

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5812998号
(P5812998)

(45) 発行日 平成27年11月17日(2015.11.17)

(24) 登録日 平成27年10月2日(2015.10.2)

(51) Int.Cl.	F I
G 1 0 L 21/0388 (2013.01)	G 1 0 L 21/0388 1 0 0
G 1 0 L 19/02 (2013.01)	G 1 0 L 19/02 1 5 0

請求項の数 14 (全 17 頁)

(21) 出願番号	特願2012-539847 (P2012-539847)	(73) 特許権者	598036300
(86) (22) 出願日	平成22年6月29日 (2010.6.29)		テレフオンアクチーボラゲット エル エム エリクソン (パブル)
(65) 公表番号	特表2013-511741 (P2013-511741A)		スウェーデン国 ストックホルム エスー164 83
(43) 公表日	平成25年4月4日 (2013.4.4)	(74) 代理人	100109726
(86) 国際出願番号	PCT/SE2010/050746		弁理士 園田 吉隆
(87) 国際公開番号	W02011/062535	(74) 代理人	100101199
(87) 国際公開日	平成23年5月26日 (2011.5.26)		弁理士 小林 義教
審査請求日	平成25年5月31日 (2013.5.31)	(72) 発明者	グランシャロヴ, ヴォロージャ
(31) 優先権主張番号	61/262, 714		スウェーデン国 エスー171 67 ソルナ, アンクダムスガタン 29
(32) 優先日	平成21年11月19日 (2009.11.19)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 オーディオコーデックにおけるラウドネスおよびシャープネスの補償のための方法および装置

(57) 【特許請求の範囲】

【請求項1】

所定の帯域幅に限られた復元スピーチ信号について知覚されるラウドネスおよびシャープネスを改善する方法であって、

前記スピーチ信号を用意するステップ(S10)と、

少なくとも前記所定の帯域幅のうちの第1の帯域部分にもとづく第1の信号部分と、前記所定の帯域幅のうちの第2の帯域部分にもとづく第2の信号部分とに、前記スピーチ信号を分割するステップ(S20)と、

前記第1の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調するように、前記第1の信号部分を調整するステップ(S30)と、

少なくとも前記調整した第1の信号部分にもとづいて前記第2の信号部分を復元するステップ(S40)と、

前記調整した第1の信号部分と前記復元した第2の信号部分とを組み合わせる(S50)、全体として知覚されるラウドネスおよびシャープネスが改善された復元スピーチ信号を生成するステップと

を含む方法。

【請求項2】

前記調整するステップ(S30)が、

前記第1の信号部分をフィルタ処理するステップを含み、前記第1の帯域部分の選択された周波数に向かって第1の信号部分のエネルギーの少なくとも一部を分配すると同時に、

前記第 1 の帯域部分の選択された高周波数区間に向かって第 1 の信号部分のエネルギーの少なくとも別の一部を分配する、請求項 1 に記載の方法。

【請求項 3】

前記フィルタ処理するステップ (S30) が以下のフィルタ関数 $H(z)$:

$$H(z) = \alpha \cdot z^{-2} + \beta \cdot z^{-1} - \gamma + \delta \cdot z^{+1} + \epsilon \cdot z^{+2}$$

に従って実行される、請求項 2 に記載の方法。

【請求項 4】

係数 $\alpha = 0.1$ 、 $\beta = 0$ 、 $\gamma = 0.85$ である、請求項 3 に記載の方法。

【請求項 5】

前記フィルタ処理するステップ (S30) が以下のフィルタ関数 $H(z)$:

$$H(z) = \alpha \cdot z^{-1} - \beta + \gamma \cdot z^{+1}$$

に従って実行される、請求項 2 に記載の方法。

【請求項 6】

係数 $\alpha = 0.06$ および $\beta = 0.66$ である、請求項 5 に記載の方法。

【請求項 7】

前記フィルタ処理するステップ (S30) が以下のフィルタ関数 $H(z)$:

$$H(z) = 1 - \mu \cdot z^{-1}$$

に従って実行される、請求項 2 に記載の方法。

【請求項 8】

係数 $\mu = 0.2$ である、請求項 7 に記載の方法。

【請求項 9】

自然の外耳 - 中耳の反応にもとづいて、前記第 1 の帯域部分の範囲内の前記周波数を選択するさらなるステップを含む、請求項 2 に記載の方法。

【請求項 10】

前記第 1 の帯域部分が、前記用意されるスピーチ信号の低周波数帯 (LB) に相当し、前記第 2 の帯域部分が、前記用意されるスピーチ信号の高周波数帯 (HB) に相当する、請求項 1 ~ 9 のいずれか 1 項に記載の方法。

【請求項 11】

前記調整するステップ (S30) が、低周波数帯 (LB) をプレフィルタ処理するステップにもとづいており、前記第 2 の信号部分を復元するステップ (S40) が、帯域拡張 (BWE) または低域通過フィルタ処理にもとづく、請求項 10 に記載の方法。

【請求項 12】

所定の帯域幅に限られた復元スピーチ信号について知覚されるラウドネスおよびシャープネスを改善するためのシステムであって、

前記スピーチ信号を生成する手段 (10) と、

少なくとも前記所定の帯域幅のうちの第 1 の帯域部分にもとづく第 1 の信号部分と、前記所定の帯域幅のうちの第 2 の帯域部分にもとづく第 2 の信号部分とに、前記スピーチ信号を分割する手段 (20) と、

前記第 1 の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調するように、前記第 1 の信号部分を調整する手段 (30) と、

少なくとも前記調整した第 1 の信号部分にもとづいて前記第 2 の信号部分を復元する手段 (40) と、

前記調整した第 1 の信号部分と前記復元した第 2 の信号部分とを組み合わせ、全体として知覚されるラウドネスおよびシャープネスが改善された復元スピーチ信号を生成する手段 (50) と

を備えるシステム。

【請求項 13】

前記手段 (30) が、プレフィルタ処理によって前記第 1 の信号部分を調整するように構成され、前記第 1 の信号部分が、前記スピーチ信号の低周波数帯 (LB) に相当し、前記手段 (40) が、帯域拡張 (BWE) または低域通過フィルタ処理にもとづいて、前記

10

20

30

40

50

スピーチ信号の高周波数帯（HB）を復元する、請求項 1 2 に記載のシステム。

【請求項 1 4】

通信システムにおいて所定の帯域幅に限られたスピーチ信号を処理するためのデコーダ装置（2）であって、

少なくとも前記所定の帯域幅のうちの第 1 の帯域部分にもとづく第 1 の信号部分と、前記所定の帯域幅のうちの第 2 の帯域部分にもとづく第 2 の信号部分とに、生成されたスピーチ信号を分割することによって得られる第 1 の信号部分を受信する手段（25）と、

前記第 1 の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調するように、前記受信される第 1 の信号部分を調整する手段（30）と、

少なくとも前記調整した第 1 の信号部分にもとづいて前記第 2 の信号部分を復元する手段（40）と、

前記調整済みの第 1 の信号部分と前記復元される第 2 の信号部分とを組み合わせ、全体として知覚されるラウドネスおよびシャープネスが改善された復元スピーチ信号を生成する手段（50）と

を備えるデコーダ装置（2）。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、広くにはオーディオコーディング/デコーディングに関し、とくにはオーディオコーディングにおけるラウドネスおよびシャープネスの限界の補償を実行またはサポートする帯域拡張の仕組みに関する。

【背景技術】

【0002】

心理音響学の分野は、音の知覚の研究に関する。それには、人間の聴覚の仕組み、その生理的な応答、ならびに人間の神経系への音楽および音の生理的な影響が含まれる。とくには最新の通信システムの開発において、音刺激が聴覚系によってどのように処理されるのかについての知見が、新たなデジタルオーディオ技術の開発および既存の技術の改善において重要である。マルチメディアおよび配信サービスにおいて不可欠な構成要素であるオーディオコーデックは、人間の聴覚系の特性についての知見に左右され、低ビットレートでの効率的な伝送および保存のためにオーディオ情報を圧縮する。さらに、やはり心理音響に関する知見に大きく依存する客観的な品質評価の仕組みも、オーディオ品質の主観的評価を模擬するために開発されてきている。

【0003】

現代のほぼすべてのオーディオコーデック [1 ~ 5] は、オーディオ信号の信号周波数成分のうちの一部だけをエンコードおよび伝送し、デコーダにおいてオーディオ信号の残りの周波数を復元するという考え方を利用している。典型的には、信号の低周波数帯（LB）だけが伝送され、信号の高周波数帯（HB）は、後にいわゆる帯域拡張（BWE）によって復元される。典型的なBWEの仕組みにおいては、信号の周波数成分が、隣接する帯域からの利用可能な周波数成分（通常は、利用可能なLB）のトランスレートまたはフリッピングによって拡張される。しかしながら、そのようなやり方で復元された信号は、復元後の信号において知覚できる特定のアーチファクトゆえ、元のオーディオ信号のHBに正確に一致するHBを有するわけではない。それらのアーチファクトの影響を最小限にするために、BWEの仕組みにおいては、復元されたHBのゲインが典型的には元のHBのゲインよりも低く保たれ、結果として復元された信号の心理音響的な特性が変質することになる。最も影響を受ける特性は、とりわけラウドネスの知覚およびシャープネスの知覚である。ラウドネスは、スピーチ信号の信号強度または音圧に関する。シャープネスは、スピーチ信号の周波数におけるエネルギー分布に関し、高周波成分の相対的增加につれて向上する。信号の帯域が限られており、あるいは従来からのBWEの仕組みが適用される場合、復元された信号について知覚されるラウドネスおよびシャープネスがどちらも元の信号と比べて低下し、主観的品質の低下につながる。

10

20

30

40

50

【発明の概要】**【発明が解決しようとする課題】****【0004】**

したがって、受信/デコード後の信号について知覚されるラウドネスおよびシャープネスを向上させることができる方法および装置について、ニーズが存在する。

【0005】

本発明は、改善された帯域拡張の仕組みに関する。

【0006】

本発明の目的は、スピーチ信号の知覚品質を改善するための方法およびシステムを提供することにある。

【0007】

さらなる目的は、復元スピーチ信号について知覚されるラウドネスおよびシャープネスの改善を可能にすることにある。

【0008】

具体的な目的は、スピーチ信号を処理するためのエンコーダおよびデコーダ装置を提供することにある。

【0009】

他の具体的な目的は、スピーチ信号の処理方法を提供することにある。

【0010】

またさらなる具体的な目的は、フィルタ装置を提供することにある。

【課題を解決するための手段】**【0011】**

所定の帯域幅に限られた復元スピーチ信号について知覚されるラウドネスおよびシャープネスを改善する第1の態様において、スピーチ信号が用意される。次いで、スピーチ信号が、少なくとも前記所定の帯域幅のうちの第1の帯域部分にもとづく第1の信号部分と、前記所定の帯域幅のうちの第2の帯域部分にもとづく第2の信号部分とに分割される。その後、第1の信号部分が、前記第1の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調するように調整される。最後に、第2の信号部分が、少なくとも第1の信号部分にもとづいて復元され、調整済みの第1の信号部分と復元された第2の信号部分とが組み合わせられ、全体として知覚されるラウドネスおよびシャープネスが改善された復元スピーチ信号が生成される。

【0012】

本発明の第2の態様においては、所定の帯域幅に限られた復元スピーチ信号について知覚されるラウドネスおよびシャープネスを改善するためのシステムが、前記スピーチ信号を生成する手段を備える。さらに、前記スピーチ信号を、少なくとも前記所定の帯域幅のうちの第1の帯域部分にもとづく第1の信号部分と、前記所定の帯域幅のうちの第2の帯域部分にもとづく第2の信号部分とに分割する手段が、システムに備えられる。さらにシステムは、第1の信号部分を前記第1の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調するように調整する手段を備える。最後に、システムは、少なくとも第1の信号部分にもとづいて前記第2の信号部分を復元する手段と、前記調整した第1の信号部分と前記復元した第2の信号部分とを組み合わせ、全体として知覚されるラウドネスおよびシャープネスが改善された復元スピーチ信号を生成する手段とを備える。

【0013】

本発明の第3の態様においては、通信システムにおいて所定の帯域幅に限られたスピーチ信号を処理するためのエンコーダ装置が、前記スピーチ信号を生成する手段を備える。さらに、このエンコーダ装置が、前記スピーチ信号を、少なくとも前記所定の帯域幅のうちの第1の帯域部分にもとづく第1の信号部分と、前記所定の帯域幅のうちの第2の帯域部分にもとづく第2の信号部分とに分割する手段を備える。加えて、このエンコーダ装置が、前記第1の信号部分を、前記第1の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調する手段と、少なくとも前記調整した第1の信号部分を他のノードへ

10

20

30

40

50

と送信する手段とを備える。

【0014】

本発明の第4の態様においては、通信システムにおいて所定の帯域幅に限られたスピーチ信号を処理するためのデコーダ装置が、スピーチ信号の調整済みの第1の信号部分を受信する手段を備える。調整済みの第1の信号部分は、生成されたスピーチ信号を少なくとも所定の帯域幅のうちの第1の帯域部分にもとづく第1の信号部分と所定の帯域幅のうちの第2の帯域部分にもとづく第2の信号部分とに分割して、前記第1の信号部分を前記第1の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調するように調整することによって得られる。さらに、デコーダ装置は、少なくとも前記受信される調整済みの第1の信号部分にもとづいて前記第2の信号部分を復元する手段を備える。最後に、デコーダ装置は、前記受信される調整済みの第1の信号部分と前記復元される第2の信号部分とを組み合わせ、全体として知覚されるラウドネスおよびシャープネスが改善された復元スピーチ信号を生成する手段を備える。

10

【0015】

本発明の第5の態様においては、通信システムにおいて所定の帯域幅に限られたスピーチ信号を処理するためのデコーダ装置が、スピーチ信号の第1の信号部分を受信する手段を備える。この第1の信号部分は、生成されたスピーチ信号を少なくとも所定の帯域幅のうちの第1の帯域部分にもとづく第1の信号部分と所定の帯域幅のうちの第2の帯域部分にもとづく第2の信号部分とに分割することによって得られる。さらに、このデコーダ装置は、受信される第1の信号部分を、前記第1の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調するように構成された手段を備える。最後に、デコーダ装置は、少なくとも前記第1の信号部分にもとづいて前記第2の信号部分を復元する手段と、前記調整済みの第1の信号部分と前記復元される第2の信号部分とを組み合わせ、全体として知覚されるラウドネスおよびシャープネスが改善された復元スピーチ信号を生成する手段とを備える。

20

【0016】

本発明の第6の態様においては、通信システムの一ノードのエンコーダ装置において所定の帯域幅に限られたスピーチ信号を処理する方法が、前記スピーチ信号を生成するステップと、前記スピーチ信号を、少なくとも前記所定の帯域幅のうちの第1の帯域部分にもとづく第1の信号部分と、前記所定の帯域幅のうちの第2の帯域部分にもとづく第2の信号部分とに分割するステップとを含む。さらにこの方法は、前記第1の信号部分を、前記第1の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調するように調整するステップと、前記調整した第1の信号部分を他のノードへと送信するステップとを含む。

30

【0017】

本発明の第7の態様においては、通信システムの一ノードのデコーダ装置において所定の帯域幅に限られたスピーチ信号を処理する方法が、調整済みの第1の信号部分を別のノードから受信するステップを含む。この調整済みの第1の信号部分は、生成されたスピーチ信号を少なくとも前記所定の帯域幅のうちの第1の帯域部分にもとづく第1の信号部分と前記所定の帯域幅のうちの第2の帯域部分にもとづく第2の信号部分とに分割して、前記第1の信号部分を前記第1の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調するように調整することによって得られる。さらにこの方法は、前記受信される調整済みの第1の信号部分にもとづいて前記第2の信号部分を復元するステップと、前記調整済みの第1の信号部分と前記復元される第2の信号部分とを組み合わせ、全体として知覚されるラウドネスおよびシャープネスが改善された復元スピーチ信号を生成するステップとを含む。

40

【0018】

本発明の第8の態様においては、通信システムの一ノードのデコーダ装置において所定の帯域幅に限られたスピーチ信号を処理する方法が、スピーチ信号の第1の信号部分を別のノードから受信するステップを含む。この第1の信号部分は、前記スピーチ信号を少な

50

くとも前記所定の帯域幅のうちの第1の帯域部分にもとづく第1の信号部分と前記所定の帯域幅のうちの第2の帯域部分にもとづく第2の信号部分とに分割することによって得られる。さらにこの方法は、前記受信した第1の信号部分を前記第1の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調するように調整するステップと、少なくとも前記第1の信号部分にもとづいて前記第2の信号部分を復元するステップとを含む。最後に、この方法は、前記調整済みの第1の信号部分と前記復元された第2の信号部分とを組み合わせ、全体として知覚されるラウドネスおよびシャープネスが改善された復元スピーチ信号を生成するステップとを含む。

【0019】

本発明の第9の態様においては、通信システムにおいて所定の帯域幅に限られたスピーチ信号を調整するためのフィルタ装置が、スピーチ信号について該スピーチ信号の前記所定の周波数帯のうちの第1の帯域部分にもとづいて生成される第1の信号部分を、前記第1の帯域部分の範囲内の少なくとも所定の周波数区間を強調するように調整するように構成される。

10

【0020】

本発明の利点として、スピーチ信号の一部をプレフィルタ処理することによって、復元スピーチ信号について知覚される全体としてのラウドネスおよびシャープネスが改善されることが挙げられる。

【0021】

本発明ならびに本発明のさらなる目的および利点を、以下の説明を参照し、添付の図面と併せて検討することによって、最もよく理解することができる。

20

【図面の簡単な説明】

【0022】

【図1】本発明による方法の実施形態の概略のフロー図である。

【図2】本発明による方法のさらなる実施形態の概略のフロー図である。

【図3】図2の実施形態の働き概略のブロック図である。

【図4】本発明による方法のまたさらなる実施形態の概略のフロー図である。

【図5】図4の実施形態の働き概略のブロック図である。

【図6】本発明による装置の実施形態の概略のブロック図である。

【図7】外耳 - 中耳の応答を示すグラフである。

30

【図8】先行技術と本発明の効果との間の比較を示すグラフである。

【図9】先行技術と本発明の効果との間の比較の聴き取りテストを示す図である。

【図10】本発明による装置のさらなる実施形態の概略のブロック図である。

【図11】本発明の実施形態の概略のブロック図である。

【発明を実施するための形態】

【0023】

本開示は、通信システム（帯域拡張の仕組みを使用するシステムなど）におけるスピーチエンコーディング/デコーディング、ならびにそのようなシステムにおいて知覚される品質を改善し、とくには知覚されるラウドネスおよびシャープネスを改善するための方法および装置に関する。本発明の実施形態が有益であると考えられる具体的なコーデックの例は、AMR-WB（適応マルチレート広帯域）コーデックである。しかしながら、帯域拡張を使用する他のコーデックについても、本発明または本発明の実施形態が有益であると考えられる。

40

【0024】

本開示の目的は、信号（例えば、復元された信号）について知覚されるラウドネスおよびシャープネスを改善するためにスピーチ信号を調整するための方法および装置を提供することにある。信号の選択された一部分だけを調整し、あるいは前もってフィルタ処理することで、信号全体について知覚される品質を改善できることが、すでに認められている。人間の耳の自然な反応を考慮に入れることによって、耳が典型的に最も敏感である周波数について、スピーチ信号を強調することができる。結果として、再結合または復元した

50

スピーチ信号全体を、あたかもラウドネスおよびシャープネスが改善されたかのように聴き手に知覚させることができる。

【 0 0 2 5 】

図 1 を参照し、本発明の所定の帯域幅に限られた自然のスピーチ信号に相当するスピーチ信号について、知覚されるラウドネスおよびシャープネスを改善する方法の実施形態を説明する。この実施形態において、本発明による方法は、特定のノードまたはネットワークデバイスに限られない。

【 0 0 2 6 】

最初に、スピーチ信号が生成される (S 1 0)。スピーチ信号を、任意の従来からの手段によって生成することができる。次いで、スピーチ信号が、所定の帯域幅の第 1 および第 2 の帯域部分にもとづいて、少なくとも第 1 および第 2 の信号部分へと分離される (S 2 0)。典型的には、これは、所定の周波数帯を低周波数帯部分 (L B) および高周波数帯部分 (H B) へと分割することによって実行される。しかしながら、帯域幅について他の分割を実行することも可能である。本発明の特定の例においては、所定の周波数帯が、0 ~ 8 . 0 k H z の周波数区間に相当し、低周波数帯が 0 ~ 6 . 4 k H z の周波数によって表わされる一方で、高周波数帯は、6 . 4 ~ 8 . 0 k H z の周波数によって表わされる。しかしながら、他の周波数区間も同様に可能である。次いで、第 1 の信号部分が、第 1 の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調するように調整される (S 3 0)。特定の例では、この所定の周波数が、内耳の反応の中心周波数 (例えば、3 . 2 k H z) によって表わされ、あるいは 3 . 2 ~ 6 . 4 k H z の周波数範囲の全体によって表わされる。最後に、第 2 の信号部分またはその表現が、第 1 の信号部分にもとづいて復元され (S 4 0)、次いで調整された第 1 の信号部分および復元された第 2 の信号部分が組み合わせられ (S 5 0)、全体として知覚されるラウドネスおよびシャープネスが改善されてなる復元スピーチ信号が生成される。

【 0 0 2 7 】

例として、分離後のスピーチ信号の第 1 の部分の調整は、第 1 の信号部分のエネルギーの少なくとも一部が第 1 の帯域部分の範囲内の選択された周波数に向かって分配されると同時に、第 1 の信号部分のエネルギーの別の一部分が第 1 の帯域部分の高周波数区間または領域に向かって分配されるようなやり方で実行される。このやり方で、後に復元される信号について知覚される全体としてのラウドネスおよびシャープネスが、フィルタ処理または調整が行なわれていないスピーチ信号の低周波数帯にもとづいて復元されるスピーチ信号と比べて、改善される。

【 0 0 2 8 】

B W E の改善を、復元される信号の全体としてのラウドネスおよびシャープネスが B W E の仕組みに起因する損失について補償されるようなやり方で、利用可能なスピーチ信号の低周波数帯 (L B) をプレフィルタ処理することによって達成することができる。プレフィルタ処理は、復元された高周波数帯 (H B) については、これによって持ち込まれる信号アーチファクトの量が増加するため、典型的には実行されない。プレフィルタ処理という用語は、開示されるフィルタ処理または調整が信号の復元または再結合に先立って実行されることを指して使用されている。したがって、フィルタ処理または調整は、好ましくは信号の一部分についてのみ適用されるが、影響または改善は、再結合または復元後の信号全体について知覚される。

【 0 0 2 9 】

調整の段階 S 3 0 は、典型的には、低周波数帯のプレフィルタ処理にもとづき、復元の段階 S 4 0 は、B W E または低域通過フィルタ処理にもとづくことができる。

【 0 0 3 0 】

以下の説明においては、機能の各段階が、例えば通信システムまたはネットワークにおける送信および受信ノードのそれぞれのエンコーダおよびデコーダなど、ネットワークの 2 つのノードに分配または共有されるものとして説明される。したがって、分離または選択された第 1 の信号部分の調整 S 3 0 またはフィルタ処理のステップを、第 1 の信号部分

10

20

30

40

50

または第1の信号部分の表現の伝送後または伝送前に実行することができる（詳細は以下で説明される）。

【0031】

図2を参照し、スピーチ信号の第1の信号部分（例えば、低周波数帯）のフィルタ処理または調整が第1のネットワークノードのデコーダまたは受信装置において実行される方法の実施形態を説明する。したがって、手順全体の種々の段階のうちの一部が、エンコーダまたは送信装置において実行され、一部がデコーダまたは受信装置において実行される。この特定の実施形態においては、スピーチ信号が公知のやり方でエンコードされる。したがって、スピーチ信号を用意する段階S10ならびにスピーチ信号をスピーチ信号の所定の帯域幅の第1および第2の帯域部分にもとづいて少なくとも第1および第2の信号部分へと分離する段階S20が、好ましくはエンコーダにおいて実行される。次いで、分離または選択された第1の信号部分または第1の信号部分の表現が送信され（S24）、ネットワークの第2のノードの受信器またはデコーダ装置において受信される（S25）。次いで、デコーダが、受信した第1の信号部分または第1の信号部分の表現を、第1の帯域部分の範囲内の所定の周波数または周波数区間が強調されるように調整する（S30）。公知の手段により、スピーチ信号の第2の信号部分または高周波数帯が、受信した第1の信号部分にもとづいて復元される（S40）。最後に、調整済みの第1の信号部分および復元された第2の信号部分が組み合わせられ（S50）、全体として知覚されるラウドネスおよびシャープネスが改善された復元スピーチ信号が生成される。

【0032】

図3を参照すると、提供されるスピーチ信号の種々の部分ならびに上述の方法の実行中のスピーチ信号の種々の部分の処理が示されている。したがって、図3において、オーディオスピーチ処理のためのスピーチ信号が、信号供給部10によって適切な形態で提供される。次いで、信号は、低周波数帯LBおよび高周波数帯HBにもとづいて第1および第2の信号部分へと信号分離部20によって分割される。次いで、第1の信号部分LBが、送信部24によって送信される。その後、送信された第1の信号部分LBが、受信部25において受信される。受信された第1の信号部分LBにもとづき、第2の信号部分HBまたは第2の信号部分の表現が、復元部40によって（例えば、好ましくはBWEを使用して）復元される一方で、第1の信号部分が、調整部30によって調整またはフィルタ処理され、フィルタ処理済みまたは調整済みの第1の信号部分LB_fが生成される。最後に、2つの部分LB_fおよびHBが、結合部50によって再び組み合わせられ、改善された復元または再結合スピーチ信号が形成される。

【0033】

図4を参照し、スピーチ信号の第1の信号部分（例えば、低周波数帯）のフィルタ処理または調整がエンコーダまたは送信装置において実行される方法の実施形態を説明する。この実施形態においては、デコーダ装置も、後述される本発明の全利益を利用できるように構成される必要がある。

【0034】

したがって、エンコーダあるいは送信ノードまたは装置において、スピーチ信号を用意するステップS10ならびにスピーチ信号をスピーチ信号の所定の帯域幅の第1および第2の帯域部分にもとづいて少なくとも第1および第2の信号部分へと分離するステップS20が実行される。次いで、エンコーダ装置が、用意された第1の信号部分を、第1の帯域部分の範囲内の所定の周波数または周波数区間が強調されるように調整する（S30）。次いで、調整後の第1の信号部分または調整後の第1の信号部分の表現が送信され（S34）、例えば受信器またはデコーダ装置などのネットワークのノードにおいて受信される（S35）。加えて、エンコーダは、使用されているコーデックの種類についての情報や、デコーダにおいて少なくとも受信される調整済みの第1の信号部分（例えば、低周波数帯）にもとづいて第2の信号部分または高周波数帯を復元（S40）することが可能であるために必要な任意の他の情報を提供する。典型的には、この補助的情報は、2つのノードの間のセッションのネゴシエーションの際にすでに入手可能にされており、あるいは前

もって知られている（コーデックおよび他のセッションパラメータについて合意がなされている）。しかしながら、いくつかの事例においては、第2の信号部分の復元を助けるためにさらなる補助的情報を提供する必要がある。最後に、デコーダが、受信した調整済みの第1の信号部分 $L B_f$ および復元した第2の信号部分 $H B$ を組み合わせ（ $S 5 0$ ）、全体として知覚されるラウドネスおよびシャープネスが改善された復元スピーチ信号を生成することができる。これが、図5にさらに説明される。

【0035】

図5を参照すると、提供されるスピーチ信号の種々の部分ならびに上述の方法の実行中のスピーチ信号の種々の部分の処理が示されている。したがって、図5において、信号供給部10がスピーチ信号を供給し、次いでこの信号が、信号分離部20によって、低周波数帯 $L B$ および高周波数帯 $H B$ にもとづいて第1および第2の信号部分へと分割される。次いで、第1の信号部分 $L B$ が、アダプタ30によって調整またはフィルタ処理され、フィルタ処理済みまたは調整済みの第1の信号部分 $L B_f$ が生成される。次いでこれが、送信部34によって送信される。その後、送信された調整済みの第1の信号部分 $L B_f$ が、受信部35において受信される。第2の信号部分 $H B$ の復元を可能にする情報が、この信号とともに生成され、あるいはセッションの初期化またはコーデックのネゴシエーションの際にすでに生成されている。受信された調整済みの第1の信号部分 $L B_f$ にもとづき、第2の信号部分 $H B$ または第2の信号部分 $H B$ の表現が、復元部40によって（例えば、好ましくは $B W E$ または低域通過フィルタ処理を使用して）復元される。最後に、2つの部分 $L B_f$ および $H B$ が、結合部50によって組み合わせられ、改善された復元または結合スピーチ信号が形成される。

【0036】

図6を参照し、全体としての方法をサポートするシステム100ならびに装置（例えば、エンコーダ装置1/デコーダ装置2、送信部/受信部、第1/第2のノード）の実施形態を説明する。さらに、第1の信号部分の調整またはフィルタ処理の機能を、分離した機能として設けることができ、例えば破線の四角30によって示されるとおり、エンコーダ装置1またはデコーダ装置2のいずれか、あるいはシステム100の何らかの他のノードに備えることができるフィルタ装置30として設けることができる。

【0037】

図6を参照すると、本発明によるシステム100の実施形態は、所定の帯域に限られたスピーチ信号を生成するための信号供給部10を備える。この信号を、システムの他のノードから生成することができ、マイクロホンまたは他のオーディオデバイスによってエンコーダ装置1において実際に記録/生成することができ、あるいはシステムの何らかの他の装置において実際に記録/生成することができる。さらに、システム100は、スピーチ信号を所定の帯域の範囲内の2つの帯域部分にもとづいて少なくとも2つの信号部分へと分割するための分割部20を備える。典型的には、2つの信号部分は、信号の低周波数帯 $L B$ および高周波数帯 $H B$ に相当するが、何らかの他の分割も実行可能である。さらに、システム100は、第1の帯域部分の範囲内の少なくとも所定の周波数または周波数区間を強調すべく第1の信号部分または $L B$ をフィルタ処理または調整する調整部30を備える。最後に、システム100は、信号の第2の信号部分または $H B$ を復元するための復元部40と、調整済みの第1の信号部分と復元した第2の信号部分とを組み合わせ、スピーチ信号の復元（例えばラウドネスおよびシャープネスなどの知覚品質が改善されている）を生成する結合部50を備える。さらに、図6を参照すると、システム100が、例えばエンコーダ装置1を備える第1のノードおよびデコーダ装置2を備える第2のノードなど、後述される実施形態の通信システムにおける2つのノードを備えている。

【0038】

エンコーダ1の実施形態によれば、エンコーダ装置1は、スピーチ信号を生成するためのスピーチ信号供給部10と、スピーチ信号を第1および第2の信号部分に分割する信号分割部20とを備えている。さらに、エンコーダ装置1は、本開示においてすでに述べた方法に従って第1の信号部分を調整するための第1の信号部分の調整部30を備えている

10

20

30

40

50

。さらに、エンコーダ 1 は、少なくとも調整後の第 1 の信号部分の表現を送信し、さらに随意によりシステム 100 のデコーダ装置 2 における第 2 の信号部分の復元を補助する情報を送信する信号送信部 34 を備えている。

【0039】

デコーダ 2 の実施形態によれば、デコーダ装置 2 が、上述したエンコーダ装置 1 と協働するように構成される。したがって、デコーダ 2 は、上述のエンコーダ 1 によって生成される調整済みの第 1 の信号部分の表現を、追加の情報とともに受信する信号受信部 35 を備える。加えて、デコーダ 2 は、受信した調整済みの第 1 の信号部分にもとづいてスピーチ信号の第 2 の信号部分を復元する復元部 40 を備える。最後に、デコーダ 2 は、受信した調整済みの第 1 の信号部分と復元した第 2 の信号部分とを組み合わせ、知覚されるラウドネスおよびシャープネスが改善されている復元信号を生成する結合部 50 を備える。

10

【0040】

エンコーダ 1 のさらなる実施形態によれば、エンコーダ装置 1 が、スピーチ信号を生成するためのスピーチ信号供給部 10、スピーチ信号を第 1 および第 2 の信号部分へと分割する信号分割部 20、ならびに第 1 の信号部分または少なくとも第 1 の信号部分の表現を通信ネットワーク内の第 2 のノードへと送信するためのユニット 24 だけを備える。

【0041】

デコーダ 2 のさらなる実施形態によれば、デコーダ装置 2 が、上述のエンコーダ装置 1 から第 1 の信号部分を受信する信号受信部 25 を備える。さらに、デコーダ 2 は、受信した第 1 の信号部分を調整またはフィルタ処理するための第 1 の信号部分の調整部 30 と、受信した第 1 の信号部分にもとづいて第 2 の信号部分を復元するための復元部 40 と、調整後の第 1 の信号部分と復元した第 2 の信号部分とを組み合わせ、全体として知覚されるラウドネスおよびシャープネスが改善されている復元信号を生成する結合部 50 とを備える。

20

【0042】

以下で、第 1 の帯域部分の範囲内の所定の周波数または周波数区間の所望の強調を生成するために、第 1 の信号部分の調整またはフィルタ処理をどのように実行できるのかについて、いくつかの例を説明する。それらはいくつでも例にすぎず、知覚されるラウドネスおよびシャープネスについて全体として同じ影響を維持する傍ら、実際の数式を変更し、あるいは別のやり方で表現することが可能であることが当業者にとって明らかである。

30

【0043】

LB の中ほどの周波数（典型的には、特定の実施形態において約 3.2 kHz）の強調を、以下の種類のフィルタ

$$H(z) = \alpha \cdot z^{-2} + \beta \cdot z^{-1} + \gamma \cdot z^{+1} + \delta \cdot z^{+2} \quad (1)$$

によって達成でき、ここで好ましい係数は、 $\alpha = 0.1$ 、 $\beta = 0$ 、および $\delta = 0.85$ である。

【0044】

LB 信号の傾斜に影響を及ぼす別のフィルタの実施例は、

$$H(z) = \alpha \cdot z^{-1} + \beta \cdot z^{+1} \quad (2)$$

であって、好ましい係数は、 $\alpha = 0.06$ および $\beta = 0.66$ であり、あるいは

40

$$H(z) = 1 - \mu \cdot z^{-1} \quad (3)$$

であって、好ましい係数は、 $\mu = 0.2$ である。

【0045】

本発明の実施形態によれば、前置フィルタモジュールが、信号の HB が BWE の仕組みによって復元されており、あるいは低域通過フィルタ処理されている場合に、信号の LB 部分をプレフィルタ処理すべく作動させられる。この文脈において、プレフィルタ処理という用語は、フィルタ処理がスピーチ信号の復元に先立って実行されることを指す。結果として、信号の一部だけがフィルタ処理されるが、フィルタ処理が、復元される信号全体の知覚品質に影響を有する。本発明の実施形態のプレフィルタ処理は、LB の中ほどの

50

周波数または高い周波数を強調することを目的とする。

【 0 0 4 6 】

すでに述べたように、0 ~ 6 . 4 k H z の周波数成分からなる典型的な L B と、6 . 4 ~ 8 k H z の周波数成分からなる復元 H B とを考える。この筋書きにおいて、プレフィルタ処理は、3 . 2 k H z に中心を有する周波数または 3 . 2 ~ 6 . 4 k H z の全範囲を強調する。強調周波数は、典型的には、通常の聴き取りテストの被験者の外耳 - 中耳の反応に関して決定される（図 7 を参照されたい）。しかしながら、強調の周波数または周波数範囲の選択について、他の基準も適用可能である。例えば、調整を、顧客（耳が不自由であり、あるいは不自由でない）の実際の聴き取りのデータにもとづいて行なうことができる。

10

【 0 0 4 7 】

本発明の効果の説明が、図 8 に示されている。この例において、実線が元のスピーチ信号を示している。点線は、従来からの B W E の仕組みの対象とされ、低域通過フィルタ処理された復元信号に相当する。最後に、破線は、本発明による復元信号に相当する。破線および点線のどちらの信号も、元の信号と比べて 6 k H z よりも上のエネルギーが少ない。それにもかかわらず、破線の信号は、3 ~ 4 k H z の領域における周波数の強調ゆえに、点線の信号と比べてラウドネスおよびシャープネスがより高いと知覚されるであろう。換言すると、高い周波数に多くのエネルギーを有するシャープネスおよびラウドネスを、H B の代わりに信号の L B を増幅することによって復元することができる。これにより、信号アーチファクトの発生を効果的に回避することができる。

20

【 0 0 4 8 】

上述のプレフィルタ処理がラウドネスおよびシャープネスの感覚または知覚にどのように影響する（すなわち、知覚品質を向上させる）のかを理解するために、それぞれの心理音響モデルを検討することが有益である。重要な帯域 k における特定ラウドネスを

$$\tilde{N}(k)$$

によって定義すると、ラウドネスおよびシャープネスを、以下のように定めることができる [6] 。

30

$$N = \sum_k \tilde{N}(k) , \quad (4)$$

$$S \propto \frac{\sum_k k \times f(k) \times \tilde{N}(k)}{\sum_k \tilde{N}(k)} . \quad (5)$$

40

【 0 0 4 9 】

合計が、信号の帯域幅の全体の重要な帯域に及び、関数 $f(k)$ は、低周波数帯において 1 に等しく、最後のいくつかの重要な周波数帯において増加する。特定ラウドネスが、

$$\tilde{N}(k) \propto (0.5 + 0.5 \times E(k) \times E^*(k))^{0.23} , \quad (6)$$

として定められ、ここで正規化係数 E^* は、静寂のしきい値の逆数または外耳 - 中耳の周

50

波数応答に関係できる（図7を参照）。励振Eを、信号波形を周波数ドメインへと変換し、周波数ピンを重要な周波数帯へとまとめることによって計算することができる。

【0050】

式(4)、(6)、および図7から、ラウドネスの感覚を、利用可能な信号エネルギーを3.2kHzの領域に向かって分配することで、たとえ全体としての信号強度を変化させることなく向上させることができると結論付けることができる。

【0051】

式(5)から、シャープネスの感覚を、エネルギーをLBの低い周波数から高い周波数に向かって分配する（より高い帯域が、 k および $f(k)$ の増加によって合計においてより大きな重みを有する）ことによって向上させることができると結論付けることができる。

10

【0052】

本発明の発明者は、定評のあるMUSHR Aの仕組み[7]に従って多数の聴き取りテストを行なった。その結果が、図9に示されている。白色の棒が、基準信号であり、灰色の棒が、本発明の結果であり、黒色の柱が、先行技術の結果である。この図から見られるように、本発明による信号の調整は、先行技術の方法と比べて基準信号により近い信号をもたらす、したがって先行技術と比べて改善された聴き取りの体感をもたらす。

【0053】

さらに、図10が、本発明によるエンコーダおよびデコーダの機能の例を示している。

【0054】

上述の各段階、機能、手順、および/またはブロックを、汎用の電子回路および特定用途向けの回路の両方を含むディスクリートな回路または集積回路の技術など、任意の従来からの技術を使用して、ハードウェアにて実現することができる。

20

【0055】

あるいは、上述の各段階、機能、手順、および/またはブロックの少なくとも一部を、マイクロプロセッサ、デジタル信号プロセッサ(DSP)、および/または任意の適切なプログラマブルな論理デバイス(フィールド・プログラマブル・ゲート・アレイ(FPGA)デバイスなど)などの適切な処理装置によって実行されるソフトウェアにて実現してもよい。

【0056】

さらに、ネットワークノードの汎用の処理能力を再使用することが可能であってもよいことが、理解されるべきである。これは、例えば既存のソフトウェアをプログラムし直すことによって行なうことができ、あるいは新たなソフトウェア構成要素を追加することによって行なうことができる。

30

【0057】

ソフトを、通常はコンピュータにとって読み取り可能な媒体上に保持されるコンピュータプログラム製品として実現することができる。したがって、ソフトウェアを、コンピュータのプロセッサによる実行のために、コンピュータの動作メモリへとロードすることができる。コンピュータ/プロセッサは、上述の各段階、機能、手順、および/またはブロックのみに専用である必要はなく、他のソフトタスクも実行可能であってよい。

【0058】

40

以下で、コンピュータによる実施の例を、図11を参照して説明する。コンピュータ200が、プロセッサ210、動作メモリ220、および入力/出力ユニット230を備えている。この特定の例においては、上述の各段階、機能、手順、および/またはブロックの少なくとも一部が、プロセッサ210による実行のために動作メモリ220へとロードされるソフトウェア225において実行される。プロセッサ210およびメモリ220は、通常のソフトウェアの実行を可能にするためにシステムバスを介して互いに接続されている。I/Oユニット230を、関連のデータ(入力パラメータおよび/または得られた出力パラメータなど)の入力および/または出力を可能にするために、I/Oバスを介してプロセッサ210および/またはメモリ220へと互いに接続することができる。

【0059】

50

上記提案のラウドネスおよびシャープネスの部分的な補償のための仕組みは、必要なビットレートおよび複雑さの制限を維持する傍ら、知覚品質を向上させる。上述の考え方は、近年のほぼすべてのオーディオコーデックまたはBWEの仕組みに適用可能である。フィルタ処理によって信号のLB部分の中または高周波数を強調することで、復元される全体の信号についてラウドネスおよびシャープネスの感覚が改善される。換言すると、信号の一部をフィルタ処理することによって、信号全体について知覚される品質の改善がもたらされる。

【先行技術文献】

【非特許文献】

【0060】

【非特許文献1】[1]3GPP TS 26.190、「Adaptive Multi-Rate-Wideband (AMR-WB) speech codec; Transcoding functions」、2008

【非特許文献2】[2]3GPP TS 26.290「Extended Adaptive Multi-Rate-Wideband (AMR-WB+) speech codec; Transcoding functions」、2005

【非特許文献3】[3]3GPP TS 26.404、「Enhanced AAC Plus encoder SBR part」、2007

【非特許文献4】[4]ITU-T Rec. G.729.1、「G.729-based embedded variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729」、2006

【非特許文献5】[5]ITU-T Rec. G.718、「Frame error robust narrowband and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s」、2008

【非特許文献6】[6]H. FastlおよびE. Zwicker、「Psychoacoustics: Facts and Models」、Chapter 8.7.1 and 9.2、Springer、2007

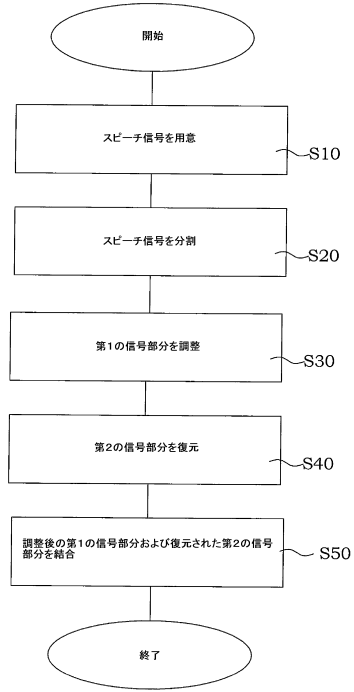
【非特許文献7】[7]G. StollおよびF. Kozamernik、「EBU listening tests on Internet audio codecs」、EBU Technical Review、June 2000。

10

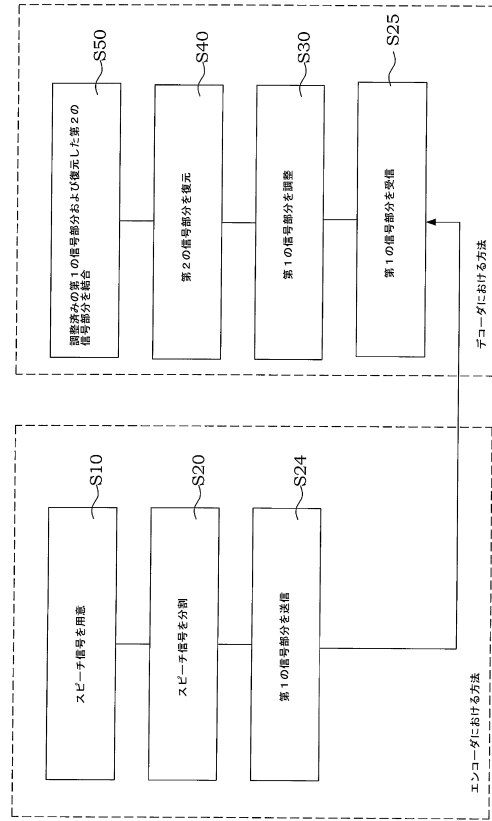
20

30

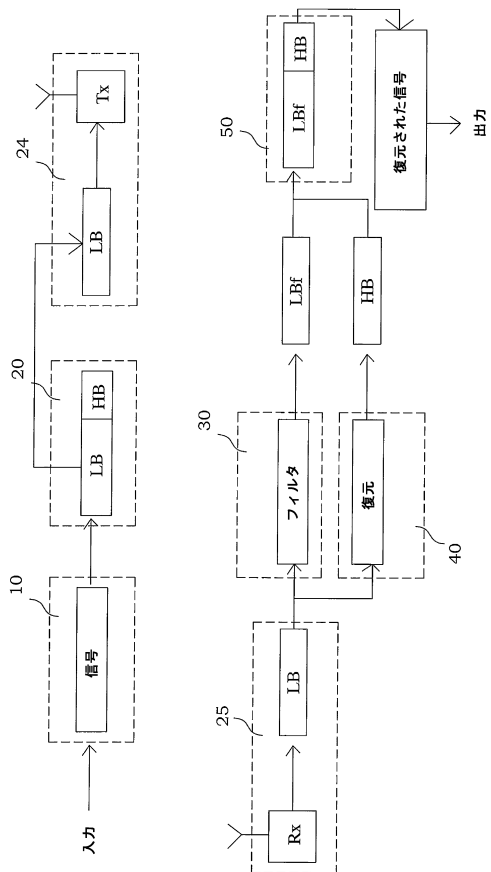
【図1】



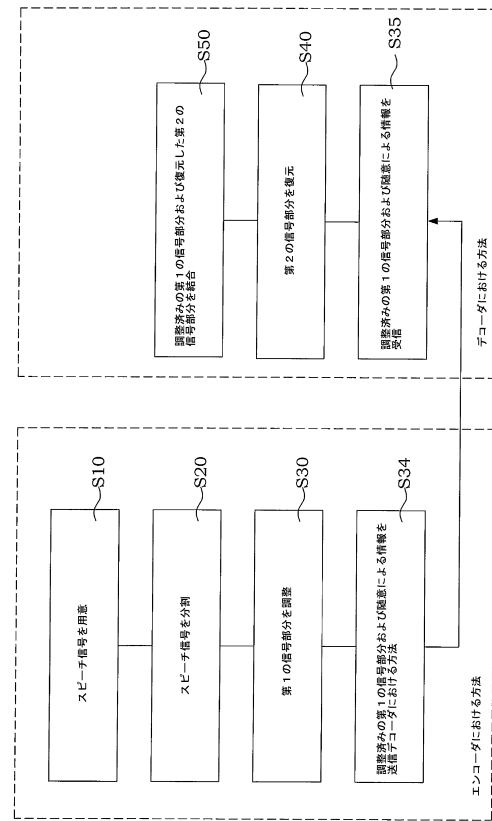
【図2】



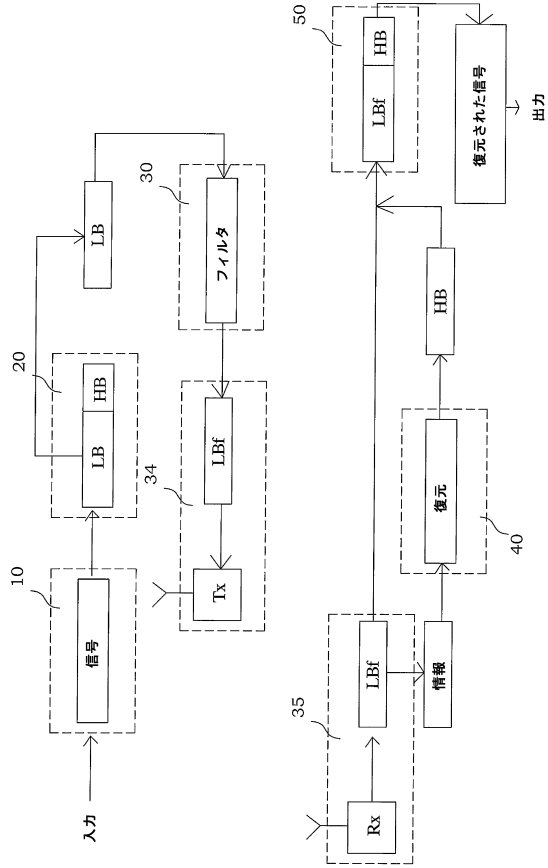
【図3】



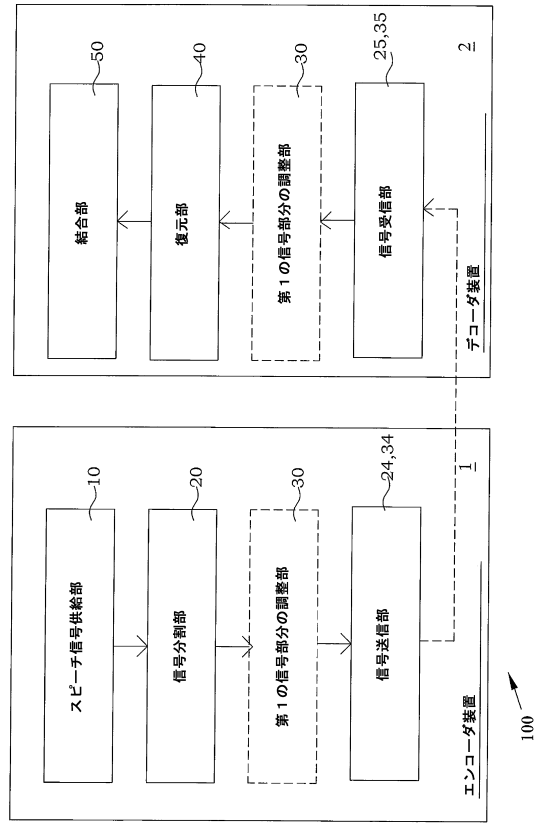
【図4】



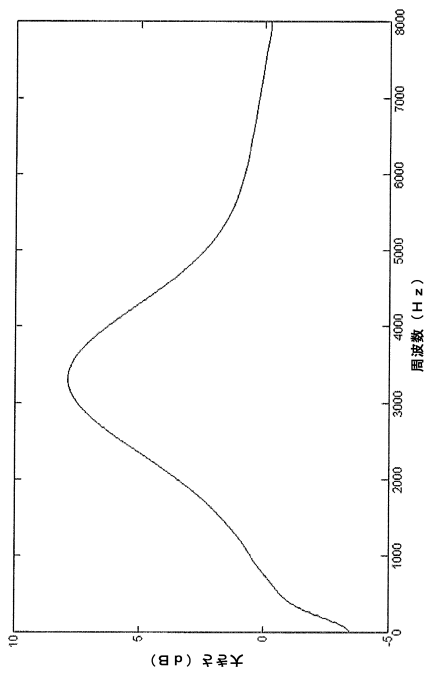
【図5】



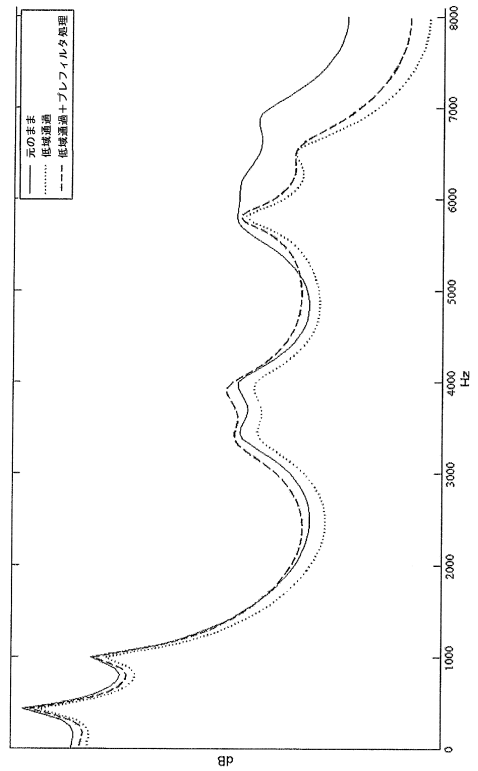
【図6】



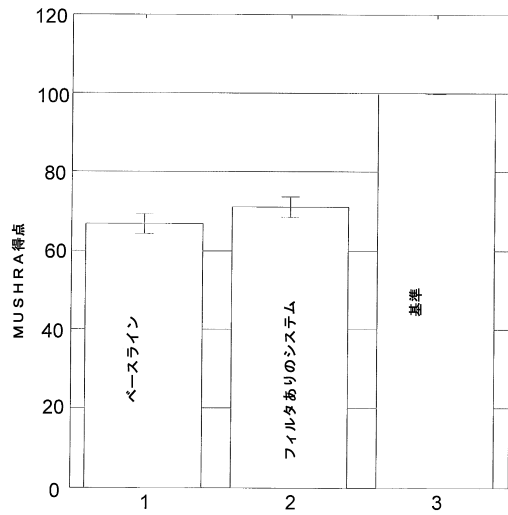
【図7】



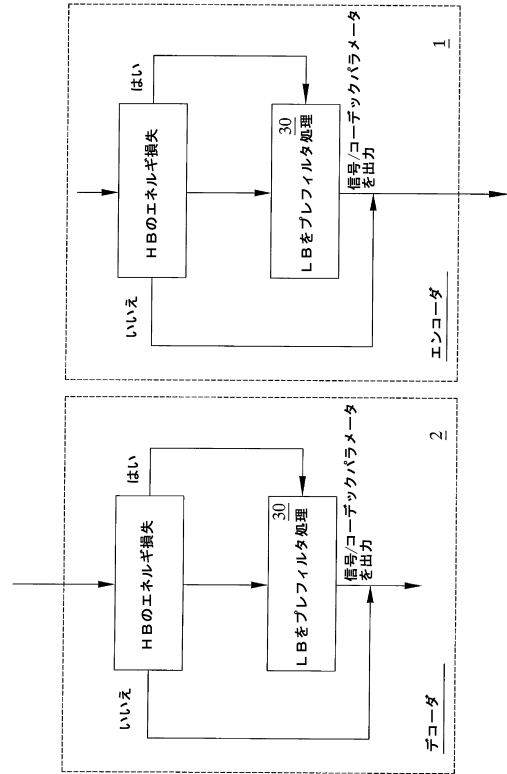
【図8】



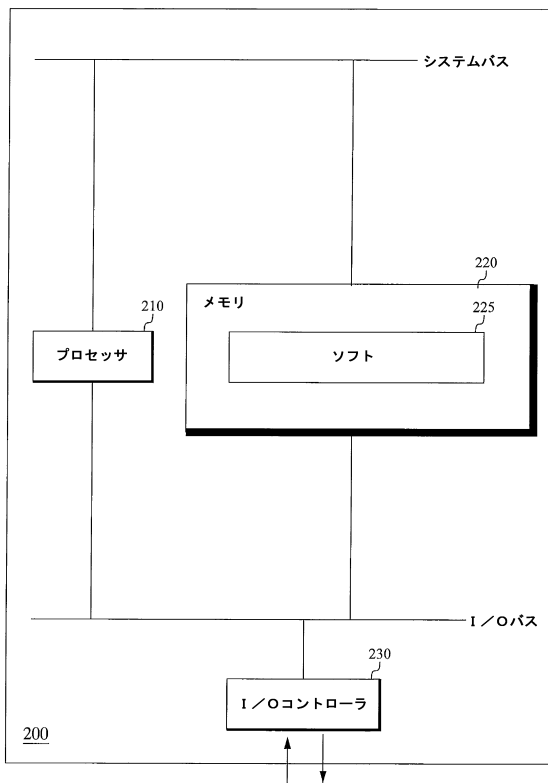
【図9】



【図10】



【図11】



フロントページの続き

(72)発明者 スヴェリルソン, シグルズール
スウェーデン国 19635 クングサンゲン, 2ティール, ハロンヴァーゲン 54

審査官 上田 雄

(56)参考文献 特開2008-107415(JP,A)
特開2010-066335(JP,A)
特開2005-010621(JP,A)
特開2007-164041(JP,A)
特開2007-178675(JP,A)

(58)調査した分野(Int.Cl., DB名)
G10L 19/00 - 19/26
G10L 21/00 - 21/18