

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4454591号
(P4454591)

(45) 発行日 平成22年4月21日(2010.4.21)

(24) 登録日 平成22年2月12日(2010.2.12)

(51) Int.Cl.

F I

G 1 O L 21/02 (2006.01)

G 1 O L 21/02 1 O 1 B

請求項の数 14 (全 32 頁)

(21) 出願番号	特願2006-31909 (P2006-31909)	(73) 特許権者	899000068
(22) 出願日	平成18年2月9日(2006.2.9)		学校法人早稲田大学
(65) 公開番号	特開2007-212704 (P2007-212704A)		東京都新宿区戸塚町1丁目104番地
(43) 公開日	平成19年8月23日(2007.8.23)	(73) 特許権者	000004075
審査請求日	平成18年11月24日(2006.11.24)		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	100111763
			弁理士 松本 隆
		(72) 発明者	風間 道子
			東京都新宿区下落合3丁目4番7号
		(72) 発明者	東山 三樹夫
			東京都杉並区成田東2丁目2番12号
		(72) 発明者	柳田 孝司
			静岡県浜松市中沢町10番1号ヤマハ株式会社内

最終頁に続く

(54) 【発明の名称】 雑音スペクトル推定方法、雑音抑圧方法及び雑音抑圧装置

(57) 【特許請求の範囲】

【請求項1】

コンピュータが、

入力音波形のサンプル列を第1のフレーム長で切り出して現フレームの第1のサンプル列を得るステップと、

前記入力音波形のサンプル列を前記第1のフレーム長よりも長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を得た後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を得るステップと、

前記現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを得るステップと、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを得た後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを得るステップと、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

10

20

前記現フレームの第 1 及び第 2 のスペクトルのうちのいずれか一方と前記前フレームの雑音スペクトルとに基づいて現フレームの雑音スペクトルを推定するステップと、

前記平均雑音パワーと前記平均音パワーとを比較し、その比較結果に応じて前記現フレームの雑音スペクトルを推定するステップにおいて前記前フレームの雑音スペクトルとともに用いる現フレームのスペクトルを選択するステップと

を実行することを特徴とする雑音スペクトル推定方法。

【請求項 2】

コンピュータが、

入力音波形のサンプル列を第 1 のフレーム長で切り出して現フレームの第 1 のサンプル列を得るステップと、

前記入力音波形のサンプル列を前記第 1 のフレーム長よりも長い第 2 のフレーム長で前記現フレームより前に切り出して前フレームの第 2 のサンプル列を得た後、前記入力音波形のサンプル列を前記第 2 のフレーム長で前記現フレームに同期して切り出して現フレームの第 2 のサンプル列を得るステップと、

前記現フレームの第 1 のサンプル列を周波数分析して現フレームの第 1 のスペクトルを得るステップと、

前記前フレームの第 2 のサンプル列を周波数分析して前フレームの第 2 のスペクトルを得た後、前記現フレームの第 2 のサンプル列を周波数分析して現フレームの第 2 のスペクトルを得るステップと、

前記前フレームの第 2 のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第 1 及び第 2 のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより大きいことを検知するステップと、

前記検知するステップで前記平均雑音パワーの方が大きいことが検知されないときは前記前フレームの雑音スペクトルと前記現フレームの第 2 のスペクトルとに基づいて現フレームの雑音スペクトルを推定し、前記検知するステップで前記平均雑音パワーの方が大きいことが検知されたときは前記前フレームの雑音スペクトルと前記現フレームの第 1 のスペクトルとに基づいて現フレームの雑音スペクトルを推定するステップと

を実行することを特徴とする雑音スペクトル推定方法。

【請求項 3】

前記現フレームの雑音スペクトルを推定するステップにおいて前記コンピュータは、前記前フレームの雑音スペクトルと前記現フレームの第 1 のスペクトルとを混合して前記現フレームの雑音スペクトルを推定し、その推定の際には前記検知するステップで前記平均雑音パワーの方が大きいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記現フレームの第 1 のスペクトルの混合比率を増大すべく制御することを特徴とする請求項 2 記載の雑音スペクトル推定方法。

【請求項 4】

コンピュータが、

入力音波形のサンプル列を第 1 のフレーム長で切り出して現フレームの第 1 のサンプル列を得るステップと、

前記入力音波形のサンプル列を前記第 1 のフレーム長よりも長い第 2 のフレーム長で前記現フレームより前に切り出して前フレームの第 2 のサンプル列を得た後、前記入力音波形のサンプル列を前記第 2 のフレーム長で前記現フレームに同期して切り出して現フレームの第 2 のサンプル列を得るステップと、

前記現フレームの第 1 のサンプル列を周波数分析して現フレームの第 1 のスペクトルを得るステップと、

前記前フレームの第 2 のサンプル列を周波数分析して前フレームの第 2 のスペクトルを得た後、前記現フレームの第 2 のサンプル列を周波数分析して現フレームの第 2 のスペクトルを得るステップと、

前記前フレームの第 2 のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第 1 及び第 2 のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより大きいことを検知するステップと、

前記前フレームの雑音スペクトルと前記現フレームの第 1 及び第 2 のスペクトルのうちのいずれか一方のスペクトルとを混合して現フレームの雑音スペクトルを推定するステップと、

前記検知するステップで前記平均雑音パワーの方が大きいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記一方のスペクトルの混合比率を増大すべく制御するステップと

を実行することを特徴とする雑音スペクトル推定方法。

【請求項 5】

コンピュータが、

入力音波形のサンプル列を第 1 のフレーム長で切り出して現フレームの第 1 のサンプル列を得るステップと、

前記入力音波形のサンプル列を前記第 1 のフレーム長よりも長い第 2 のフレーム長で前記現フレームより前に切り出して前フレームの第 2 のサンプル列を得た後、前記入力音波形のサンプル列を前記第 2 のフレーム長で前記現フレームに同期して切り出して現フレームの第 2 のサンプル列を得るステップと、

前記現フレームの第 1 のサンプル列を周波数分析して現フレームの第 1 のスペクトルを得るステップと、

前記前フレームの第 2 のサンプル列を周波数分析して前フレームの第 2 のスペクトルを得た後、前記現フレームの第 2 のサンプル列を周波数分析して現フレームの第 2 のスペクトルを得るステップと、

前記前フレームの第 2 のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第 1 及び第 2 のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより小さいことを検知するステップと、

前記前フレームの雑音スペクトルと前記現フレームの第 2 のスペクトルとを混合して現フレームの雑音スペクトルを推定するステップと、

前記検知するステップで前記平均雑音パワーの方が小さいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記現フレームの第 2 のスペクトルの混合比率を増大すべく制御するステップと

を実行することを特徴とする雑音スペクトル推定方法。

【請求項 6】

前記現フレームの雑音スペクトルを推定するステップにおいて前記コンピュータは、前記前フレームの平均雑音パワーと該前フレームより前の複数フレームの平均雑音パワーとを記憶すると共に記憶に係る平均雑音パワーのうちで最小の平均雑音パワーを検知し、検知に係る最小の平均雑音パワーと前記前フレームの平均雑音パワーとを比較して前記前フ

10

20

30

40

50

レームの平均雑音パワーの方が前記検知に係る最小の平均雑音パワーより小さいことを条件として前記前フレームの雑音スペクトルに対する前記現フレームの第2のスペクトルの混合比率を増大すべく制御することを特徴とする請求項5記載の雑音スペクトル推定方法。

【請求項7】

コンピュータが、

入力音波形のサンプル列を第1のフレーム長で切り出して現フレームの第1のサンプル列を得るステップと、

前記入力音波形のサンプル列を前記第1のフレーム長よりも長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を得た後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を得るステップと、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータを付加するステップと、

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを得るステップと、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを得た後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを得るステップと、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記現フレームの第1及び第2のスペクトルのうちのいずれか一方と前記前フレームの雑音スペクトルとに基づいて現フレームの雑音スペクトルを推定するステップと、

前記平均雑音パワーと前記平均音パワーとを比較し、その比較結果に応じて前記現フレームの雑音スペクトルを推定する際に前記前フレームの雑音スペクトルとともに用いる現フレームのスペクトルを選択するステップと、

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを得るステップと

を実行することを特徴とする雑音抑圧方法。

【請求項8】

コンピュータが、

入力音波形のサンプル列を第1のフレーム長で切り出して現フレームの第1のサンプル列を得るステップと、

前記入力音波形のサンプル列を前記第1のフレーム長よりも長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を得た後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を得るステップと、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータを付加するステップと、

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを得るステップと、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを得た後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを得るステップと、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより大きいことを検知するステップと、

前記検知するステップで前記平均雑音パワーの方が大きいことが検知されないときは前記前フレームの雑音スペクトルと前記現フレームの第2のスペクトルとに基づいて現フレームの雑音スペクトルを推定し、前記検知するステップで前記平均雑音パワーの方が大きいことが検知されたときは前記前フレームの雑音スペクトルと前記現フレームの第1のスペクトルとに基づいて現フレームの雑音スペクトルを推定するステップと、

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを得るステップと

を実行することを特徴とする雑音抑圧方法。

【請求項9】

コンピュータが、

入力音波形のサンプル列を第1のフレーム長で切り出して現フレームの第1のサンプル列を得るステップと、

前記入力音波形のサンプル列を前記第1のフレーム長よりも長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を得た後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を得るステップと、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータを付加するステップと、

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを得るステップと、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを得た後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを得るステップと、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより大きいことを検知するステップと、

前記前フレームの雑音スペクトルと前記現フレームの第1及び第2のスペクトルのうちのいずれか一方のスペクトルとを混合して現フレームの雑音スペクトルを推定するステップと、

前記検知するステップで前記平均雑音パワーの方が大きいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記一方のスペクトルの混合比率を増大すべく制御するステップと、

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを得るステップと

を実行することを特徴とする雑音抑圧方法。

【請求項10】

コンピュータが、

入力音波形のサンプル列を第 1 のフレーム長で切り出して現フレームの第 1 のサンプル列を得るステップと、

前記入力音波形のサンプル列を前記第 1 のフレーム長よりも長い第 2 のフレーム長で前記現フレームより前に切り出して前フレームの第 2 のサンプル列を得た後、前記入力音波形のサンプル列を前記第 2 のフレーム長で前記現フレームに同期して切り出して現フレームの第 2 のサンプル列を得るステップと、

前記現フレームの第 1 のサンプル列のフレーム長が前記現フレームの第 2 のサンプル列のフレーム長と等しくなるように前記現フレームの第 1 のサンプル列の末尾にゼロデータを付加するステップと、

10

前記ゼロデータが付加された現フレームの第 1 のサンプル列を周波数分析して現フレームの第 1 のスペクトルを得るステップと、

前記前フレームの第 2 のサンプル列を周波数分析して前フレームの第 2 のスペクトルを得た後、前記現フレームの第 2 のサンプル列を周波数分析して現フレームの第 2 のスペクトルを得るステップと、

前記前フレームの第 2 のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第 1 及び第 2 のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

20

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより小さいことを検知するステップと、

前記前フレームの雑音スペクトルと前記現フレームの第 2 のスペクトルとを混合して現フレームの雑音スペクトルを推定するステップと、

前記検知するステップで前記平均雑音パワーの方が小さいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記現フレームの第 2 のスペクトルの混合比率を増大すべく制御するステップと、

前記現フレームの第 1 のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを得るステップと

30

を実行することを特徴とする雑音抑圧方法。

【請求項 11】

入力音波形のサンプル列を第 1 のフレーム長で切り出して現フレームの第 1 のサンプル列を生成する第 1 の切出手段と、

前記入力音波形のサンプル列を前記第 1 のフレーム長よりも長い第 2 のフレーム長で前記現フレームより前に切り出して前フレームの第 2 のサンプル列を生成した後、前記入力音波形のサンプル列を前記第 2 のフレーム長で前記現フレームに同期して切り出して現フレームの第 2 のサンプル列を生成する第 2 の切出手段と、

前記現フレームの第 1 のサンプル列のフレーム長が前記現フレームの第 2 のサンプル列のフレーム長と等しくなるように前記現フレームの第 1 のサンプル列の末尾にゼロデータを付加する付加手段と、

40

前記ゼロデータが付加された現フレームの第 1 のサンプル列を周波数分析して現フレームの第 1 のスペクトルを生成する第 1 の分析手段と、

前記前フレームの第 2 のサンプル列を周波数分析して前フレームの第 2 のスペクトルを生成した後、前記現フレームの第 2 のサンプル列を周波数分析して現フレームの第 2 のスペクトルを生成する第 2 の分析手段と、

前記前フレームの第 2 のスペクトルに基づいて前フレームの雑音スペクトルを推定する第 1 の推定手段と、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出する第 1 の算出手段と、

50

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出する第2の算出手段と、

前記現フレームの第1及び第2のスペクトルのうちのいずれか一方と前記前フレームの雑音スペクトルとに基づいて現フレームの雑音スペクトルを推定する第2の推定手段と、

前記平均雑音パワーと前記平均音パワーとを比較し、その比較結果に応じて前記第2の推定手段において現フレームの雑音スペクトルを推定する際に前記前フレームの雑音スペクトルとともに用いる現フレームのスペクトルを選択する制御手段と、

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを生成する減算手段と

を備えた雑音抑圧装置。

10

【請求項12】

入力音波形のサンプル列を第1のフレーム長で切り出して現フレームの第1のサンプル列を生成する第1の切出手段と、

前記入力音波形のサンプル列を前記第1のフレーム長よりも長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を生成した後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を生成する第2の切出手段と、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータを付加する付加手段と、

20

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを生成する第1の分析手段と、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを生成した後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを生成する第2の分析手段と、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定する第1の推定手段と、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出する第1の算出手段と、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出する第2の算出手段と、

30

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより大きいことを検知する検知手段と、

前記検知手段で前記平均雑音パワーの方が大きいことが検知されないときは前記前フレームの雑音スペクトルと前記現フレームの第2のスペクトルとに基づいて現フレームの雑音スペクトルを推定し、前記検知手段で前記平均雑音パワーの方が大きいことが検知されたときは前記前フレームの雑音スペクトルと前記現フレームの第1のスペクトルとに基づいて現フレームの雑音スペクトルを推定する第2の推定手段と、

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを生成する減算手段と

40

を備えた雑音抑圧装置。

【請求項13】

入力音波形のサンプル列を第1のフレーム長で切り出して現フレームの第1のサンプル列を生成する第1の切出手段と、

前記入力音波形のサンプル列を前記第1のフレーム長よりも長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を生成した後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を生成する第2の切出手段と、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータ

50

を付加する付加手段と、

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを生成する第1の分析手段と、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを生成した後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを生成する第2の分析手段と、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定する第1の推定手段と、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出する第1の算出手段と、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出する第2の算出手段と、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより大きいことを検知する検知手段と、

前記前フレームの雑音スペクトルと前記現フレームの第1及び第2のスペクトルのうちのいずれか一方のスペクトルとを混合して現フレームの雑音スペクトルを推定する第2の推定手段と、

前記検知手段で前記平均雑音パワーの方が大きいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記一方のスペクトルの混合比率を増大すべく制御する制御手段と、

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを生成する減算手段と

を備えた雑音抑圧装置。

【請求項14】

入力音波形のサンプル列を第1のフレーム長で切り出して現フレームの第1のサンプル列を生成する第1の切出手段と、

前記入力音波形のサンプル列を前記第1のフレーム長より長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を生成した後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を生成する第2の切出手段と、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータを付加する付加手段と、

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを生成する第1の分析手段と、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを生成した後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを生成する第2の分析手段と、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定する第1の推定手段と、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出する第1の算出手段と、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出する第2の算出手段と、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより小さいことを検知する検知手段と、

前記前フレームの雑音スペクトルと前記現フレームの第2のスペクトルとを混合して現フレームの雑音スペクトルを推定する第2の推定手段と、

前記検知手段で前記平均雑音パワーの方が小さいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記現フレームの第2のスペクトルの混合比率を増大

10

20

30

40

50

すべく制御する制御手段と、

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを生成する減算手段と
を備えた雑音抑圧装置。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、雑音抑圧等に用いるに好適な雑音スペクトル推定方法と、この雑音スペクトル推定方法及びスペクトルサブトラクション法を用いて雑音を抑圧する方法及び装置と

10

【背景技術】

【0002】

従来、音声通信技術や音声認識技術において用いられる雑音抑圧方法としては、スペクトルサブトラクション法が知られている（例えば、特許文献1～3参照）。スペクトルサブトラクション法は、雑音が混入した音声信号から雑音のスペクトルを推定し、この推定に係る雑音のスペクトルを音声信号のスペクトルから差し引くことにより雑音が抑圧された音声信号を得るものである。

【特許文献1】特開平11-3094号公報

【特許文献2】特開2002-14694号公報

20

【特許文献3】特開2003-223186号公報

【発明の開示】

【発明が解決しようとする課題】

【0003】

従来のスペクトルサブトラクション法によると、有音声区間（音声と雑音が混在する区間）と無音声区間（音声がなく、雑音のみ存在する区間）とで同様に雑音推定が行なわれるため、主として無音声区間において新たな雑音を雑音スペクトルの推定に十分反映させることができず、十分な雑音抑圧が困難であった。このような問題点を解決するため、現フレームの音入力（音入力）のスペクトル包絡と前フレームの推定雑音のスペクトル包絡との相関値を求め、相関値が大きい無音声区間では新たな雑音を雑音スペクトルの推定に十分に反映させるようにした雑音スペクトル推定方法が本願と同一出願人により先に提案された（特願2004-100935号）。

30

【0004】

また、従来のスペクトルサブトラクション法によると、雑音スペクトルの推定に用いる音入力（音入力）のフレーム長（サンプル数）と、推定に係る雑音スペクトルを減算すべき被減算側の音入力（音入力）のフレーム長とが例えば512サンプルで等しかったため、雑音スペクトルの推定では周波数分解能が不足し、十分な雑音抑圧が困難であった。雑音スペクトルの推定に用いる音入力（音入力）のフレーム長を十分な周波数分解能となるように例えば4096サンプルとすることも考えられるが、このようにすると、被減算側の音入力（音入力）のフレーム長も4096サンプルとなるため、被減算側の振幅スペクトルの時間分解能が不足する事態を招く。このような問題点を解決するため、雑音スペクトルの推定に用いる音入力（音入力）のフレーム長を被減算側の音入力（音入力）のフレーム長（例えば512サンプル）より長く設定（例えば4096サンプルに設定）することで高精度の雑音スペクトルの推定を可能にする方法が本願と同一出願人により先に提案された（特願2005-144744号）。

40

【0005】

図8は、本願の発明者の研究に係る雑音抑圧処理を示すもので、この処理では、上記した2つの先行出願技術が採用されている。

【0006】

図8において、音入力 $x(n)$ （ $n = 0, 1, 2, \dots$ ）は、入力音波形をA/D（アナログ/デジタル）変換した波形データのサンプル列からなるものである。入力音波形とし

50

ては、雑音付き音声信号波形のことが多いが、音声信号波形のみであったり、雑音波形のみであったりすることもある。ステップS1では、音入力 $x(n)$ を雑音抑圧用フレーム（雑音抑圧の対象となるフレーム）として比較的短いフレーム長（時間窓長）で切り出す。また、ステップS2では、音入力 $x(n)$ を雑音推定用フレーム（雑音スペクトルの推定に用いられるフレーム）として比較的長いフレーム長で切り出す。ステップS1, S2でのフレーム切り出しは、フレームの頭を時間的に揃えて行なわれる。

【0007】

一例として、サンプリング周波数が16kHzである場合、ステップS1での切り出しは、512サンプルの長さで行なわれ、ステップS2での切り出しは、4096サンプルの長さで行なわれる。また、S1, S2のいずれのステップでも、フレームの切り出しは、雑音抑圧用フレームの半分の周期で行なわれる。

【0008】

ステップS3では、後述するステップS11での減算処理の際にサンプル数を揃える必要があるため、ステップS1で切り出された短いフレームにゼロデータを付加する。上記した512サンプルの例では、 $4096 - 512 = 3584$ サンプル分のゼロデータを短いフレーム毎に付加する。

【0009】

ステップS4では、ゼロデータを付加した各フレームデータに高速フーリエ変換（FFT: Fast Fourier Transform）処理を施して周波数分析を行なう。そして、周波数分析データに基づいて各フレーム毎に振幅スペクトル $|X_1(k)|$ （ $k = 0, 1, 2 \dots N-1$: N はサンプル数）及び位相スペクトルを算出する。ステップS5では、各長いフレームデータにFFT処理を施して周波数分析を行ない、周波数分析データに基づいて振幅スペクトル $|X_2(k)|$ （ $k = 0, 1, 2 \dots N-1$ ）を算出する。

【0010】

ステップS6では、振幅スペクトル $|X_2(k)|$ に平滑化処理を施す。そして、ステップS7では、平滑化処理を受けた振幅スペクトル $|X_2'(k)|$ からスペクトル包絡 $|X_2''(k)|$ を抽出すると共に、後述のステップS10で得られる推定雑音振幅スペクトル $|No(k)|$ からスペクトル包絡 $|No'(k)|$ を抽出する。スペクトル $|X_2(k)|$ を今回算出されたものとする、スペクトル $|No(k)|$ は、前回算出されたものである。

【0011】

ステップS8では、スペクトル包絡 $|X_2''(k)|$ とスペクトル包絡 $|No'(k)|$ との相関を求め、相関値を算出する。相関値の計算式については後述する。相関値は、0～1の範囲内の値を取り、雑音のある有音声区間では小さく、雑音のみある無音声区間では大きくなる。

【0012】

ステップS9では、今回算出された振幅スペクトル $|X_2(k)|$ と、前回算出された雑音振幅スペクトル $|No(k)|$ とを相関値に応じた比率で加算（混合）することにより今回用いるべき雑音振幅スペクトル $|N(k)|$ を所定の推定式に従って算出する。スペクトル $|N(k)|$ の推定式については後述する。相関値が小さい有音声区間では、相関値が大きいときに比べて、前回の雑音振幅スペクトル $|No(k)|$ の比率を高くすると共に今回の振幅スペクトル $|X_2(k)|$ の比率を低くすることで推定に係る雑音振幅スペクトル $|N(k)|$ の更新を抑制する。これに対し、相関値が大きい無音声区間では、相関値が小さいときに比べて、前回の雑音振幅スペクトル $|No(k)|$ の比率を低くすると共に今回の振幅スペクトル $|X_2(k)|$ の比率を高くすることで推定に係る雑音振幅スペクトル $|N(k)|$ の更新を促進する。

【0013】

ステップS10では、推定に係る雑音振幅スペクトル $|N(k)|$ を半フレーム相当の時間だけ遅延させて雑音振幅スペクトル $|No(k)|$ として前述のステップS7, S9に供給する。なお、システム起動時において、ステップS10では所定の雑音振幅スペク

10

20

30

40

50

トル初期値をステップ S 7, S 9 に供給する。

【 0 0 1 4 】

ステップ S 1 1 では、ステップ S 4 で得られた今回の振幅スペクトル $|X_1(k)|$ からステップ S 9 で得られた今回の雑音振幅スペクトル $|N(k)|$ を減算することにより雑音抑圧を行なう。この後、ステップ S 1 1 では、雑音抑圧に係る振幅スペクトルとステップ S 4 で得られた対応する位相スペクトルとを合成して各フレーム毎に音声スペクトル $G(k)$ を作成する。

【 0 0 1 5 】

ステップ S 1 2 では、各フレームの音声スペクトル $G(k)$ に逆高速フーリエ変換 (IFFT) 処理を施して時間領域の音声信号 (音声波形データのサンプル列) を得る。そして、ステップ S 1 3 では、各フレームの音声信号からステップ S 3 での付加に係る 3854 サンプル分のゼロデータを削除する。この後、ステップ S 1 3 では、各フレームの音声信号に三角窓関数を乗算すると共に順次の 2 フレームの音声信号を半フレームずつオーバーラップさせるように加算することによりフレーム合成を行なう。この結果、雑音が抑圧された音声出力 $g(n)$ が得られる。

【 0 0 1 6 】

上記した雑音抑圧処理によれば、ステップ S 9 において無音声区間での推定雑音スペクトルの更新を促進するようにしたので、無音声区間で十分な雑音抑圧が可能となる。また、ステップ S 2 において音入力をステップ S 1 に比べて長いフレームで切り出すようにしたので、ステップ S 9 では長いフレームの振幅スペクトル $|X_2(k)|$, $|N_o(k)|$ に基づいて高精度の雑音スペクトルの推定を行なうことができ、十分な雑音抑圧が可能となる。しかしながら、上記した雑音抑圧処理では、例えば空調雑音のオン/オフ等の雑音変動に対応することを考慮した場合、ステップ S 9 において長いフレーム長の振幅スペクトルを用いているため、雑音のオン又はオフに対する推定雑音振幅スペクトル $|N(k)|$ の更新が緩慢であり、雑音変動に対する追従性が十分でない (応答速度が遅い) という問題点がある。

【 0 0 1 7 】

この発明の目的は、雑音変動に対する追従性が良好な新規な雑音スペクトル推定方法と、この雑音スペクトル推定方法及びスペクトルサブトラクション法を用いて雑音を抑圧するための新規な方法及び装置とを提供することにある。

【課題を解決するための手段】

【 0 0 1 8 】

この発明に係る雑音スペクトル推定方法は、

入力音波形のサンプル列を比較的短い第 1 のフレーム長で切り出して現フレームの第 1 のサンプル列を得るステップと、

前記入力音波形のサンプル列を比較的長い第 2 のフレーム長で前記現フレームより前に切り出して前フレームの第 2 のサンプル列を得た後、前記入力音波形のサンプル列を前記第 2 のフレーム長で前記現フレームに同期して切り出して現フレームの第 2 のサンプル列を得るステップと、

前記現フレームの第 1 のサンプル列を周波数分析して現フレームの第 1 のスペクトルを得るステップと、

前記前フレームの第 2 のサンプル列を周波数分析して前フレームの第 2 のスペクトルを得た後、前記現フレームの第 2 のサンプル列を周波数分析して現フレームの第 2 のスペクトルを得るステップと、

前記前フレームの第 2 のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第 1 及び第 2 のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

10

20

30

40

50

前記前フレームの雑音スペクトルに基づいて現フレームの雑音スペクトルを推定するステップと、

前記平均雑音パワーと前記平均音パワーとを比較し、その比較結果に応じて前記現フレームの雑音スペクトルの推定態様を制御するステップとを含むものである。この雑音スペクトル推定方法は、以下に述べる第１～第３の雑音スペクトル推定方法として具体化することができる。

【００１９】

この発明に係る第１の雑音スペクトル推定方法は、

入力音波形のサンプル列を比較的短い第１のフレーム長で切り出して現フレームの第１のサンプル列を得るステップと、

前記入力音波形のサンプル列を比較的長い第２のフレーム長で前記現フレームより前に切り出して前フレームの第２のサンプル列を得た後、前記入力音波形のサンプル列を前記第２のフレーム長で前記現フレームに同期して切り出して現フレームの第２のサンプル列を得るステップと、

前記現フレームの第１のサンプル列を周波数分析して現フレームの第１のスペクトルを得るステップと、

前記前フレームの第２のサンプル列を周波数分析して前フレームの第２のスペクトルを得た後、前記現フレームの第２のサンプル列を周波数分析して現フレームの第２のスペクトルを得るステップと、

前記前フレームの第２のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第１及び第２のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより大きいことを検知するステップと、

前記検知するステップで前記平均雑音パワーの方が大きいことが検知されないときは前記前フレームの雑音スペクトルと前記現フレームの第２のスペクトルとに基づいて現フレームの雑音スペクトルを推定し、前記検知するステップで前記平均雑音パワーの方が大きいことが検知されたときは前記前フレームの雑音スペクトルと前記現フレームの第１のスペクトルとに基づいて現フレームの雑音スペクトルを推定するステップとを含むものである。

【００２０】

第１の雑音スペクトル推定方法によれば、前フレーム内の平均雑音パワーの方が大きいことが検知されないときは前フレームの雑音スペクトルと現フレームの第２のスペクトルとに基づいて現フレームの雑音スペクトルを推定すると共に、前フレーム内の平均雑音パワーの方が大きいことが検知されたときは前フレームの雑音スペクトルと現フレームの第１のスペクトルとに基づいて現フレームの雑音スペクトルを推定するようにしたので、前フレーム内の平均雑音パワーが現フレーム内の平均音パワー以下である時は、フレーム長が長い現フレームの第２のスペクトルを用いて精度良く雑音スペクトルの推定を行なえると共に、前フレーム内の平均雑音パワーの方が大きくなる雑音減少時には、フレーム長が短い現フレームの第１のスペクトルを用いて追従性良く雑音スペクトルの推定を行なえる。

【００２１】

この発明に係る第２の雑音スペクトル推定方法は、

入力音波形のサンプル列を比較的短い第１のフレーム長で切り出して現フレームの第１のサンプル列を得るステップと、

前記入力音波形のサンプル列を比較的長い第２のフレーム長で前記現フレームより前に切り出して前フレームの第２のサンプル列を得た後、前記入力音波形のサンプル列を前記

第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を得るステップと、

前記現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを得るステップと、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを得た後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを得るステップと、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより大きいことを検知するステップと、

前記前フレームの雑音スペクトルと前記現フレームの第1及び第2のスペクトルのうちのいずれか一方のスペクトルとを混合して現フレームの雑音スペクトルを推定するステップと、

前記検知するステップで前記平均雑音パワーの方が大きいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記一方のスペクトルの混合比率を増大すべく制御するステップとを含むものである。

【0022】

第2の雑音スペクトル推定方法によれば、前フレームの雑音スペクトルと現フレームの第1及び第2のスペクトルのうちのいずれか一方のスペクトルとを混合して現フレームの雑音スペクトルを推定する際に、前フレーム内の平均雑音パワーの方が大きいことが検知されるのに応答して前フレームの雑音スペクトルに対する前記一方のスペクトルの混合比率を増大すべく制御するようにしたので、前フレーム内の平均雑音パワーの方が大きくなる雑音減少時には、前記一方のスペクトルの混合比率の増大により現フレームの雑音スペクトルの更新が促進され、追従性良好な雑音スペクトルの推定が可能となる。

【0023】

第2の雑音スペクトル推定方法において、前記一方のスペクトルとして前記第1のスペクトルを用いる場合は、前述した第1の雑音スペクトル推定方法と組合せるのに好適であり、このようにすると、雑音減少時に雑音スペクトルを推定する際に追従性を一層向上させることができる。また、第2の雑音スペクトル推定方法において、前記一方のスペクトルとして前記第2のスペクトルを用いる場合は、前記第2のスペクトルのフレーム長が長いので、精度良く雑音スペクトルの推定を行なうことができる。

【0024】

この発明に係る第3の雑音スペクトル推定方法は、

入力音波形のサンプル列を比較的短い第1のフレーム長で切り出して現フレームの第1のサンプル列を得るステップと、

前記入力音波形のサンプル列を比較的長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を得た後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を得るステップと、

前記現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを得るステップと、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを得た後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを得るステップと、

10

20

30

40

50

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより小さいことを検知するステップと、

前記前フレームの雑音スペクトルと前記現フレームの第2のスペクトルとを混合して現フレームの雑音スペクトルを推定するステップと、

前記検知するステップで前記平均雑音パワーの方が小さいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記現フレームの第2のスペクトルの混合比率を増大すべく制御するステップとを含むものである。

【0025】

第3の雑音スペクトル推定方法によれば、前フレームの雑音スペクトルと現フレームの第2のスペクトルとを混合して現フレームの雑音スペクトルを推定する際に、前フレーム内の平均雑音パワーの方が小さいことが検知されるのに応答して前フレームの雑音スペクトルに対する現フレームの第2のスペクトルの混合比率を増大すべく制御するようにしたので、前フレーム内の平均雑音パワーの方が小さくなる雑音増大時には、現フレームの第2のスペクトルの混合比率の増大により現フレームの雑音スペクトルの更新が促進され、追従性良好な雑音スペクトルの推定が可能となる。また、第3の雑音スペクトル推定方法では、前フレームの雑音スペクトルと現フレームの第2のスペクトルとを混合して現フレームの雑音スペクトルを推定しているので、フレーム長が長いスペクトルを用いて精度良く雑音スペクトルを推定することができる。

【0026】

第3の雑音スペクトル推定方法において、前記現フレームの雑音スペクトルを推定するステップでは、前記前フレームの平均雑音パワーと該前フレームより前の複数フレームの平均雑音パワーとを記憶すると共に記憶に係る平均雑音パワーのうちで最小の平均雑音パワーを検知し、検知に係る最小の平均雑音パワーと前記前フレームの平均雑音パワーとを比較して前記前フレームの平均雑音パワーの方が前記検知に係る最小の平均雑音パワーより小さいことを条件として前記前フレームの雑音スペクトルに対する前記現フレームの第2のスペクトルの混合比率を増大すべく制御するようにしてもよい。このようにすると、雑音増大を一層確実に検知可能となる。

【0027】

この発明に係る雑音抑圧方法は、

入力音波形のサンプル列を比較的短い第1のフレーム長で切り出して現フレームの第1のサンプル列を得るステップと、

前記入力音波形のサンプル列を比較的長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を得た後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を得るステップと、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータを付加するステップと、

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを得るステップと、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを得た後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを得るステップと、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記前フレームの雑音スペクトルに基づいて現フレームの雑音スペクトルを推定するステップと、

前記平均雑音パワーと前記平均音パワーとを比較し、その比較結果に応じて前記現フレームの雑音スペクトルの推定態様を制御するステップと、

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを得るステップと

を含むものである。この雑音抑圧方法は、以下に述べる第1～第3の雑音抑圧方法として具体化することができる。

【0028】

この発明に係る第1の雑音抑圧方法は、

入力音波形のサンプル列を比較的短い第1のフレーム長で切り出して現フレームの第1のサンプル列を得るステップと、

前記入力音波形のサンプル列を比較的長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を得た後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を得るステップと、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータを付加するステップと、

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを得るステップと、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを得た後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを得るステップと、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより大きいことを検知するステップと、

前記検知するステップで前記平均雑音パワーの方が大きいことが検知されないときは前記前フレームの雑音スペクトルと前記現フレームの第2のスペクトルとに基づいて現フレームの雑音スペクトルを推定し、前記検知するステップで前記平均雑音パワーの方が大きいことが検知されたときは前記前フレームの雑音スペクトルと前記現フレームの第1のスペクトルとに基づいて現フレームの雑音スペクトルを推定するステップと、

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを得るステップとを含むものである。

【0029】

第1の雑音抑圧方法によれば、前述した第1の雑音スペクトル推定方法及びスペクトルサブトラクション法を用いて雑音抑圧を行なうので、雑音抑圧性能を向上させることができ、特に雑音減少時の雑音抑圧性能を向上させることができる。第1の雑音抑圧方法では

10

20

30

40

50

、現フレームの第1のサンプル列の末尾にゼロデータを付加して現フレームの第1のサンプル列と現フレームの第2のサンプル列とでフレーム長を等しくしている。これは、現フレームの第1のスペクトルから現フレームの雑音スペクトルを減算する際にデータ長（サンプル数）を揃える必要があるためである。

【0030】

この発明に係る第2の雑音抑圧方法は、

入力音波形のサンプル列を比較的短い第1のフレーム長で切り出して現フレームの第1のサンプル列を得るステップと、

前記入力音波形のサンプル列を比較的長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を得た後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を得るステップと、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータを付加するステップと、

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを得るステップと、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを得た後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを得るステップと、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより大きいことを検知するステップと、

前記前フレームの雑音スペクトルと前記現フレームの第1及び第2のスペクトルのうちのいずれか一方のスペクトルとを混合して現フレームの雑音スペクトルを推定するステップと、

前記検知するステップで前記平均雑音パワーの方が大きいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記一方のスペクトルの混合比率を増大すべく制御するステップと、

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを得るステップとを含むものである。

【0031】

第2の雑音抑圧方法によれば、前述した第2の雑音スペクトル推定方法及びスペクトルサブトラクション法を用いて雑音抑圧を行なうので、雑音減少時の雑音抑圧性能を向上させることができる。

【0032】

この発明に係る第3の雑音抑圧方法は、

入力音波形のサンプル列を比較的短い第1のフレーム長で切り出して現フレームの第1のサンプル列を得るステップと、

前記入力音波形のサンプル列を比較的長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を得た後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を得るステップと、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列

10

20

30

40

50

のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータを付加するステップと、

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを得るステップと、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを得た後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを得るステップと、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定するステップと、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出するステップと、

10

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出するステップと、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより小さいことを検知するステップと、

前記前フレームの雑音スペクトルと前記現フレームの第2のスペクトルとを混合して現フレームの雑音スペクトルを推定するステップと、

前記検知するステップで前記平均雑音パワーの方が小さいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記現フレームの第2のスペクトルの混合比率を増大すべく制御するステップと、

20

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを得るステップとを含むものである。

【0033】

第3の雑音抑圧方法によれば、前述した第3の雑音スペクトル推定方法及びスペクトルサブトラクション法を用いて雑音抑圧を行なうので、雑音抑圧性能を向上させることができ、特に雑音増大時の雑音抑圧性能を向上させることができる。

【0034】

この発明に係る雑音抑圧装置は、

入力音波形のサンプル列を比較的短い第1のフレーム長で切り出して現フレームの第1のサンプル列を生成する第1の切出手段と、

30

前記入力音波形のサンプル列を比較的長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を生成した後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を生成する第2の切出手段と、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータを付加する付加手段と、

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを生成する第1の分析手段と、

40

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを生成した後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを生成する第2の分析手段と、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定する第1の推定手段と、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出する第1の算出手段と、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出する第2の算出手段と、

前記前フレームの雑音スペクトルに基づいて現フレームの雑音スペクトルを推定する第

50

2 の推定手段と、

前記平均雑音パワーと前記平均音パワーとを比較し、その比較結果に応じて前記現フレームの雑音スペクトルの推定態様を制御する制御手段と、

前記現フレームの第 1 のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを生成する減算手段と

を備えたものである。この雑音抑圧装置は、以下に述べる第 1 ～ 第 3 の雑音抑圧装置として具体化することができる。

【 0 0 3 5 】

この発明に係る第 1 の雑音抑圧装置は、

入力音波形のサンプル列を比較的短い第 1 のフレーム長で切り出して現フレームの第 1 のサンプル列を生成する第 1 の切出手段と、

前記入力音波形のサンプル列を比較的長い第 2 のフレーム長で前記現フレームより前に切り出して前フレームの第 2 のサンプル列を生成した後、前記入力音波形のサンプル列を前記第 2 のフレーム長で前記現フレームに同期して切り出して現フレームの第 2 のサンプル列を生成する第 2 の切出手段と、

前記現フレームの第 1 のサンプル列のフレーム長が前記現フレームの第 2 のサンプル列のフレーム長と等しくなるように前記現フレームの第 1 のサンプル列の末尾にゼロデータを付加する付加手段と、

前記ゼロデータが付加された現フレームの第 1 のサンプル列を周波数分析して現フレームの第 1 のスペクトルを生成する第 1 の分析手段と、

前記前フレームの第 2 のサンプル列を周波数分析して前フレームの第 2 のスペクトルを生成した後、前記現フレームの第 2 のサンプル列を周波数分析して現フレームの第 2 のスペクトルを生成する第 2 の分析手段と、

前記前フレームの第 2 のスペクトルに基づいて前フレームの雑音スペクトルを推定する第 1 の推定手段と、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出する第 1 の算出手段と、

前記現フレームの第 1 及び第 2 のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出する第 2 の算出手段と、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより大きいことを検知する検知手段と、

前記検知手段で前記平均雑音パワーの方が大きいことが検知されないときは前記前フレームの雑音スペクトルと前記現フレームの第 2 のスペクトルとに基づいて現フレームの雑音スペクトルを推定し、前記検知手段で前記平均雑音パワーの方が大きいことが検知されたときは前記前フレームの雑音スペクトルと前記現フレームの第 1 のスペクトルとに基づいて現フレームの雑音スペクトルを推定する第 2 の推定手段と、

前記現フレームの第 1 のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを生成する減算手段と

を備えたものである。

【 0 0 3 6 】

第 1 の雑音抑圧装置は、前述した第 1 の雑音抑圧方法を実行するもので、第 1 の雑音抑圧方法と同様の作用効果を奏する。

【 0 0 3 7 】

この発明に係る第 2 の雑音抑圧装置は、

入力音波形のサンプル列を比較的短い第 1 のフレーム長で切り出して現フレームの第 1 のサンプル列を生成する第 1 の切出手段と、

前記入力音波形のサンプル列を比較的長い第 2 のフレーム長で前記現フレームより前に切り出して前フレームの第 2 のサンプル列を生成した後、前記入力音波形のサンプル列を前記第 2 のフレーム長で前記現フレームに同期して切り出して現フレームの第 2 のサンプル列を生成する第 2 の切出手段と、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータを付加する付加手段と、

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを生成する第1の分析手段と、

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを生成した後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを生成する第2の分析手段と、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定する第1の推定手段と、

10

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出する第1の算出手段と、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出する第2の算出手段と、

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより大きいことを検知する検知手段と、

前記前フレームの雑音スペクトルと前記現フレームの第1及び第2のスペクトルのうちのいずれか一方のスペクトルとを混合して現フレームの雑音スペクトルを推定する第2の推定手段と、

前記検知手段で前記平均雑音パワーの方が大きいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記一方のスペクトルの混合比率を増大すべく制御する制御手段と、

20

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを生成する減算手段とを備えたものである。

【0038】

第2の雑音抑圧装置は、前述した第2の雑音抑圧方法を実行するもので、第2の雑音抑圧方法と同様の作用効果を奏する。

【0039】

この発明に係る第3の雑音抑圧装置は、
入力音波形のサンプル列を比較的短い第1のフレーム長で切り出して現フレームの第1のサンプル列を生成する第1の切出手段と、

30

前記入力音波形のサンプル列を比較的長い第2のフレーム長で前記現フレームより前に切り出して前フレームの第2のサンプル列を生成した後、前記入力音波形のサンプル列を前記第2のフレーム長で前記現フレームに同期して切り出して現フレームの第2のサンプル列を生成する第2の切出手段と、

前記現フレームの第1のサンプル列のフレーム長が前記現フレームの第2のサンプル列のフレーム長と等しくなるように前記現フレームの第1のサンプル列の末尾にゼロデータを付加する付加手段と、

前記ゼロデータが付加された現フレームの第1のサンプル列を周波数分析して現フレームの第1のスペクトルを生成する第1の分析手段と、

40

前記前フレームの第2のサンプル列を周波数分析して前フレームの第2のスペクトルを生成した後、前記現フレームの第2のサンプル列を周波数分析して現フレームの第2のスペクトルを生成する第2の分析手段と、

前記前フレームの第2のスペクトルに基づいて前フレームの雑音スペクトルを推定する第1の推定手段と、

前記前フレームの雑音スペクトルに基づいて前フレーム内の平均雑音パワーを算出する第1の算出手段と、

前記現フレームの第1及び第2のサンプル列のうちのいずれか一方のサンプル列に基づいて現フレーム内の平均音パワーを算出する第2の算出手段と、

50

前記平均雑音パワーと前記平均音パワーとを比較して前記平均雑音パワーの方が前記平均音パワーより小さいことを検知する検知手段と、

前記前フレームの雑音スペクトルと前記現フレームの第2のスペクトルとを混合して現フレームの雑音スペクトルを推定する第2の推定手段と、

前記検知手段で前記平均雑音パワーの方が小さいことが検知されるのに応答して前記前フレームの雑音スペクトルに対する前記現フレームの第2のスペクトルの混合比率を増大すべく制御する制御手段と、

前記現フレームの第1のスペクトルから前記現フレームの雑音スペクトルを減算して雑音が抑圧された音スペクトルを生成する減算手段と
を備えたものである。

10

【0040】

第3の雑音抑圧装置は、前述した第3の雑音抑圧方法を実行するもので、第3の雑音抑圧方法と同様の作用効果を奏する。

【発明の効果】

【0041】

この発明によれば、前フレーム内の平均雑音パワーと現フレーム内の平均音パワーとを比較し、平均雑音パワーの方が大きい雑音減少時には現フレームの音入力振幅スペクトルとしてフレーム長が短い振幅スペクトルを用いたり、前フレームの雑音スペクトルに対する現フレームの音入力振幅スペクトルの混合比率を増大させたりして雑音スペクトルの更新を促進し、平均雑音パワーの方が小さい雑音増大時には前フレームの雑音スペクトルに対する現フレームの音入力振幅スペクトルの混合比率を増大させて雑音スペクトルの更新を促進するようにしたので、雑音変動に対する追従性が良好な雑音スペクトルの推定が可能になる効果が得られる。

20

【0042】

また、この発明の雑音スペクトル推定方法及びスペクトルサブトラクション法を用いて雑音抑圧を行なうようにしたので、雑音減少時や雑音増大時の雑音抑圧性能が向上する効果も得られる。

【発明を実施するための最良の形態】

【0043】

図1は、この発明の一実施形態に係る雑音抑圧装置の構成を示すもので、この雑音抑圧装置は、雑音抑圧部12及び雑音スペクトル出力部14を備えている。

30

【0044】

図1において、音入力 $x(n)$ は、図8に関して前述したと同様の入力音波形のサンプル列からなるもので、雑音抑圧部12のフレーム切出部32と、雑音スペクトル出力部14のフレーム切出部16とに共通に供給される。フレーム切出部32は、図2に示すように音入力 $x(n)$ のサンプル列を比較的短いフレーム長 T_1 (M サンプル)で順次に切り出すものである。切出部32の切出動作は、あるフレームのサンプル列 D_{11b} の切り出しから次のフレームのサンプル列 D_{21b} の切り出しが半フレーム $T_1/2$ ($M/2$ サンプル)相当の時間だけ遅れるようにして行なわれる。このことは、サンプル列 D_{21b} と D_{31b} 、 D_{31b} と D_{41b} ...についても同様である。図2において、 $D_{11b} \sim D_{41b}$ は、それぞれ音入力 $x(n)$ のサンプル列を半フレーム $T_1/2$ ($M/2$ サンプル)相当のタイミング $D_1 \sim D_4$ で切り出したもので、いずれも雑音抑圧用フレーム(雑音抑圧の対象となるフレーム) F_b として用いられる。

40

【0045】

フレーム切出部16は、音入力 $x(n)$ のサンプル列を比較的長いフレーム長 T_2 (N サンプル)で順次に切り出すものである。切出部16の切出動作は、あるフレームのサンプル列 D_{11a} の切り出しから次のフレームのサンプル列 D_{21a} の切り出しが半フレーム $T_1/2$ ($M/2$ サンプル)相当の時間だけ遅れるようにして行なわれる。このことは、サンプル列 D_{21a} と D_{31a} 、 D_{31a} と D_{41a} ...についても同様である。図2において、 $D_{11a} \sim D_{41a}$ は、それぞれ音入力 $x(n)$ のサンプル列を $T_1/2$ ($M/$

50

2 サンプル) 相当のタイミング $D_1 \sim D_4$ で切り出したもので、いずれも雑音推定用フレーム (雑音スペクトルの推定に用いられるフレーム) F_a となる。雑音推定用フレームのサンプル列 D_{11a} の切り出しと雑音抑圧用フレームのサンプル列 D_{11b} の切り出しとは、互いに同期して (時間的に頭 $H D_a$, $H D_b$ を揃えて) 行なわれる。このことは、サンプル列 D_{21a} と D_{21b} 、 D_{31a} と D_{31b} 、 D_{41a} と $D_{41b} \dots$ についても同様である。

【0046】

一例として、サンプリング周波数が 16 kHz である場合、フレーム長 T_1 を $M = 512$ サンプル (32 msec) とし、フレーム長 T_2 を $N = 4096$ サンプル (256 msec) とする ($N = 8M$ とする) ことができる。

10

【0047】

ゼロデータ発生部 34 は、フレーム長 T_1 , T_2 の差に相当する $(N - M) = 3584$ サンプルのゼロデータ D_0 を発生するものである。ゼロデータ D_0 は、加算器 36 に供給され、図 2 に示すように D_{11b} 等の各雑音抑圧用フレームのサンプル列の末尾 T_L に付加される。これは、雑音抑圧演算部 40 でのスペクトル減算処理においてデータ長 (サンプル数) を揃える必要があるためである。

【0048】

加算器 36 から送出される各フレームデータ (ゼロデータ D_0 が付加された各切出データ) は、FFT 部 38 に供給され、FFT 処理により周波数分析 (周波数領域のデータに変換) される。FFT 部 38 から送出される各フレームの周波数分析データ $X_1(k) = (k = 0, 1, 2 \dots N - 1)$ は、雑音抑圧演算部 40 に供給される。演算部 40 では、各フレーム毎に周波数分析データ $X_1(k)$ に基づいて振幅スペクトル $|X_1(k)|$ が算出され、各フレームの振幅スペクトル $|X_1(k)|$ は、雑音推定部 22 に供給される。

20

【0049】

一方、フレーム切出部 16 から送出される各フレームの切出データは、FFT 部 18 に供給され、FFT 処理により周波数分析される。FFT 部 18 から送出される各フレームの周波数分析データ $X_2(k) = (k = 0, 1, 2 \dots N - 1)$ は、振幅スペクトル演算部 20 に供給される。演算部 20 では、各フレーム毎に周波数分析データ $X_2(k)$ に基づいて振幅スペクトル $|X_2(k)|$ が算出され、各フレームの振幅スペクトル $|X_2(k)|$ は、雑音推定部 22 に供給される。雑音推定部 22 には、フレーム切出部 32 から各フレームの切出データ (各フレーム毎に 512 サンプルのデータ) も供給される。

30

【0050】

雑音推定部 22 では、ある振幅スペクトル $|N_2(k)|$ を現フレームの振幅スペクトルとすると、 $|N_2(k)|$ の前のフレーム (前フレーム) の振幅スペクトルに基づいて前フレームの雑音振幅スペクトルが推定されると共に、前フレームの雑音振幅スペクトルと現フレームの振幅スペクトル $|N_2(k)|$ とを後述するように混合して現フレームの雑音振幅スペクトル $|N(k)|$ が推定される。このようにして、雑音推定部 22 からは、各フレーム毎に雑音振幅スペクトル $|N(k)|$ が次々に送出され、各フレームの雑音振幅スペクトル $|N(k)|$ は、平滑化処理部 24 に供給される。

【0051】

40

平滑化処理部 24 では、各フレームの雑音振幅スペクトル $|N(k)|$ に平滑化処理が施される。平滑化処理のアルゴリズムとしては、例えば移動平均法を用いることができる。移動平均法では、所定数の連続した周波数ポイント (所定の周波数帯域幅) における振幅の平均値を該周波数帯域の中央の周波数ポイントの振幅値として置き換える。1 回の平均で使用する連続した周波数ポイントの点数 (平均値を求める周波数帯域幅) は、例えば 8 点とすれば、平滑化された振幅スペクトル (雑音推定用振幅スペクトル) の実質的な周波数分解能は、雑音抑圧用振幅スペクトルの実質的な周波数分解能に等しくなる。この平均値算出及び振幅値の置き換えを、周波数ポイントを 1 ポイントずつずらして実行し、全周波数帯域にわたり平滑化した振幅スペクトルを求める。

【0052】

50

平滑化処理部 24 における平滑化処理アルゴリズムとしては、移動平均法の他に、移動メディアン法を用いることもできる。移動メディアン法では、所定数（例えば 8 点）の連続した周波数ポイント（所定の周波数帯域幅）の中で、振幅値の中央値を該周波数帯域の中央の周波数ポイントの振幅値として置き換える。そして、この振幅値の中央値の抽出及び振幅値の置き換えを、周波数ポイントを 1 ポイントずつずらして実行し、全周波数帯域にわたり平滑化した振幅スペクトルを求める。

【0053】

平滑化処理部 24 においては、ディップ除去処理を追加してもよい。ディップ除去処理では、平滑化処理された振幅スペクトルと平滑化処理される前の振幅スペクトルとを比較して周波数ポイント毎に大きい方の値を選択し、選択に係る値をつなぐことによりディップが除去された振幅スペクトルを得る。ディップを除去することで処理雑音の低減が可能になる。平滑化処理部 24 からは、各フレーム毎に平滑化処理された（更にはディップ除去処理された）振幅スペクトル $|N'(k)|$ が送出され、雑音抑圧演算部 40 に供給される。

【0054】

演算部 40 では、周波数分析データ $X_1(k)$ に基づいて各フレーム毎に振幅スペクトル及び位相スペクトルが算出される。そして、現フレームの振幅スペクトルから現フレームの雑音振幅スペクトル $|N'(k)|$ を減算することにより雑音が抑圧された振幅スペクトルが生成される。この後、雑音抑圧に係る振幅スペクトルに対応する位相スペクトルと合成することにより各フレーム毎に音声スペクトル $G(k)$ が生成され、逆高速フーリエ変換（I - FFT）部 42 に供給される。

【0055】

I - FFT 部 42 は、演算部 40 からの音声スペクトル $G(k)$ を I - FFT 処理により時間領域の音声信号に変換するもので、変換に係る時間領域の音声信号は、フレーム合成部 44 に供給される。

【0056】

フレーム合成部 44 では、図 2 に示すように各フレームの音声信号から先に付加された $(N - M)$ サンプルのゼロデータが削除されると共に各フレームの音声信号 (M) サンプルのデータに三角窓関数が乗算される。図 2 において、 $D1s \sim D4s$ は、いずれも三角窓関数が乗算された音声信号であり、フレーム合成に用いられるフレーム Fs となる。この後、図 2 に示すように順次の 2 フレームの音声信号を半フレームずつオーバーラップさせるように加算することによりフレーム合成が行なわれる。この結果、雑音が抑圧された音声出力 $g(n)$ がフレーム間に段差や切れ目がない連続した状態で得られる。

【0057】

次に、図 3 を参照して雑音推定部 22 の構成及び動作を説明する。スペクトル包絡抽出部 54 は、図 1 の雑音抑圧演算部 40 から供給される振幅スペクトル $|X_1(k)|$ に含まれる細かな凹凸特性を除去してスペクトル包絡 $|X_1'(k)|$ を抽出するものである。これは、後述する相関値算出において振幅スペクトル $|X_1(k)|$ そのものを用いると、スペクトルの相関値が低くなり、「音声区間」と「雑音区間」の区別が明確でなくなるためである。すなわち、雑音は、長時間観測を繰り返してそのスペクトルを平均してみれば、そのスペクトルは、広い帯域にわたってほぼ一様となる滑らかな分布となるが、短時間で見れば多くの山谷を有するスペクトルの変動が観察される。一方、音声は、雑音とは異なり、その全体的な周波数特性は特定の周波数帯域に大きな振幅値を持っており、全周波数帯域に一様に分布していない。この実施形態では、「全周波数帯域に一様に分布する雑音」と、「ある特定の周波数帯域に大きな振幅値を持つ音声」とをスペクトル包絡の相関値の大小で区別して雑音振幅スペクトルを推定するので、雑音振幅スペクトルが持っている細かな凹凸特性を除去するようにしている。

【0058】

スペクトル包絡抽出部 54 では、一例として、振幅スペクトル $|X_1(k)|$ を時間波形と見立ててローパスフィルタ処理をすることによりスペクトル包絡を抽出する。ローパ

10

20

30

40

50

スフィルタ処理は、振幅スペクトル $|X_1(k)|$ を直接ローパスフィルタにかける、あるいは振幅スペクトル $|X_1(k)|$ に移動平均処理を施す等の方法により行なうことができる。また、スペクトル包絡抽出部 54 においてスペクトル包絡 $|X_1'(k)|$ を抽出する別の方法としては、振幅スペクトル $|X_1(k)|$ を更にフーリエ変換してケプストラムによって求める方法を用いることもできる。抽出部 54 で抽出された各フレームのスペクトル包絡 $|X_1'(k)|$ は、相関値算出部 56 に供給される。

【0059】

雑音振幅スペクトル初期値出力部 46 は雑音振幅スペクトルの初期値を出力するものである。すなわち、システムの起動当初は、参照する雑音振幅スペクトルデータがないため、初期値を設定する。雑音振幅スペクトル初期値の設定方法としては、例えば、次の方法が考えられる。

10

【0060】

(方法1) 起動直後に入力され、音声の混入していない背景雑音のみのデータをフーリエ変換し、該フーリエ変換されたデータから求められる振幅スペクトルデータを雑音振幅スペクトル初期値として設定する。

【0061】

(方法2) 予め背景雑音に相当する振幅スペクトルデータをメモリに保持しておき、起動時にそれを読み出して雑音振幅スペクトル初期値として設定する。あるいは、予め背景雑音に相当する振幅スペクトルの包絡データをメモリに保持しておき、起動時にそれを読み出して雑音振幅スペクトル包絡データの初期値として設定する。

20

【0062】

(方法3) ホワイトノイズやピンクノイズの振幅スペクトルデータを雑音振幅スペクトル初期値として設定する。

【0063】

雑音振幅スペクトル遅延部 48 は、後述する雑音振幅スペクトル算出部 50 で半フレーム ($T_1/2$) 毎に求められる雑音振幅スペクトル $|N(k)|$ を順次入力し、半フレーム分遅延して前フレームの雑音振幅スペクトル $|N_o(k)|$ として順次出力するものである。システムの起動当初は雑音振幅スペクトル $|N(k)|$ は未だ推定されていないので、遅延部 48 は、出力部 46 で設定された雑音振幅スペクトル初期値を雑音振幅スペクトル $|N_o(k)|$ としてスペクトル包絡抽出部 52 へ供給する。抽出部 52 は、前述したスペクトル包絡抽出部 54 と同様の方法により、雑音振幅スペクトル $|N_o(k)|$ のスペクトル包絡 $|N_o'(k)|$ を抽出し、相関値算出部 56 へ供給する。

30

【0064】

相関値算出部 56 は、スペクトル包絡抽出部 54 で抽出された現フレームのスペクトル包絡 $|X_1'(k)|$ と、スペクトル包絡抽出部 52 で抽出された前フレームのスペクトル包絡 $|N_o'(k)|$ との相関を求め、相関値 を算出するものである。相関値 は、

$$\text{スペクトル包絡 } |X_1'(k)| = x_k \quad (k = 1, 2, \dots, K)$$

$$\text{スペクトル包絡 } |N_o'(k)| = y_k \quad (k = 1, 2, \dots, K)$$

とすると、次の数1の式(1)により求められ、0～1の範囲内の値をとる。

【0065】

40

【数 1】

$$\rho = \frac{C_{XY}}{\sqrt{C_{XX}} \sqrt{C_{YY}}} \dots (1)$$

ここで、

$$C_{XY} = \sum_{k=1}^K \left\{ (x_k - (\sum_{k=1}^K x_k)/K) (y_k - (\sum_{k=1}^K y_k)/K) \right\}$$

10

$$C_{XX} = \sum_{k=1}^K (x_k - (\sum_{k=1}^K x_k)/K)^2$$

$$C_{YY} = \sum_{k=1}^K (y_k - (\sum_{k=1}^K y_k)/K)^2$$

【0066】

20

雑音振幅スペクトル算出部50は、算出部56から供給される相関値と、セクタ76から供給される更新制御値 x と、遅延部48から供給される前フレームの雑音振幅スペクトル $|N_o(k)|$ と、セクタ66から供給される現フレームの音入力の振幅スペクトル $|X(k)|$ （これは、後述するように $|X_2(k)|$ 又は $|X_1(k)|$ からなる）に基づいて次の式（2）により現フレームの雑音振幅スペクトル $|N(k)|$ を推定し、算出するものである。

【0067】

$$|N(k)| = (1 - x) \cdot |N_o(k)| + x \cdot |X(k)| \dots (2)$$

式（2）は、前フレームの雑音振幅スペクトル $|N_o(k)|$ と現フレームの音入力の振幅スペクトル $|X(k)|$ とを相関値及び更新制御値 x に応じた比率で加算（混合）して現フレームの雑音振幅スペクトル $|N(k)|$ を推定することを示すものである。式（2）によれば、相関値が小さい有音声区間では、相関値が大きいときに比べて、雑音振幅スペクトル $|N_o(k)|$ の比率を高くすると共に振幅スペクトル $|X(k)|$ の比率を低くして加算が行なわれる。このため、音声成分が多い有音声区間では、推定に係る雑音振幅スペクトル $|N(k)|$ が音声成分の影響によりさほど変化しない。これに対し、相関値が大きい無音声区間では、相関値が小さいときに比べて、雑音振幅スペクトル $|N_o(k)|$ の比率を低くすると共に振幅スペクトル $|X(k)|$ の比率を高くして加算が行なわれる。このため、無音声区間では、推定に係る雑音振幅スペクトル $|N(k)|$ が新たな雑音成分に応じて変化し、 $|N(k)|$ の更新が促進される。

30

【0068】

40

パワー算出部60は、遅延部48からの前フレームの雑音振幅スペクトル $|N_o(k)|$ に基づいて前フレーム内の平均雑音パワー P_n を算出するものである。平均雑音パワー P_n は、前フレームの雑音振幅スペクトルにおける振幅値を $A_1 \sim A_N$ （ $N = 2048$ ）とすると、 $(A_1^2 + A_2^2 + A_3^2 + \dots + A_N^2) / N$ なる式で求められる。平均雑音パワー P_n を示すパワーデータは比較部64に比較入力 A として供給される。

【0069】

パワー算出部62は、図1の切出部32からの現フレームの切出データ D_b に基づいて現フレーム内の平均音パワー P_b を算出するものである。平均音パワー P_b は、現フレームの切出データ D_b をFFTすることにより求めた振幅スペクトルにおいて振幅値を $a_1 \sim a_M$ （ $M = 256$ ）とすると、 $(a_1^2 + a_2^2 + a_3^2 + \dots + a_M^2) / M$ なる式で求

50

められる。平均音パワー P_b を示すパワーデータは比較部 64 に比較入力 B として供給される。なお、周波数スペクトルでパワーを算出する場合、FFT フレーム長の半分以上のサンプル数とするのが好ましい。

【0070】

比較部 64 は、比較入力 A, B (パワー P_n , P_b) を比較し、 $A > B$ ($P_n > P_b$) ならば比較出力 $C_1 = 1$ を、 $A < B$ ($P_n < P_b$) ならば比較出力 $C_2 = 1$ をそれぞれ送出する。ここで、 $C_1 = 1$ は、前フレーム内の平均雑音パワー P_n の方が現フレーム内の平均音パワー P_b より大きいので、雑音の減少を意味する。また、 $C_2 = 1$ は、前フレーム内の平均雑音パワー P_n の方が現フレーム内の平均音パワー P_b より小さいので、雑音の増大を意味する。

10

【0071】

セクタ 66 は、選択入力 $SB = 0$ のときは入力 A としての振幅スペクトル $|X_2(k)|$ を選択し、算出部 50 に振幅スペクトル $|X(k)|$ として供給する。比較部 64 の比較出力 C_1 が 1 になると、セクタ 66 は、 $C_1 = 1$ からなる選択入力 SB に応じて入力 B としての振幅スペクトル $|X_1(k)|$ を選択し、算出部 50 に振幅スペクトル $|X(k)|$ として供給する。前述したように、振幅スペクトル $|X_2(k)|$ は、振幅スペクトル $|X_1(k)|$ に比べてサンプル数が 8 倍も多いので、 $A > B$ ($P_n > P_b$) 以外の場合は、算出部 50 において $|X(k)|$ として $|X_2(k)|$ を用いることで精度良く雑音スペクトルの推定を行なうことができる。しかしながら、 $A > B$ ($P_n > P_b$) である雑音減少時にも振幅スペクトル $|X_2(k)|$ を用いると、振幅スペクトル $|X_2(k)|$ のフレーム長 (サンプル数) が長いため、過去の推定結果の影響から脱却するのに時間がかかる。そこで、この発明では、雑音減少時には、 $|X(k)|$ として $|X_2(k)|$ の代りにフレーム長の短い $|X_1(k)|$ を用いることで雑音減少に対する追従性を改善している。

20

【0072】

比較部 64 からの比較出力 $C_1 = 1$ は、セクタ 76 に選択入力 SC として供給される。セクタ 76 は、選択入力 SB , SC のいずれも 0 のときは、入力値 $A = 13.5$ を選択し、算出部 50 に更新制御値 x として供給する。 $x = 13.5$ のときは、前述の式 (2) において x が例えば 0.5 であれば x の値は極めて小さな値となり、雑音振幅スペクトル $|N(k)|$ の更新が抑制される。この発明では、 $C_1 = 1$ となる雑音減少時において、セクタ 76 で選択入力 $SC = C_1 = 1$ に応じて入力値 1 を選択し、算出部 50 に更新制御値 x として供給する。この結果、算出部 50 では、前述の式 (2) において $x = 1$ であり且つ x が例えば 0.5 であれば $|N(k)| = \{ |N_o(k)| + |X_1(k)| \} / 2$ となり、雑音振幅スペクトル $|N(k)|$ の更新が促進される。従って、雑音減少に対する追従性が $x = 1$ としたことでも改善される。

30

【0073】

記憶部 68 は、算出部 62 で算出された平均音パワー P_b を示すパワーデータを現フレーム分とその前の 32 フレーム分 (合計 33 フレーム分) 記憶するものである。このように現フレームまでの一定期間分のパワーデータを記憶するのは、雑音増大の検出を確実にするためである。最小値検出部 70 は、記憶部 68 に記憶された 33 フレーム分のパワーデータ中からフレーム内平均音パワーの最小値を検出するもので、検出に係る最小値のパワー P_m を示すパワーデータを比較部 72 に比較入力 B として供給する。比較部 72 の比較入力 A としては、算出部 60 から前フレーム内の平均雑音パワー P_n を示すパワーデータが供給される。

40

【0074】

比較部 72 は、比較入力 A, B (パワー P_n , P_m) を比較し、 $A < B$ ($P_n < P_m$) ならば比較出力 $C_3 = 1$ を AND ゲート 74 に一方の入力として供給する。AND ゲート 74 の他方の入力としては、比較部 64 から $A < B$ ($P_n < P_b$) に応じて比較出力 $C_2 = 1$ が供給される。AND ゲート 74 では、前フレーム内の平均雑音パワー P_n の方が現フレーム内の平均音パワー P_b より小さく且つ前フレーム内の平均雑音パワー P_n の方が

50

検出部 70 での検出に係る最小値のパワー P_m より小さいときに AND 条件が成立し、出力 $AD = 1$ が発生される。この出力 $AD = 1$ の発生は、雑音の増大を意味する。

【0075】

セレクタ 76 は、AND ゲート 74 からの出力 $AD = 1$ からなる選択入力 SB に応じて入力値 7 を選択し、算出部 50 に更新制御値 x として供給する。このとき、算出部 50 には、振幅スペクトル $|X(k)|$ としてセレクタ 66 からフレーム長が長い振幅スペクトル $|X_2(k)|$ が供給されている。すなわち、算出部 50 では、前述の式 (2) において $x = 7$ であるので、 $|N(k)| = (1 - 7) \cdot |N_0(k)| + 7 \cdot |X_2(k)|$ となり、 $x = 13.5$ の場合に比べて $|N(k)|$ の更新が促進される。従って、雑音増大に対する追従性が改善される。

10

【0076】

次に、図 4 を参照して雑音抑圧演算部 40 の構成及び動作を説明する。振幅スペクトル計算部 56 及び位相スペクトル計算部 58 には、図 1 の FFT 部 38 から周波数分析データ $X_1(k)$ が共通に供給される。計算部 56 は、周波数分析データ $X_1(k)$ に基づいて次の式 (3) により振幅スペクトル $|X_1(k)|$ を算出する。

【0077】

$$|X_1(k)| = \{X_R(k)^2 + X_I(k)^2\}^{1/2} \dots (3)$$

ここで、 $X_R(k)$: $X_1(k)$ の実数部

$X_I(k)$: $X_1(k)$ の虚数部

また、計算部 58 は、周波数分析データ $X_1(k)$ に基づいて次の式 (4) により位相スペクトル $\angle X_1(k)$ を算出する。

20

【0078】

$$\angle X_1(k) = \tan^{-1} \{X_I(k) / X_R(k)\} \dots (4)$$

スペクトル減算部 60 は、次の式 (5) に従って減算を行なう。すなわち、計算部 56 で求めた現フレームの振幅スペクトル $|X_1(k)|$ から、図 1 の平滑化処理部 24 から供給される現フレームの雑音振幅スペクトル $|N'(k)|$ を減算することにより、雑音振幅スペクトルが除去された現フレームの振幅スペクトル $|Y(k)|$ を求める。

【0079】

$$|Y(k)| = |X_1(k)| - |N'(k)| \dots (5)$$

なお、 $|X_1(k)| - |N'(k)|$ が負の値となる周波数ポイントでは、引き過ぎであるので、減算値 $|Y(k)|$ を負の値のままとせず、ゼロとするのがよい。

30

【0080】

合成部 62 は、減算部 60 で求めた現フレームの振幅スペクトル $|Y(k)|$ と、計算部 58 で求めた現フレームの位相スペクトル $\angle X_1(k)$ とを合成して次の式 (6) に示す複素スペクトル (雑音が抑圧された音声スペクトル) $G(k)$ を作成する。

【0081】

$$G(k) = |Y(k)| e^{j\angle X_1(k)} \dots (6)$$

作成された音声スペクトル $G(k)$ は、図 1 の IFFT 部 42 に供給される。

【0082】

図 5 は、上記した雑音抑圧装置について雑音抑圧性能をテストした結果を示すものである。図 5 (A) は、クリーンスピーチ区間 Z1 - ノイジスピーチ区間 Z2 - クリーンスピーチ区間 Z3 について入力音信号の振幅変化を示し、図 5 (B) は、図 5 (A) と同様の区間 Z1 - Z2 - Z3 について音声パワー SP 及び推定雑音パワー NP の変化を示すものである。ここで、「クリーンスピーチ区間」とは、雑音がない会話区間であり、「ノイジスピーチ区間」とは、雑音がある会話区間である。図 5 (B) によれば、クリーンスピーチ区間 Z1 からノイジスピーチ区間 Z2 に至る雑音増大部で推定雑音パワー NP が追従性良く増大すると共にノイジスピーチ区間 Z2 からクリーンスピーチ区間 Z3 に至る雑音減少部で推定雑音パワー NP が追従性良く減少しているのがわかる。

40

【0083】

上記した実施形態の変形例としては、前述した推定式 (2) の代りに次の推定式 (7)

50

を用いてもよい。

【0084】

$$|N(k)| = [1 - \{x / (1 + x)\}^m] \cdot |No(k)| + \{x / (1 + x)\}^m \cdot |X(k)| \dots (7)$$

ここで、 $|N(k)|$ ：現フレームの雑音振幅スペクトル

$|No(k)|$ ：前フレームの雑音振幅スペクトル

$|X(k)|$ ：現フレームの音入力の振幅スペクトルであって、 $|X_1(k)|$ 又は $X_2(k)$ からなるもの

：スペクトル包絡 $|No'(k)|$ 、 $|X_1'(k)|$ の相関値

x 、 m ：更新制御値（ x は1以上の値、 m は0以上の値）

式(7)は、前フレームの雑音振幅スペクトル $|No(k)|$ と、現フレームの音入力の振幅スペクトル $|X(k)|$ とを相関値及び更新制御値 x 、 m に応じた比率で加算（混合）して現フレームの雑音振幅スペクトル $|N(k)|$ を推定することを示すものである。式(7)によれば、相関値が小さい有音声区間では、相関値が大きいときに比べて、雑音振幅スペクトル $|No(k)|$ の比率を高くすると共に振幅スペクトル $|X(k)|$ の比率を低くして加算が行なわれる。このため、音声成分が多い有音声区間では、推定に係る雑音振幅スペクトル $|N(k)|$ が音声成分の影響によりさほど変化しない。これに対し、相関値が大きい無音声区間では、相関値が小さいときに比べて、雑音振幅スペクトル $|No(k)|$ の比率を低くすると共に振幅スペクトル $|X(k)|$ の比率を高くして加算が行なわれる。このため、無音声区間では、推定に係る雑音振幅スペクトル $|N(k)|$ が新たな雑音成分に応じて変化し、 $|N(k)|$ の更新が促進される。

【0085】

式(7)において、 x 、 m は、 $|N(k)|$ の更新を制御するためのそれぞれ第1、第2の更新制御値である。図6は、相関値に対する係数値 $[1 - \{x / (1 + x)\}^m]$ 、 $\{x / (1 + x)\}^m$ の変化を第1の更新制御値 x をパラメータとして示すもので、第2の更新制御値 m を1とした場合である。図6によれば、例えば $x = 0.8$ としたとき x が大きいほど $\{x / (1 + x)\}^m$ が小さくなり、 $|N(k)|$ の更新が抑制されることがわかる。

【0086】

図7は、相関値に対する係数値 $[1 - \{x / (1 + x)\}^m]$ 、 $\{x / (1 + x)\}^m$ の変化を第2の更新制御値 m をパラメータとして示すもので、第1の更新制御値 x を2とした場合である。図7によれば、例えば $x = 0.8$ としたとき m が大きいほど $\{x / (1 + x)\}^m$ が小さくなり、 $|N(k)|$ の更新が抑制されることがわかる。

【0087】

式(7)を用いてこの発明を実施する際には、図3の回路において、セクタ76では、選択入力SB、SCのいずれも0のときは例えば $x = 70$ を算出部50に供給し、SB = 1のときは $x = 8$ を算出部50に供給し、SC = 1のときは $x = 2$ を算出部50に供給する。別の方法としては、SB、SCのいずれも0のときは $m = 4$ を算出部50に供給し、SB = 1のときは $m = 2$ を算出部50に供給し、SC = 1のときは $m = 1$ を算出部50に供給する。このように x 又は m を減少させると、雑音振幅スペクトル $|N(k)|$ の更新が促進される。なお、ここで述べた x を用いる方法と、 m を用いる方法とは適宜併用してもよい。

【0088】

この発明は、上記した実施形態に限定されるものではなく、種々の改変形態で実施可能である。例えば、次のような変更が可能である。

【0089】

(a) 上記した実施形態では、低相関時に現フレームの音入力の振幅スペクトル $|X(k)|$ の混合比率を低減するようにしたが、相関値が所定値以下のときは、スペクトル $|X(k)|$ の混合比率をゼロとする（すなわち、雑音振幅スペクトル $|N(k)|$ を更

新しない)ようにしてもよい。

【0090】

(b) 上記した実施形態では、雑音推定処理を所定時間間隔 ($T_1 / 2$ 時間毎) に必ず行なうようにしたが、適宜の時間毎に行なうようにしてもよい。例えば、無音声区間、微小音声区間等の雑音推定が容易な区間を実時間で検出し、該雑音推定が容易な区間でのみ雑音推定処理を行ない、それ以外の区間では雑音推定処理を行なわない(一時停止する)ことができる。また、雑音変動が少ない区間や処理負荷を減らしたい区間も雑音推定処理を行なわない(一時停止する)ことができる。これらの場合、雑音推定処理を一時停止している区間では、雑音振幅スペクトル遅延部 48 に保持されている最新の(一時停止直前の)雑音振幅スペクトル $|N_o(k)|$ に基づいて雑音抑圧処理を行なうことができる。

10

【0091】

(c) 上記した実施形態では、音入力から切り出す雑音抑圧用フレームの長さ T_1 (M サンプル分の時間)を、切り出しを行なう時間間隔 ($M / 2$ サンプル分の時間)より長く設定したが、これはフレーム合成の際にオーバーラップ処理を行なうためであり、オーバーラップ処理を行なわない場合には、フレーム長 T_1 を切り出し時間間隔と等しく設定してもよい。

【0092】

(d) 上記した実施形態では、周波数分析手法として FFT を用いた場合について説明したが、この発明では、FFT 以外の周波数分析手法を用いることもできる。

【0093】

20

(e) 上記した実施形態では、振幅スペクトルサブトラクション法を用いる代りに、パワースペクトルサブトラクション法を用いることもできる。この場合、音入力のパワースペクトル $|X_1(k)|^2$ のスペクトル包絡に基づいて雑音パワースペクトル $|N(k)|^2$ を推定し、音入力のパワースペクトル $|X_1(k)|^2$ から推定に係る雑音のパワースペクトル $|N(k)|^2$ を減算して雑音抑圧を行なう。雑音のパワースペクトル $|N(k)|^2$ の推定処理にこの発明を適用できる。

【0094】

(f) 上記した実施形態では、振幅スペクトルサブトラクション法を用いる代りに、複素スペクトルサブトラクション法を用いることもできる。この場合、音入力の複素スペクトル(振幅情報と位相情報とに分離されていないもの) $X_1(k)$ のスペクトル包絡に基づいて雑音の複素スペクトル $N(k)$ を推定し、音入力の複素スペクトル $X_1(k)$ から推定に係る雑音の複素スペクトル $N(k)$ を減算して雑音抑圧を行なう。雑音の複素スペクトル $N(k)$ の推定処理にこの発明を適用できる。

30

【0095】

(g) 上記した実施形態では、前フレーム内の平均雑音パワー P_n が現フレーム内の平均音パワー P_b より大きいとき、セクタ 66 でフレーム長が短い振幅スペクトル $|X_1(k)|$ を選択すると共にセクタ 76 で更新制御値 $x = 1$ を選択するようにしたが、 $|X_1(k)|$ の選択又は $x = 1$ の選択のいずれか一方を省略することもできる。 $|X_1(k)|$ の選択を省略した場合、算出部 50 には、 $|X(k)|$ として $|X_2(k)|$ を供給すればよい。

40

【0096】

(h) 上記した実施形態では、切出部 32 で切り出した 512 サンプルのデータに基づいてフレーム内の平均音パワーを算出部 62 で算出したが、算出部 62 では、切出部 16 で切り出した 4096 サンプルのデータに基づいてフレーム内の平均音パワーを算出するようにしてもよい。

【0097】

(i) 上記した実施形態では、スペクトルの時間変化により敏感な短いフレーム長の振幅スペクトル $|X_1(k)|$ からスペクトル包絡を抽出するようにしたので、より時間反応のよい雑音振幅スペクトルを算出できる。しかしながら、スペクトル包絡の抽出対象は、振幅スペクトル $|X_1(k)|$ に限らず、振幅スペクトル $|X_2(k)|$ でもよく、あ

50

るいはユーザが $|X_1(k)|$ 又は $|X_2(k)|$ を任意に切換えできるようにしてもよい。

【0098】

(j) 雑音増大判定処理において、記憶した過去のパワーデータに基づく処理(図3のANDゲート74の右側入力に至る処理)は省略してもよい。また、雑音減少判定処理において、過去のパワーデータに基づく処理を採用するようにしてもよい。

【0099】

(k) この発明の雑音スペクトル推定方法は、雑音抑圧以外の用途にも適用することができる。また、この発明は、小型コンピュータ等を用いて実施することもできる。

【図面の簡単な説明】

10

【0100】

【図1】この発明の一実施形態に係る雑音抑圧装置の構成を示すブロック図である。

【図2】雑音抑圧装置の動作を説明するためのタイムチャートである。

【図3】雑音推定部を示すブロック図である。

【図4】雑音抑圧演算部を示すブロック図である。

【図5】(A)は、クリーンスピーチ区間及びノイジスピーチ区間について音声信号の振幅変化を示す波形図、(B)は、(A)と同様の区間について音声パワー及び推定雑音パワーの変化を示す波形図である。

【図6】変形例に係る雑音推定式における相関値 に対する2種類の係数値の変化を第1の更新制御値 x をパラメータとして示すグラフである。

20

【図7】変形例に係る雑音推定式における相関値 に対する2種類の係数値の変化を第2の更新制御値 m をパラメータとして示すグラフである。

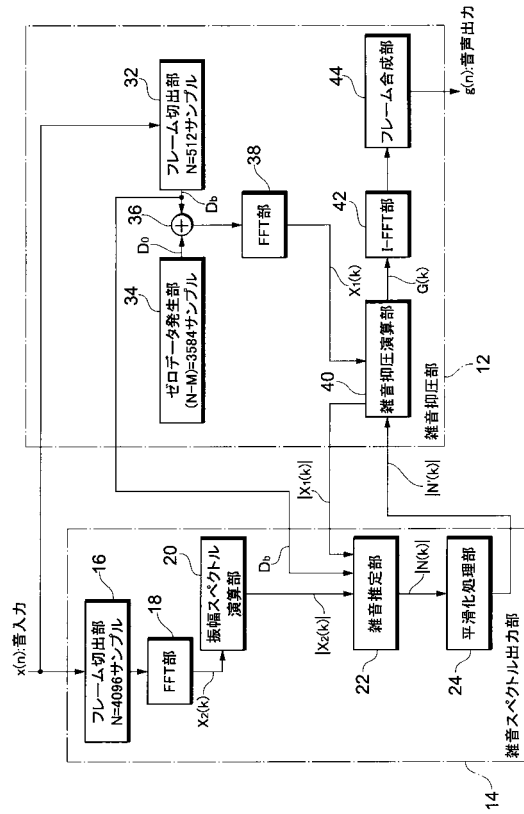
【図8】発明者の研究に係る雑音抑圧処理を示すフローチャートである。

【符号の説明】

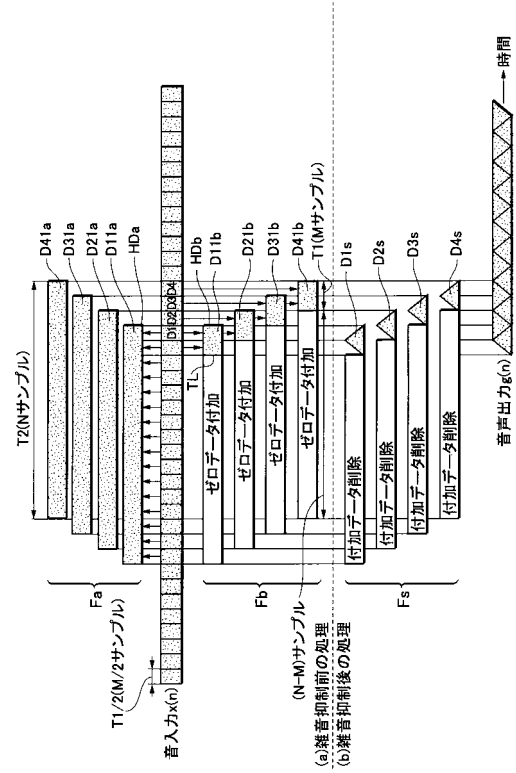
【0101】

12：雑音抑圧部、14：雑音スペクトル出力部、22：雑音推定部、40：雑音抑圧演算部、60, 62：パワー算出部、64, 72：比較部、66, 76：セレクト、68：記憶部、70：最小値検出部、74：ANDゲート。

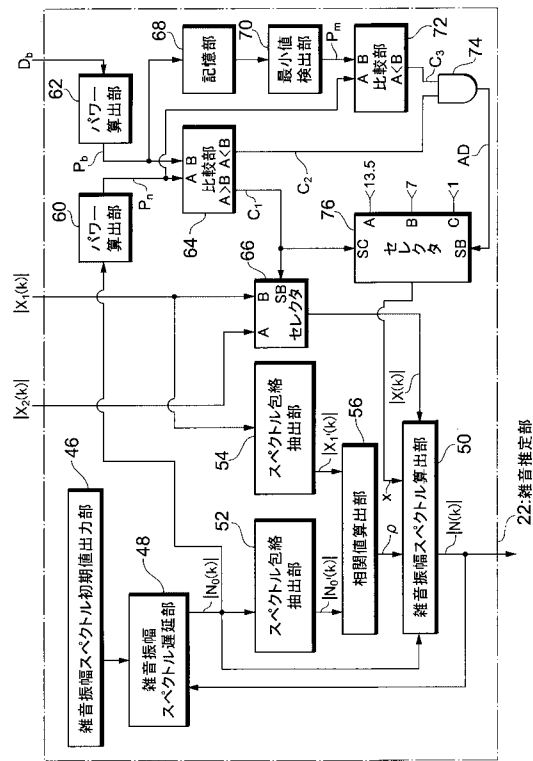
【図 1】



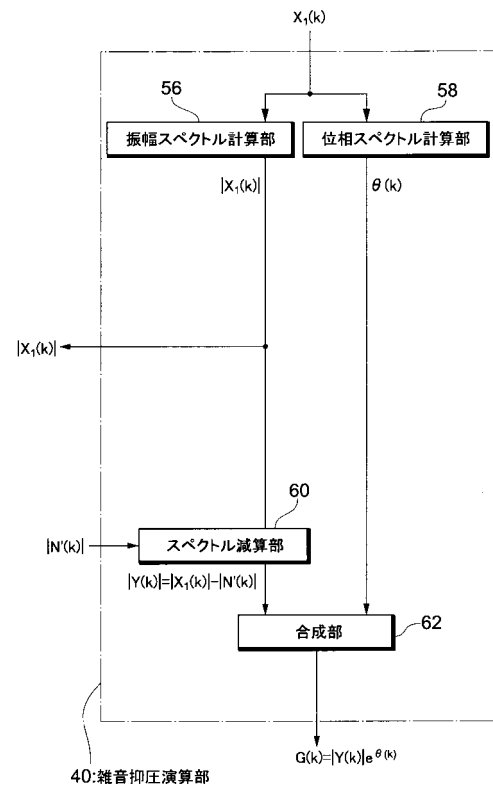
【図 2】



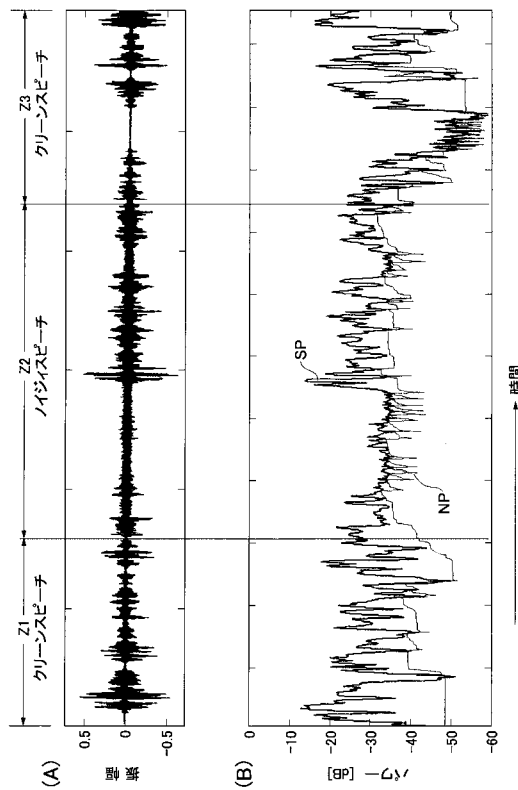
【図 3】



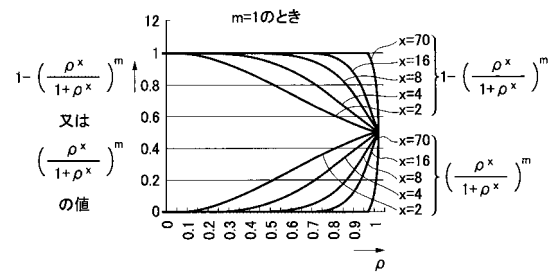
【図 4】



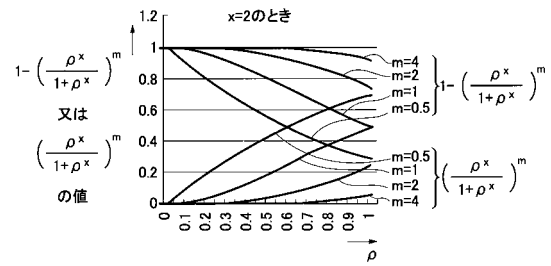
【図5】



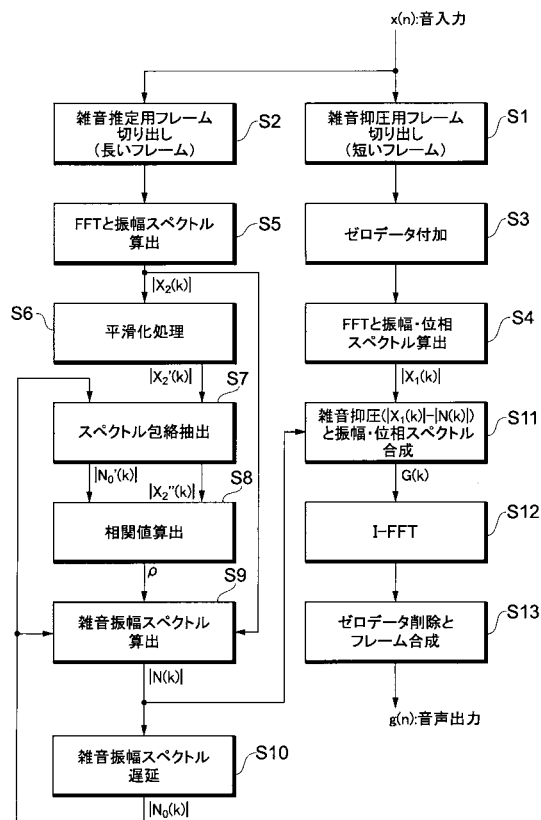
【図6】



【図7】



【図8】



フロントページの続き

審査官 涌井 智則

(56)参考文献 特開2004-020679(JP,A)
国際公開第2006/123721(WO,A1)
特開2005-284163(JP,A)
特開2005-099405(JP,A)
特開2001-134286(JP,A)
特許第3591068(JP,B2)
特開2005-077731(JP,A)

(58)調査した分野(Int.Cl., DB名)
G10L 11/00-21/06
Cini