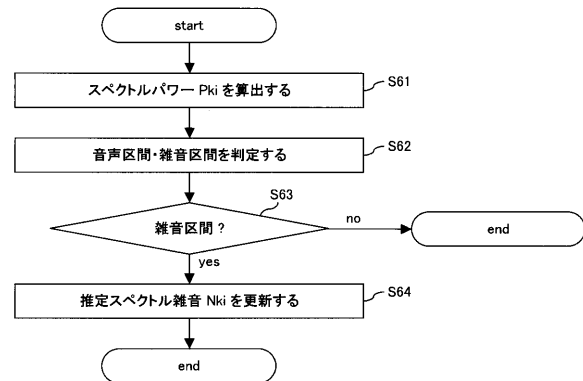
第2の実施形態における
雑音低減装置の構成を示すブロック図



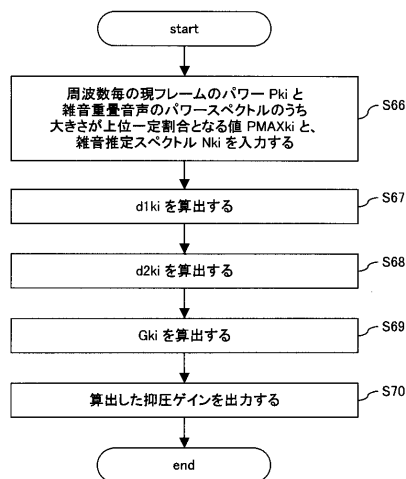
【図 13】

第2の実施形態における
雑音低減処理の全体フローチャート

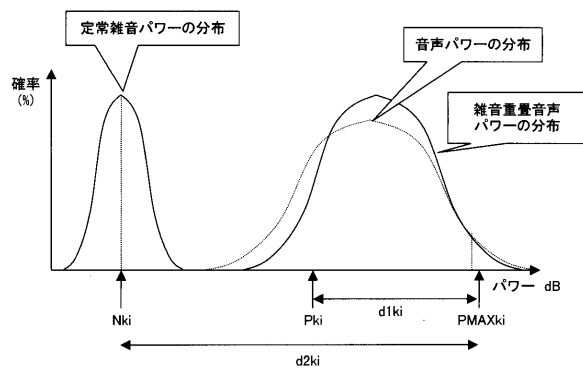
【図 14】

第2の実施形態における
雑音推定処理の詳細フローチャート

【図 15】

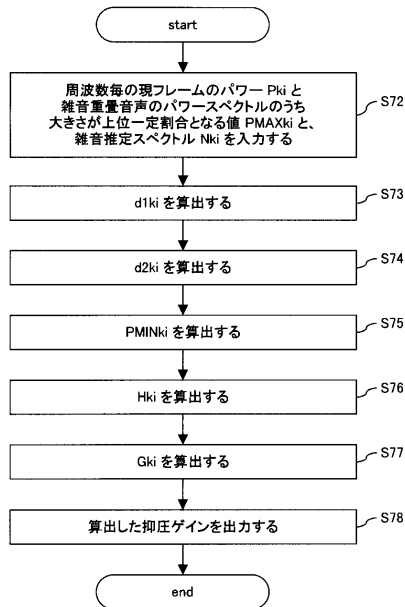
第2の実施形態における
抑圧ゲイン算出処理の詳細フローチャート

【図 16】

図15の抑圧ゲイン算出処理を
説明するためのパワー分布の説明図

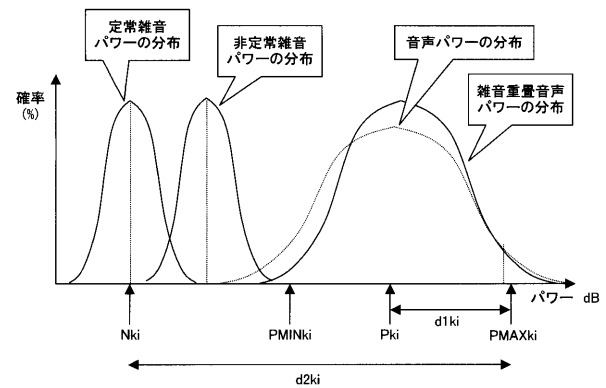
【図 17】

抑圧ゲイン算出処理の異なる
実施例の詳細フローチャート



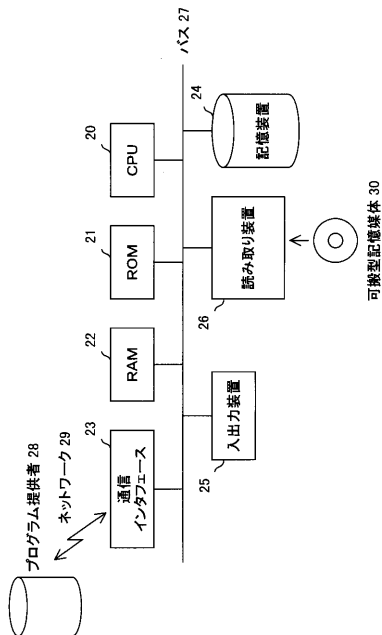
【図 18】

図17の抑圧ゲイン算出処理に
おけるパワー分布の説明図



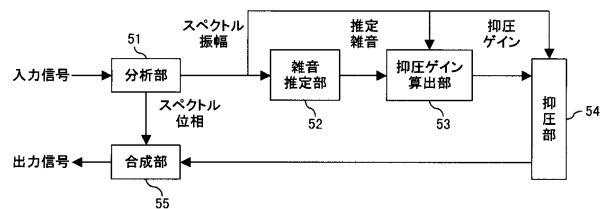
【図 19】

本発明を実現するためのプログラムの
コンピュータへのローディングを説明する図



【図 20】

雑音低減装置の従来例の構成を示すブロック図



フロントページの続き

(72)発明者 松原 光良

神奈川県横浜市港北区新横浜三丁目9番18号 富士通ネットワークテクノロジーズ株式会社内

(72)発明者 大田 恭士

神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

審査官 清水 正一

(56)参考文献 特開平5 - 204395 (JP, A)

特開2000 - 047697 (JP, A)

特開2000 - 330597 (JP, A)

特開平04 - 340599 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G10L 21/02