

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2018-528480
(P2018-528480A)

(43) 公表日 平成30年9月27日(2018.9.27)

(51) Int.Cl.	F I	テーマコード (参考)
G 1 0 L 19/16 (2013.01)	G 1 0 L 19/16 2 0 0 A	
G 1 0 L 19/06 (2013.01)	G 1 0 L 19/06 B	
G 1 0 L 21/0216 (2013.01)	G 1 0 L 21/0216	

審査請求 有 予備審査請求 未請求 (全 31 頁)

(21) 出願番号 特願2018-515646 (P2018-515646)
 (86) (22) 出願日 平成28年9月23日 (2016.9.23)
 (85) 翻訳文提出日 平成30年4月16日 (2018.4.16)
 (86) 国際出願番号 PCT/EP2016/072701
 (87) 国際公開番号 W02017/050972
 (87) 国際公開日 平成29年3月30日 (2017.3.30)
 (31) 優先権主張番号 15186901.3
 (32) 優先日 平成27年9月25日 (2015.9.25)
 (33) 優先権主張国 欧州特許庁 (EP)
 (31) 優先権主張番号 16175469.2
 (32) 優先日 平成28年6月21日 (2016.6.21)
 (33) 優先権主張国 欧州特許庁 (EP)

(71) 出願人 500341779
 フラウンホーファー・ゲゼルシャフト・ツ
 ール・フェルデルング・デル・アンゲヴァ
 ンテン・フォルシュング・アインゲトラ
 ー
 ゲネル・フェライン
 ドイツ連邦共和国, 80686 ミュンヘ
 ン, ハンザシュトラッセ 27ツェー
 (74) 代理人 100085497
 弁理士 筒井 秀隆
 (72) 発明者 フィッシャー, ヨハネス
 ドイツ連邦共和国 91330 バンメル
 スドルフ (エッゴルシェイム) ユーラケ
 レ 18

最終頁に続く

(54) 【発明の名称】 線形予測符号化を使用して低減された背景ノイズを有するオーディオ信号を符号化する符号器および方法

(57) 【要約】

線形予測符号化を使用して低減された背景ノイズを有するオーディオ信号を符号化する符号器が示される。この符号器は、オーディオ信号の背景ノイズを推定する背景ノイズ推定部と、オーディオ信号の推定された背景ノイズをオーディオ信号から差し引くことによって背景ノイズ低減済みオーディオ信号を生成する背景ノイズ低減部と、オーディオ信号に対し線形予測分析を行って線形予測フィルタ(LPC)係数の第1セットを取得し、背景ノイズ低減済みオーディオ信号に対し線形予測分析を行って線形予測フィルタ(LPC)係数の第2セットを取得する予測部を含む。符号器は更に、取得されたLPC係数の第1セットと取得されたLPC係数の第2セットとによって制御される時間ドメインフィルタのカスケードからなる分析フィルタを含む。

【選択図】 図1

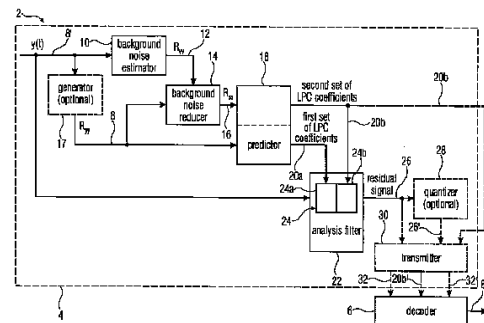


Fig. 1

【特許請求の範囲】**【請求項 1】**

線形予測符号化を使用して低減された背景ノイズを有するオーディオ信号(8')を符号化する符号器(4)であって、

オーディオ信号(8')の背景ノイズ表現(12)を推定するよう構成された背景ノイズ推定部(10)と、

オーディオ信号表現(8)から前記オーディオ信号(8')の推定された背景ノイズ表現(12)を差し引くことによって、背景ノイズ低減済みオーディオ信号表現(16)を生成するよう構成された背景ノイズ低減部(14)と、

前記オーディオ信号表現(8)に対し線形予測分析を行って線形予測フィルタ(LPC)係数の第1セット(20a)を取得し、前記背景ノイズ低減済みオーディオ信号表現(12)に対し線形予測分析を行って線形予測フィルタ(LPC)係数の第2セット(20b)を取得するよう構成された予測部(18)と、

取得されたLPC係数の第1セット(20a)と取得されたLPC係数の第2セット(20b)とによって制御される時間ドメインフィルタのカスケード(24, 24a, 24b)から構成され、前記オーディオ信号(8')から残差信号(26)を得る分析フィルタ(22)と、

を含む符号器(4)。

【請求項 2】

前記時間ドメインフィルタのカスケード(24)は、前記取得されたLPC係数の第1セット(20a)を使用する2回の線形予測フィルタ(24a)と、前記取得されたLPC係数の第2セット(20b)を使用する1回の他の線形予測フィルタの逆(24b)とを含む、請求項1に記載の符号器(4)。

【請求項 3】

前記時間ドメインフィルタのカスケード(24)はウィナーフィルタである、請求項1又は2に記載の符号器(4)。

【請求項 4】

前記背景ノイズ推定部(10)は、前記オーディオ信号(8)の背景ノイズ表現(12)として、前記背景ノイズの自己相関を推定するよう構成され、

前記背景ノイズ低減部(14)は、オーディオ信号の自己相関(8)から前記背景ノイズの自己相関(12)を差し引くことにより、前記背景ノイズ低減済みオーディオ信号表現(16)を生成するよう構成され、前記オーディオ信号の自己相関(8)は前記オーディオ信号表現であり、前記背景ノイズ低減済みオーディオ信号表現(16)は背景ノイズ低減済みオーディオ信号の自己相関である、請求項1乃至3のいずれかに記載の符号器(4)。

【請求項 5】

前記表現は自己相関である、請求項1乃至4のいずれかに記載の符号器(4)。

【請求項 6】

前記LPC係数の第2セット(20b)を伝送するよう構成されたトランスミッタ(30)をさらに含む、請求項1乃至5のいずれかに記載の符号器(4)。

【請求項 7】

前記残差信号(26)を伝送するよう構成されたトランスミッタをさらに含む、請求項1乃至6のいずれかに記載の符号器(4)。

【請求項 8】

前記残差信号(26)を伝送の前に量子化及び/又は符号化するよう構成された量子化部(28)をさらに含む、請求項1乃至7のいずれかに記載の符号器(4)。

【請求項 9】

前記LPC係数の第2セット(20b)を伝送の前に量子化及び/又は符号化するよう構成された量子化部(28)をさらに含む、請求項1乃至8のいずれかに記載の符号器(4)。

10

20

30

40

50

【請求項 10】

前記量子化部は、符号励振線形予測 (CELP)、エントロピー符号化、又は変換符号化励振 (TCX) を使用するよう構成された、請求項 8 又は 9 に記載の符号器。

【請求項 11】

請求項 1 乃至 10 のいずれかに記載の符号器 (4) と、符号化されたオーディオ信号を復号化するよう構成された復号器 (6) と、を含むシステム (2)。

【請求項 12】

線形予測符号化を使用して低減された背景ノイズを有するオーディオ信号を符号化する方法 (800) であって、

オーディオ信号の背景ノイズ表現を推定するステップ (S802) と、オーディオ信号表現から前記オーディオ信号の推定された背景ノイズ表現を差し引くことにより、背景ノイズ低減済みオーディオ信号表現を生成するステップ (S804) と、前記オーディオ信号表現に対し線形予測分析を行って線形予測フィルタ (LPC) 係数の第 1 セットを取得し、前記背景ノイズ低減済みオーディオ信号表現に対し線形予測分析を行って線形予測フィルタ (LPC) 係数の第 2 セットを取得するステップ (S806) と、

取得された LPC 係数の第 1 セットと取得された LPC 係数の第 2 セットとによって時間ドメインフィルタのカスケードを制御して、前記オーディオ信号から残差信号を得るステップ (S808) と、

を含む方法。

【請求項 13】

請求項 12 に記載の方法を実行するための、プログラムコードを有するコンピュータプログラム。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、線形予測符号化を使用して低減された背景ノイズを有するオーディオ信号を符号化する符号器と、対応する方法と、符号器および復号器を含むシステムと、に関する。換言すれば、本発明は、結合型のスピーチ強化及び/又は符号化の手法に関し、例えば C

【背景技術】**【0002】**

スピーチ及び通信の装置が広範に行き渡るようになり、悪条件下で使用される可能性が高まるにつれて、悪条件に対処し得るようなスピーチ強化方法への需要が高まってきた。その結果、例えばモバイルホンにおいて、スピーチ符号化などの全ての後続のスピーチ処理のための前処理ブロック/ステップとして、ノイズ減衰方法を使用することが今では普通に行われている。スピーチ強化をスピーチ符号器内へと組み込む手法は様々存在する (非特許文献 1, 2, 3, 4)。そのような設計は、伝送されるスピーチの品質を確かに向上させるが、カスケード処理に起因して、品質の一体的な知覚的最適化/最小化ができず、又は、量子化ノイズと干渉との一体的な最小化は少なくとも困難であった。

【0003】

スピーチコーデックの目標は、最少量の伝送データを用いて高品質のスピーチを伝送できるようにすることである。この目標に到達するため、信号の効率的な表現が必要となる。例えば、線形予測によるスピーチ信号のスペクトル包絡と、長期予測器による基本周波数と、ノイズ符号帳を用いた残差 (remainder) とのモデリングである。このような表現は、符号励振線形予測 (CELP) パラダイムを使用するスピーチコーデックの基礎であり、適応型マルチレート (AMR)、AMR - 広帯域 (AMR - WB)、統合型スピーチ及びオーディオ符号化 (USAC) 及び強化されたボイスサービス (EVS) などの主要な

10

20

30

40

50

スピーチ符号化標準において使用されている（非特許文献 5, 6, 7, 8, 9, 10, 11）。

【0004】

自然なスピーチ通信において、話者は手を使わないモードで装置を使用することが多い。そのようなシナリオでは、マイクロホンは通常口から離れた位置にあり、従って信号は反響または背景ノイズなどの干渉によって容易に歪みを受ける。そのような劣化は知覚されるスピーチ品質だけでなく、スピーチ信号の了解度にも影響を及ぼすので、会話の自然な流れの重大な障害となる。通信体験を向上させるためには、スピーチ強化方法を適用して、ノイズを減衰させ、反響の影響を低減することが有益である。スピーチ強化の分野は成熟しており、多くの方法が既に利用可能となっている（非特許文献 12）。しかしながら、現存するアルゴリズムの多くは、短時間フーリエ変換（STFT）のような変換などのオーバーラップ加算法に基づくものであり、それはオーバーラップ加算に基づく窓掛けスキームを適用する。これとは対照的に、CELPコーデックは、線形予測器/線形予測フィルタを用いて信号をモデル化し、残差だけに対して窓掛けを適用する。そのような基本的な相違により、強化と符号化の方法を合体させることが困難となる。とはいえ、強化と符号化との一体的な最適化は、潜在的に品質を向上させ、遅延と演算量とを低減させることが明白である。

10

【0005】

よって、改善された手法に対する必要性が生じる。

【先行技術文献】

20

【非特許文献】

【0006】

【非特許文献 1】M. Jeub and P. Vary, "Enhancement of reverberant speech using the CELP postfilter," in Proc. ICASSP, April 2009, pp. 3993-3996.

【非特許文献 2】M. Jeub, C. Herghlotz, C. Nelke, C. Beaugeant, and P. Vary, "Noise reduction for dual-microphone mobile phones exploiting power level differences," in Proc. ICASSP, March 2012, pp. 1693-1696.

【非特許文献 3】Martin, I. Wittke, and P. Jax, "Optimized estimation of spectral parameters for the coding of noisy speech," in Proc. ICASSP, vol. 3, 2000, pp. 1479-1482 vol.3.

30

【非特許文献 4】H. Taddei, C. Beaugeant, and M. de Meuleneire, "Noise reduction on speech codec parameters," in Proc. ICASSP, vol. 1, May 2004, pp. 1-497-500 vol.1.

【非特許文献 5】3GPP, "Mandatory speech CODEC speech processing functions; AMR speech Codec; General description," 3rd Generation Partnership Project (3GPP), TS 26.071, 12 2009. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/26071.htm>

【非特許文献 6】"Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Transcoding functions," 3rd Generation Partnership Project (3GPP), TS 26.190, 12 2009. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/26190.htm>

40

【非特許文献 7】B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukki la, J. Vainio, H. Mikkola, and K. Jarvinen, "The adaptive multirate wideband speech codec (AMR-WB)," IEEE Transactions on Speech and Audio Processing, vol. 10, no. 8, pp. 620-636, Nov 2002.

【非特許文献 8】ISO/IEC 23003-3:2012, "MPEG-D (MPEG audio technologies), Part 3: Unified speech and audio coding," 2012.

【非特許文献 9】M. Neuendorf, P. Gournay, M. Multrus, J. Lecomte, B. Bessette, R. Geiger, S. Bayer, G. Fuchs, J. Hilpert, N. Rettelbach, R. Salami, G. Schuller, R. Lefebvre, and B. Grill, "Unified speech and audio coding scheme for high qua

50

lity at low bitrates," in Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on, April 2009, pp. 1-4.

【非特許文献10】3GPP, "TS 26.445, EVS Codec Detailed Algorithmic Description; 3GPP Technical Specification (Release 12)," 3rd Generation Partnership Project (3GPP), TS 26.445, 12 2014. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/26445.htm>

【非特許文献11】M. Dietz, M. Multrus, V. Eksler, V. Malenovsky, E. Norvell, H. Pobloth, L. Miao, Z.Wang, L. Laaksonen, A. Vasilache, Y. Kamamoto, K. Kikuri, S. Ragot, J. Faure, H. Ehara, V. Rajendran, V. Atti, H. Sung, E. Oh, H. Yuan, and C. Zhu, "Overview of the EVS codec architecture," in Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on, April 2015, pp. 5698-5702. 10

【非特許文献12】J. Benesty, M. Sondhi, and Y. Huang, Springer Handbook of Speech Processing. Springer, 2008.

【非特許文献13】T. Baeckstroem, "Computationally efficient objective function for algebraic codebook optimization in ACELP," in Proc. Interspeech, Aug. 2013.

【非特許文献14】"Comparison of windowing in speech and audio coding," in Proc. WASPAA, New Paltz, USA, Oct. 2013.

【非特許文献15】J. Fischer and T. Baeckstroem, "Comparison of windowing schemes for speech coding," in Proc EUSIPCO, 2015. 20

【非特許文献16】M. Schroeder and B. Atal, "Code-excited linear prediction (CELP): High-quality speech at very low bit rates," in Proc. ICASSP. IEEE, 1985, pp. 937-940.

【非特許文献17】T. Baeckstroem and C. R. Helmrich, "Decorrelated innovative codebooks for ACELP using factorization of autocorrelation matrix," in Proc. Interspeech, 2014, pp. 2794-2798.

【非特許文献18】soundeffects.ch, "Civilisation soundscapes library," accessed: 23.09.2015. [Online]. Available: <https://www.soundeffects.ch/de/geraeusch-archiv/soundeffects.ch-produkte/civilisation-soundscapes-d.php>

【非特許文献19】Method for the subjective assessment of intermediate quality levels of coding systems, ITU-R Recommendation BS.1534, 2003. [Online]. Available: <http://www.itu.int/rec/R-REC-BS.1534/en>. 30

【非特許文献20】P. P. Vaidyanathan, "The theory of linear prediction," in Synthesis Lectures on Signal Processing, vol. 2, pp. 1{184. Morgan & Claypool publishers, 2007.

【非特許文献21】J. Allen, "Short-term spectral analysis, and modification by discrete Fourier transform," IEEE Trans. Acoust., Speech, Signal Process., vol. 25, pp. 235{238, 1977.

【発明の概要】

【発明が解決しようとする課題】 40

【0007】

本発明の目的は、線形予測符号化を使用してオーディオ信号を処理する、改善された概念を提供することである。この目的は独立請求項の主題により解決される。

【課題を解決するための手段】

【0008】

本発明の実施形態は、線形予測符号化を使用して低減された背景ノイズを有するオーディオ信号を符号化する符号器を示す。その符号器は、オーディオ信号の背景ノイズを推定するよう構成された背景ノイズ推定部と、オーディオ信号の推定された背景ノイズをオーディオ信号から差し引くことによって背景ノイズ低減済みのオーディオ信号を生成するよう構成された背景ノイズ低減部と、オーディオ信号に対し線形予測分析を行って線形予測フ 50

フィルタ(LPC)係数の第1セットを取得し、背景ノイズ低減済みのオーディオ信号に対し線形予測分析を行って線形予測フィルタ(LPC)係数の第2セットを取得するよう構成された予測部と、を含む。更に、符号器は、取得されたLPC係数の第1セットと取得されたLPC係数の第2セットとによって制御される時間ドメインフィルタのカスケードから構成される分析フィルタを含む。

【0009】

本発明は、線形予測符号化の環境における改善された分析フィルタにより、符号器の信号処理特性が強化されるという知見に基づいている。より具体的には、直列接続された時間ドメインフィルタのカスケード又は連鎖を使用することは、それらフィルタが線形予測符号化の環境の分析フィルタに適用された場合には、入力オーディオ信号の処理速度または処理時間を改善することになる。これが有利である理由は、ノイズに支配された周波数帯域をフィルタリングすることにより背景ノイズを低減するために入来時間ドメインオーディオ信号に対して典型的に使用される、時間-周波数変換とその逆の周波数-時間変換とを省略できるからである。換言すれば、背景ノイズの低減または除去を分析フィルタの一部として実行することで、背景ノイズの低減が時間ドメインにおいて実行され得る。従って、時間/周波数/時間の変換のために使用され得る、例えばMDCT/IDMCT(〔逆〕修正離散コサイン変換)のオーバーラップ加算処理が省略される。背景ノイズ低減は単一フレームに対して実行され得ず、連続フレームに対してのみ実行され得るので、このようなオーバーラップ加算方法は、符号器のリアルタイム処理特性を制限してしまう。

10

【0010】

換言すれば、上述の符号器は、背景ノイズ低減および従って分析フィルタの全体処理を単一オーディオフレームに対して実行することができ、それによりオーディオ信号のリアルタイム処理が可能になる。リアルタイム処理とは、参加しているユーザーにとって知覚可能な遅延無しにオーディオ信号を処理することを指す。知覚可能な遅延は、例えばテレビ会議においてオーディオ信号の処理の遅延に起因して一人のユーザーが他のユーザーの応答を待たねばならないときに発生し得る。この場合に許可される最大遅延は1秒未満であり、好ましくは0.75秒未満であり、より好ましくは0.25秒未満であり得る。注意すべきは、これらの処理時間は、送信者から受信者へのオーディオ信号の全体的な処理を指しているのであり、従って、符号器の信号処理に加えてオーディオ信号の伝送時間および対応する復号器内での信号処理をも含んでいる。

20

30

【0011】

実施形態によれば、時間ドメインフィルタのカスケード、つまり分析フィルタは、取得されたLPC係数の第1セットを使用する2回の線形予測フィルタと、取得されたLPC係数の第2セットを使用する1回の他の線形予測フィルタの逆とを含む。この信号処理はウィナーフィルタリング(Wiener filtering)と称されてもよい。従って、換言すれば、時間ドメインフィルタのカスケードはウィナーフィルタを含み得る。

【0012】

他の実施形態によれば、背景ノイズ推定部は、背景ノイズの自己相関(autocorrelation)をオーディオ信号の背景ノイズ表現として推定してもよい。更に、背景ノイズ低減部は、オーディオ信号の推定された自己相関から背景ノイズの自己相関を差し引くことにより、背景ノイズ低減済みオーディオ信号表現を生成してもよく、その場合、オーディオ信号の推定された自己相関はオーディオ信号表現であり、背景ノイズ低減済みオーディオ信号表現は、背景ノイズ低減済みのオーディオ信号の自己相関である。自己相関の推定を使用することは、LPC係数を計算するために時間ドメインオーディオ信号を使用することの代わりとして機能し、背景ノイズ低減を行うことは、時間ドメインにおける完全な信号処理を可能にする。従って、オーディオ信号の自己相関と背景ノイズの自己相関とは、1つのオーディオフレーム又はそのオーディオフレームの一部分の畳み込みにより、又は畳み込み積分(convolution integral)を使用することにより、計算することができる。このように、背景ノイズの自己相関は1つのフレーム内又は1つのサブフレーム内においてさえ実行されてもよく、そのフレーム又はサブフレームは、スピーチなどの前景オーディ

40

50

オ信号が（殆ど）何も存在しないようなフレーム又はフレームの一部として定義され得る。更に、背景ノイズ低減済みオーディオ信号の自己相関は、背景ノイズの自己相関と（背景ノイズを含む）オーディオ信号の自己相関とを差し引くことで計算され得る。背景ノイズ低減済みオーディオ信号と（典型的には背景ノイズを有する）オーディオ信号との自己相関を使用することにより、背景ノイズ低減済みオーディオ信号とオーディオ信号とのためのLPC係数をそれぞれ計算することが可能となる。背景ノイズ低減済みオーディオ信号のLPC係数はLPC係数の第2セットと称されてもよく、他方、オーディオ信号のLPC係数はLPC係数の第1セットと称されてもよい。従って、オーディオ信号は時間ドメインにおいて完全に処理され得る。なぜなら、時間ドメインフィルタのカスケードの適用は、オーディオ信号に対するそれらのフィルタリングを時間ドメインで実行するから

10

【0013】

添付図面を用いて実施形態を詳細に説明する前に、同一若しくは同等の機能を有する構成要素は、図中で同一の参照番号が付与されており、同一の参照番号が付与された構成要素の繰り返し説明が省略されていることを指摘しておく。従って、同一の参照番号を有する構成要素について記載された説明は相互に交換可能である。

【0014】

以下に、本発明の実施形態について添付図面を参照しながら説明する。

【図面の簡単な説明】

【0015】

【図1】オーディオ信号を符号化する符号器および復号器を含むシステムの概略ブロック図を示す。

20

【図2】(a)カスケード型の強化・符号化のスキーム、(b)CELPスピーチ符号化スキーム、及び(c)本発明の結合型強化・符号化スキームの概略ブロック図を示す。

【図3】異なる表記法を用いた、図2の実施形態の概略ブロック図を示す。

【図4】本提案の結合型手法(J)及びカスケード型方法(C)について式23で定義された知覚的大きさSNR(信号対ノイズ比)の概略線グラフを示し、ここで、入力信号が非定常な車両ノイズによって劣化しており、その結果が2つの異なるビットレート(下付き文字7で示された7.2kbit/s及び下付き文字13で示された13.2kbit/s)について表現されている。

30

【図5】本提案の結合型手法(J)及びカスケード接続された方法(C)について式23で定義された知覚的大きさSNRの概略線グラフを示し、ここで、入力信号が定常な白色ノイズによって劣化しており、その結果が2つの異なるビットレート(下付き文字7で示された7.2kbit/s及び下付き文字13で示された13.2kbit/s)について表現されている。

【図6】提案の結合型手法(JE)及びカスケード接続された強化(CE)に関し、2つの異なる干渉(白色ノイズ(W)及び車両ノイズ(C))についての異なる英語話者(女性(F)及び男性(M))に関するMUSHRAスコアを表す概略プロットを示し、ここで、全てのアイテムは2つの異なるビットレート(7.2kbit/s(7)と13.2kbit/s(13))で符号化されており、REFは隠れ基準であり、LPは3.5k

40

Hzのローパスアンカーを示し、Mixは歪みを受けた混合を示す。

【図7】2つの異なるビットレートに関してシミュレートされた異なるMUSHRAスコアのプロットを示し、新たな結合型強化(JE)とカスケード接続された手法(CE)との比較を示す。

【図8】線形予測符号化を使用して低減された背景ノイズを有するオーディオ信号を符号化する方法の概略フローチャートを示す。

【発明を実施するための形態】

【0016】

以下の説明において、本発明の実施形態をより詳細に説明する。各図の中で示され、同一又は類似の機能を有する構成要素には同じ参照番号が関連付けられている。

50

【 0 0 1 7 】

以下に、ウィナーフィルタリング（非特許文献 1 2）及び C E L P 符号化に基づく結合型の強化及び符号化の方法を説明する。このような融合は、（ 1 ）処理連鎖の中にウィナーフィルタリングを含むことで、C E L P コーデックのアルゴリズム上の低い遅延を増大させない、（ 2 ）これと同時に、結合型の最適化は、量子化及び背景ノイズに起因する歪みを最小化するという利点がある。更に、結合型スキームの演算量はカスケード手法の 1 つよりも低い。その実施は、C E L P 方式のコーデックにおける残差窓掛け(residual-windowing)についての最近の研究成果に基づいており（非特許文献 1 3 , 1 4 , 1 5）、その手法は C E L P コーデックのフィルタの中に新たな方法でウィナーフィルタリングを組み込むことを可能にする。この手法を使えば、カスケードシステムと比べて客観的品質および主観的品質の両方が改善されることを示し得る。

10

【 0 0 1 8 】

よって、スピーチの結合型強化および符号化のための本提案方法は、カスケード処理に起因するエラーの蓄積を防止し、更に知覚的な出力品質を向上させる。換言すれば、知覚ドメインにおける最適なウィナーフィルタリングにより干渉および量子化歪みの一体的な最小化が実現するので、本提案方法は、カスケード処理に起因するエラーの蓄積を防止する。

【 0 0 1 9 】

図 1 は、符号器 4 と復号器 6 とを含むシステム 2 の概略ブロック図を示す。符号器 4 は、線形予測符号化を使用して低減された背景ノイズを有するオーディオ信号 8 ' を符号化するよう構成されている。よって、符号器 4 は、オーディオ信号 8 ' の背景ノイズ表現 1 2 を推定するよう構成された推定部 1 0 を含んでもよい。符号器は更に、オーディオ信号表現 8 からオーディオ信号 8 ' の推定された背景ノイズ表現 1 2 を差し引くことにより、背景ノイズ低減済みオーディオ信号表現 1 6 を生成するよう構成された背景ノイズ低減部 1 4 を含んでもよい。このように、背景ノイズ低減部 1 4 は背景ノイズ推定部 1 0 から背景ノイズ表現 1 2 を受け取ってもよい。背景ノイズ低減部の更なる入力は、オーディオ信号 8 ' 又はオーディオ信号表現 8 であってもよい。任意選択的に、背景ノイズ低減部は、例えばオーディオ信号 8 ' の自己相関 8 のようなオーディオ信号表現 8 を内部的に生成するよう構成された生成部を含んでもよい。

20

【 0 0 2 0 】

更に、符号器 4 は、オーディオ信号表現 8 に対し線形予測分析を行って線形予測フィルタ（L P C）係数の第 1 セット 2 0 a を取得し、かつ背景ノイズ低減済みオーディオ信号表現 1 6 に対し線形予測分析を行って線形予測フィルタ（L P C）係数の第 2 セット 2 0 b を取得するよう構成された、予測部 1 8 を含んでもよい。背景ノイズ低減部 1 4 と同様に、予測部 1 8 は、オーディオ信号 8 ' からオーディオ信号表現 8 を内部的に生成する生成部を含んでもよい。しかしながら、共通の又は中央の生成部 1 7 を使用して、オーディオ信号 8 ' の表現 8 を一度に計算し、例えばオーディオ信号 8 ' の自己相関などのオーディオ信号表現を背景ノイズ低減部 1 4 及び予測部 1 8 へと提供することは有利であり得る。このように、予測部は、オーディオ信号表現 8 及び背景ノイズ低減済みオーディオ信号表現 1 6、例えばオーディオ信号の自己相関および背景ノイズ低減済みオーディオ信号の自己相関をそれぞれ受け取り、入来信号に基づいて L P C 係数の第 1 セット及び L P C 係数の第 2 セットをそれぞれ決定してもよい。

30

40

【 0 0 2 1 】

換言すれば、L P C 係数の第 1 セットはオーディオ信号表現 8 から決定されてもよく、L P C 係数の第 2 セットは背景ノイズ低減済みオーディオ信号表現 1 6 から決定されてもよい。予測部は、それぞれの自己相関から L P C 係数の第 1 セットおよび第 2 セットを計算するためにレビンソン・ダービン（Levinson Durbin）アルゴリズムを実行してもよい。

【 0 0 2 2 】

更に、符号器は、時間ドメインフィルタ 2 4 a , 2 4 b のカスケード 2 4 からなる分析フィルタ 2 2 を含み、その時間ドメインフィルタ 2 4 a , 2 4 b は取得された L P C 係数の

50

第1セット20aと取得されたLPC係数の第2セット20bとによって制御される。分析フィルタは、オーディオ信号8'に対して時間ドメインフィルタのカスケードを適用して残差信号26を決定してもよく、その場合、第1の時間ドメインフィルタ24aのフィルタ係数はLPC係数の第1セットであり、第2の時間ドメインフィルタ24bのフィルタ係数はLPC係数の第2セットである。残差信号は、LPC係数の第1及び/又は第2セットを有する線形フィルタによって表現されない、オーディオ信号8'の信号成分を含んでもよい。

【0023】

実施形態によれば、残差信号は、残差信号及び/又はLPC係数の第2セット24bを伝送の前に量子化及び/又は符号化するように構成された、量子化部28へと提供されてもよい。量子化部は、例えば変換符号化励振(TCX)、符号励振線形予測(CELP)又は例えばエントロピー符号化などのロスレス符号化などを実行してもよい。

10

【0024】

更なる実施形態によれば、残差信号の符号化は、量子化部28における符号化の代替例として、トランスミッタ30において実行されてもよい。このように、トランスミッタは、例えば変換符号化励振(TCX)、符号励振線形予測(CELP)又は例えばエントロピー符号化などのロスレス符号化を実行して、残差信号を符号化する。更に、トランスミッタは、LPC係数の第2セットを伝送するように構成されてもよい。任意選択的な受信機は復号器6である。このように、トランスミッタ30は、残差信号26又は量子化済み残差信号26'を受信してもよい。一実施例によれば、少なくとも量子化済み残差信号が量子化部内で既に符号化されていない場合、トランスミッタは残差信号又は量子化済み残差信号を符号化してもよい。残差信号又は代替的に量子化済み残差信号を任意選択的に符号化した後、トランスミッタへと提供されたそれぞれの信号は、符号化済み残差信号32又は符号化されかつ量子化された残差信号32'として伝送される。更に、トランスミッタは、LPC係数の第2セット20b'を受け取り、任意選択的に例えば残差信号を符号化するために使用されたのと同じ符号化方法を用いてそれを符号化して、その符号化されたLPC係数の第2セット20b'を、LPC係数の第1セットを伝送せずに、例えば復号器6へと更に伝送してもよい。換言すれば、LPC係数の第1セット20aは伝送される必要がない。

20

【0025】

復号器6は、符号化済み残差信号32又は代替的に符号化された量子化済み残差信号32'と、この残差信号32又は32'の1つに加えて符号化されたLPC係数の第2セット20b'とを、更に受信してもよい。この復号器は受信信号を復号化して、その復号化済み残差信号26を合成フィルタへと提供してもよい。合成フィルタは、LPC係数の第2セットをフィルタ係数として持つ、線形予測FIR(有限インパルス応答)フィルタの逆であってもよい。換言すれば、LPC係数の第2セットを持つフィルタが復号器6の合成フィルタを形成するために逆転される。合成フィルタの出力であり、よって復号器の出力は、復号化済みオーディオ信号8''である。

30

【0026】

実施形態によれば、背景ノイズ推定部は、オーディオ信号の背景ノイズの自己相関12をオーディオ信号の背景ノイズ表現として推定してもよい。更に、背景ノイズ低減部は、オーディオ信号の自己相関8から背景ノイズの自己相関12を差し引くことにより、背景ノイズ低減済みオーディオ信号表現16を生成してもよく、その場合、オーディオ信号の推定自己相関8はオーディオ信号表現であり、また、背景ノイズ低減済みオーディオ信号表現16は背景ノイズ低減済みオーディオ信号の自己相関である。

40

【0027】

図2と図3とは、同じ実施形態に関するものであるが、異なる表記法を使用している。つまり、図2は、カスケード型および結合型の強化/符号化手法の説明であって、 W_N と W_C とはノイズの多い信号とクリーンな信号とのホワイトニングをそれぞれ表し、 W_N^{-1} と W_C^{-1} とはそれらに対応する逆を表す。他方、図3は、カスケード型および結合型の強化/符

50

号化手法の説明であって、 A_y と A_s とはノイズの多い信号とクリーンな信号とのホワイトニングフィルタをそれぞれ表し、 H_y と H_s とは再構築（又は合成）フィルタであってそれらに対応する逆を表す。

【0028】

図2aと図3aとは、カスケード型の強化および符号化を実行する信号処理連鎖の強化部分と符号化部分とを示す。強化部分34は周波数ドメインで作動してもよく、ブロック36a及び36bは、例えばMDCTを使用する時間-周波数変換および例えばIMDCTを使用する周波数-時間変換を実行してもよく、又は任意の他の適切な変換を使用して時間-周波数変換および周波数-時間変換を実行してもよい。フィルタ38と40とは、周波数変換されたオーディオ信号42の背景ノイズ低減を実行してもよい。ここで、背景ノイズの周波数部分は、オーディオ信号8'の周波数スペクトルに対するそれらの影響力を低減することでフィルタリングされてもよい。よって、周波数-時間変換部36bは、周波数ドメインから時間ドメインへの逆変換を実行してもよい。強化部分34において背景ノイズ低減が実行された後で、符号化部分35は低減された背景ノイズを有するオーディオ信号の符号化を実行してもよい。従って、分析フィルタ22'は適切なLPC係数を用いて残差信号26''を計算する。残差信号は量子化されて合成フィルタ44へと提供されてもよく、その合成フィルタは図2aと図3aの場合には分析フィルタ22'の逆である。合成フィルタ42は図2aと図3aの場合には分析フィルタ22'の逆であるため、残差信号26を決定するために使用されたLPC係数が復号器へと伝送されて、復号化済みオーディオ信号8''を決定する。

10

20

【0029】

図2bと図3bとは、背景ノイズ低減が事前に実行されない符号化ステージ35を示す。符号化ステージ35は図2aと図3aとに関して既に説明したので、繰り返しを避けるため更なる説明を省略する。

【0030】

図2cと図3cとは、結合型の強化・符号化の主要な概念に関係する。分析フィルタ22が、フィルタ A_y と H_s とを使用する時間ドメインフィルタのカスケードを含むことが示されている。より詳細には、時間ドメインフィルタのカスケードは、取得されたLPC係数の第1セット20aを使用する2回の線形予測フィルタ(A_y^2)と、取得されたLPC係数の第2セット20bを使用する1回の他の線形予測フィルタの逆(H_s)とを含む。フィルタのこの配列またはこのフィルタ構造は、ウィナーフィルタと呼んでもよい。しかしながら、1つの予測フィルタ H_s が分析フィルタ A_s を打ち消すことに注目すべきである。換言すれば、フィルタ A_y を2回適用し(A_y^2 で表される)、フィルタ H_s を2回適用し(H_s^2 で表される)、フィルタ A_s を1回適用してもよい。

30

【0031】

図1に関して既に説明したように、これらフィルタのためのLPC係数は、例えば自己相関を使用して決定された。自己相関は時間ドメインで実行され得るので、結合型の強化および符号化を実施するために時間-周波数変換を実行する必要がない。更に、この手法が有利な理由は、量子化の更なる処理連鎖と合成フィルタリングの伝送とが、図2a及び図3aに関して説明した符号化ステージ35と同じであることである。しかし、背景ノイズ低減済み信号に基づくLPCフィルタ係数が、適切な合成フィルタリングのために復号器に伝送されなければならないことに注意すべきである。しかしながら、別の実施形態によれば、LPC係数を伝送する代わりに、LPC係数を有する線形フィルタの更なる逆転を防止し、合成フィルタ42を導出するために、フィルタ24bの既に計算済みのフィルタ係数(フィルタ係数20bの逆転により表現される)が伝送されてもよい。なぜなら、この逆転は符号器内で既に実行されていたからである。換言すれば、フィルタ係数20bを伝送する代わりに、これら係数の逆行列が伝送され、それにより逆転を2回実行することを防止してもよい。更に、符号器側のフィルタ24bと合成フィルタ42とは、符号器と復号器とにおいてそれぞれ適用される同じフィルタであってよい。

40

【0032】

50

換言すれば、図 2 に関し、C E L P モデルに基づくスピーチコーデックは、あるスピーチ生成モデルに基づいており、そのモデルは、入力スピーチ信号 S_n の相関が係数

$$\mathbf{a} = [\alpha_0, \alpha_1, \dots, \alpha_M]^T$$

を有し、M がモデル次数である線形予測フィルタによってモデル化され得ると仮定する（非特許文献 16）。残差

$$r_n = a_n * S_n$$

は、スピーチ信号の線形予測フィルタにより予測できない部分であるが、次にベクトル量子化を用いて量子化される。

10

【 0 0 3 3 】

$$\mathbf{s}_k = [s_k, s_{k-1}, \dots, s_{k-M}]^T$$

が入力信号のベクトルであり、上付き文字 T が転置を表すとする。この場合、残差は次式で表される。

【 数 1 】

$$r_k = \mathbf{a}^T \mathbf{s}_k. \quad (1)$$

【 0 0 3 4 】

スピーチ信号ベクトル \mathbf{s}_k の自己相関行列 \mathbf{R}_{ss} が次式で与えられると、

20

【 数 2 】

$$\mathbf{R}_{ss} = E\{\mathbf{s}_k \mathbf{s}_k^T\} \quad (2)$$

次数 M の予測フィルタの推定は次式で与えられる（非特許文献 20）。

【 数 3 】

$$\mathbf{a} = \sigma_e^2 \mathbf{R}_{ss}^{-1} \mathbf{u} \quad (3)$$

ここで、

$$\mathbf{u} = [1, 0, 0, \dots, 0]^T$$

30

であり、スカラー予測誤差 σ_e^2 は $\sigma_e^2 = 1$ となるように選択される。線形予測フィルタ \hat{S}_n がホワイトニングフィルタであることに着目すると、 r_k は無相関のホワイトノイズである。更に、オリジナル信号 S_n は、予測器 \hat{S}_n を用いた I I R フィルタリングを介して、残差 r_n から再構築され得る。次のステップは、知覚的歪みが最小化されるように、ベクトル量子化器を用いて、残差のベクトル

$$\mathbf{r}_k = [r_{kN}, r_{kN-1}, \dots, r_{kN-N+1}]^T$$

を

$$\tilde{\mathbf{r}}_k,$$

40

へ量子化することである。出力信号のベクトルが

$$\mathbf{s}'_k = [s_{kN}, s_{kN-1}, \dots, s_{kN-N+1}]^T$$

であり、

$$\tilde{\mathbf{s}}'_k$$

がその量子化された同等物であり、W が出力に知覚的重み付けを適用する畳み込み行列であるとする。その場合、知覚的最適化の問題は次式で表すことができる。

【数 4】

$$\min_{\mathbf{r}_k} \|\mathbf{W}(\mathbf{s}'_k - \tilde{\mathbf{s}}'_k)\|^2 = \min_{\mathbf{r}_k} \|\mathbf{WH}(\mathbf{r}_k - \tilde{\mathbf{r}}_k)\|^2 \quad (4)$$

ここで、H は予測器 \hat{s}_n のインパルス応答に対応する畳み込み行列である。

【0035】

C E L P 方式のスピーチ符号化の処理を図 2 b に示す。入力信号はまずフィルタ

$$A(z) = \sum_{m=0}^M \alpha_m z^{-m}$$

を用いてホワイトニングされ、残差信号を得る。残差のベクトルは、次にブロック Q において量子化される。最後に、スペクトル包絡構造が I I R フィルタリング $A^{-1}(z)$ によって再構築され、量子化された出力信号

10

$$\tilde{\mathbf{s}}_k$$

を得る。再合成された信号は知覚ドメインにおいて評価されるので、この手法は合成による分析の方法として知られている。

【0036】

ウィナーフィルタリング

単一チャンネルのスピーチ強化において、所望のクリーンなスピーチ信号 s_n と何らかの所望でない干渉 v_n との加算的混合である信号 y_n が得られたと仮定すると、次式が成り立つ。

20

【数 5】

$$y_n = s_n + v_n \quad (5)$$

【0037】

強化処理の目標はクリーンなスピーチ信号 s_n を推定することであるが、他方、ノイズの多い信号 y_n と相関行列の推定とにのみアクセス可能である。即ち、

【数 6】

$$\mathbf{R}_{ss} = E\{s_k s_k^T\} \quad \text{及び} \quad \mathbf{R}_{yy} = E\{y_k y_k^T\} \quad (6)$$

【0038】

ここで、

$$\mathbf{y}_k = [y_k, y_{k-1}, \dots, y_{k-M}]^T$$

である。フィルタ行列 H を使用すると、クリーンなスピーチ信号の推定

$$\hat{s}_n$$

は次式により定義される。

【数 7】

$$\hat{\mathbf{s}}_k = \mathbf{H} \mathbf{y}_k \quad (7)$$

40

【0039】

ウィナーフィルタとして知られる、最小二乗平均誤差 (M M S E) の意味における最適なフィルタは、非特許文献 1 2 のように導出され得る。

【数 8】

$$\mathbf{H} = \mathbf{R}_{ss} \mathbf{R}_{yy}^{-1} \quad (8)$$

【0040】

通常、ウィナーフィルタリングは、入力信号のオーバーラップしている窓に対して適用され、オーバーラップ加算法 (非特許文献 2 1、1 2) を使用して再構築される。この手法は図 2 a の強化ブロックの中で示される。しかし、その手法は、窓同士のオーバーラッ

50

プの長さに応じてアルゴリズム上の遅延を増大させてしまう。そのような遅延を回避するため、目的は、ウィナーフィルタリングを線形予測に基づく方法と合体させることである。

【 0 0 4 1 】

そのような関係を得るため、推定スピーチ信号

$$\hat{s}_k$$

は式 1 へと代入され、次式が得られる。

【 数 9 】

$$\begin{aligned} r_k &= \mathbf{a}^T \hat{s}_k = \mathbf{a}^T \mathbf{H} \mathbf{y}_k = \sigma_e^2 \mathbf{u}^T \mathbf{R}_{ss}^{-1} \mathbf{R}_{ss} \mathbf{R}_{yy}^{-1} \mathbf{y}_k \\ &= \sigma_e^2 \mathbf{u}^T \mathbf{R}_{yy}^{-1} \mathbf{y}_k = \gamma \mathbf{a}'^T \mathbf{y}_k \end{aligned} \quad (9)$$

ここで、 γ はスケーリング係数であり、

【 数 1 0 】

$$\mathbf{a}' = \hat{\sigma}_e^2 \mathbf{R}_{yy}^{-1} \mathbf{u} \quad (10)$$

はノイズの多い信号 y_n の最適な予測子である。換言すれば、ノイズの多い信号を \mathbf{a}' でフィルタリングすることで、推定クリーン信号の（スケーリング済み）残差が得られる。スケーリングは、クリーンな信号とノイズの多い信号、つまり

$$\sigma_e^2 \text{ と } \hat{\sigma}_e^2$$

のそれぞれの予想される残差誤差間の比、即ち

$$\gamma = \sigma_e^2 / \hat{\sigma}_e^2$$

である。この導出は、ウィナーフィルタリングと線形予測とが密接に関係した方法であることを示しており、以下では、この関係が結合型の強化および符号化方法を発展させるために用いられるであろう。

【 0 0 4 2 】

ウィナーフィルタリングの C E L P コーデックへの合体

ここでの目的は、ウィナーフィルタリングと C E L P コーデック（3 章および 2 章で説明した）とを 1 つの結合型アルゴリズムへと合体させることである。これらアルゴリズムを合体することで、ウィナーフィルタリングの通常の実施で必要とされるオーバーラップ加算窓掛けの遅延が回避され得、演算量が低減される。

【 0 0 4 3 】

結合型構造の実施は複雑ではない。強化されたスピーチ信号の残差は式 9 によって取得され得ると示される。よって、強化されたスピーチ信号は、クリーン信号の線形予測モデル \hat{s}_k を用いて残差を I I R フィルタリングすることにより、再構築され得る。

【 0 0 4 4 】

残差の量子化のために、クリーン信号 s_k' をその推定信号

$$\tilde{s}_k'$$

で置き換えることにより、式 4 が修正され得る。その結果、次式が得られる。

【 数 1 1 】

$$\min_{\mathbf{r}_k} \|\mathbf{W}(\tilde{s}_k' - \tilde{s}_k)\|^2 = \min_{\mathbf{r}_k} \|\mathbf{W}\mathbf{H}(\mathbf{r}_k - \tilde{\mathbf{r}}_k)\|^2 \quad (11)$$

【 0 0 4 5 】

換言すると、強化された目標信号

$$\tilde{s}_k'$$

10

20

30

40

50

を用いる目的関数は、クリーンな入力信号 s'_k とのアクセスを有する場合と同じである。

【0046】

結論として、標準 CELP に対する修正は、クリーンな信号の分析フィルタ a を、ノイズの多い信号の分析フィルタ a' と置き換えるだけである。CELP アルゴリズムの残りの部分は変更なしのままである。本提案の手法は図 2 (c) に示される。

【0047】

明白なことであるが、本提案方法は、ノイズ減衰が望まれる場合はいつでも、かつクリーンなスピーチ信号の自己相関 R_{ss} の推定にアクセスを有する場合には、任意の CELP コーデックにおいて、最小の変更で適用可能である。クリーンなスピーチ信号の自己相関の推定が利用不可の場合には、ノイズ信号の自己相関 R_w の推定を使用して、次式

$$R_{ss} \approx R_{yy} - R_{ww}$$

又は他の通常の推定により推測され得る。

【0048】

本提案方法は、時間ドメインフィルタを使用してクリーン信号の推定を取得できる限り、ビーム形成を伴う多チャンネルアルゴリズムなどのシナリオに対しても容易に拡張され得る。

【0049】

本提案方法の演算量における利点は、以下のように特徴付けられる。従来手法では、式 8 で与えられた行列フィルタ H を決定する必要があることに注意されたい。必要となる行列反転の複雑さは

$$O(M^3)$$

である。しかし、提案手法では、ノイズの多い信号について式 3 だけが解ければよく、これはレビンソン・ダービン・アルゴリズム (又は同類) を用いて実施されることができ、その複雑さは

$$O(N^2)$$

である。

【0050】

符号励振線形予測

図 3 に関して言い換えると、CELP パラダイムに基づくスピーチコーデックはスピーチ生成モデルを利用しており、そのモデルでは以下のことを仮定している。即ち、相関、従って入力スピーチ信号 S_n のスペクトル包絡は、係数

$$\mathbf{a} = [\alpha_0, \alpha_1, \dots, \alpha_M]^T$$

を有する線形予測フィルタによりモデル化され得ると仮定しており、ここで、 M は根底となるチューブモデル (非特許文献 16) により決定されるモデル次数である。線形予測フィルタ (予測子 18 とも称される) によって予測できないスピーチ信号の部分である、残差

$$r_n = a_n * S_n$$

は、次にベクトル量子化を用いて量子化される。

【0051】

入力信号 s の 1 つのフレームのための線形予測フィルタ a_s は、次式の最小化によって取得され得る。

【数 12】

$$\min_{\mathbf{a}_s} \mathcal{E} \{ \|\mathbf{s}^* \mathbf{a}_s\|^2 - 2\sigma_s^2 (\mathbf{u}^* \mathbf{a}_s - 1) \} \quad (12)$$

10

20

30

40

50

ここで、

$$\mathbf{u} = [1 \ 0 \ 0 \ \dots \ 0]^T$$

である。その解は次式の通りである。

【数 1 3】

$$\mathbf{a}_s = \sigma_e^2 \mathbf{R}_{ss}^{-1} \mathbf{u} \quad (13)$$

【 0 0 5 2 】

\mathbf{a}_s のフィルタ係数 から成る 畳み込み行列 \mathbf{A}_s の以下の定義

【数 1 4】

$$\mathbf{A}_s = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \alpha_1 & \ddots & & \vdots \\ \alpha_2 & \ddots & 1 & \ddots \\ \vdots & \ddots & \alpha_1 & 1 & 0 \\ \alpha_M & \dots & \alpha_2 & \alpha_1 & 1 \end{bmatrix} \quad (14)$$

10

を用いて、残差信号は入力スピーチフレームを畳み込み行列 \mathbf{A}_s で乗算することにより取得され得る。

20

【数 1 5】

$$\mathbf{e}_s = \mathbf{A}_s \cdot \mathbf{s} \quad (15)$$

【 0 0 5 3 】

ここで、入力信号からゼロ入力応答を差し引き、それを再合成の中に再導入することにより（非特許文献 1 5）、窓掛けが C E L P コーデックと同様に実行される。

【 0 0 5 4 】

式 1 5 の乗算は、予測フィルタを用いた入力信号の畳み込みと同じであり、従って、F I R フィルタリングと対応する。オリジナル信号は、次式のように再構築フィルタ \mathbf{H}_s を用いた乗算によって、残差から再構築され得る。

30

【数 1 6】

$$\mathbf{s} = \mathbf{H}_s \cdot \mathbf{e}_s \quad (16)$$

ここで、 \mathbf{H}_s は予測フィルタのインパルス応答

$$\boldsymbol{\eta} = [1, \eta_1, \dots, \eta_{N-1}]$$

から構成される。

【数 1 7】

$$\mathbf{H}_s = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \eta_1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \eta_{N-1} & \dots & \eta_1 & 1 \\ \vdots & & \vdots & \vdots \end{bmatrix} \quad (17)$$

40

結果として、この演算は I I R フィルタリングに相当する。

【 0 0 5 5 】

残差ベクトルはベクトル量子化を適用して量子化される。従って、量子化されたベクトル

$\hat{\mathbf{e}}_s$

は、ノルム - 2 の意味において、知覚的距離を最小化するように、所望の再構築クリーン信号へと選択される。

【数 1 8】

$$\min_{\hat{\mathbf{e}}_s} \|\mathbf{WH}(\hat{\mathbf{e}}_s - \mathbf{e}_s)\|^2 \quad (18)$$

ここで、 \mathbf{e}_s は量子化されていない残差であり、

$$\mathbf{W}(z) = A(0.92z)$$

10

は、AMR-WB スピーチコーデック（非特許文献 6）で使用されているような知覚的加重み付けフィルタである。

【0056】

CELP コーデックにおけるウィナーフィルタリングの適用

単一チャネルのスピーチ強化の適用に関し、獲得されたマイクロホン信号 y_n が所望のクリーンなスピーチ信号 S_n と何らかの所望でない干渉 V_n との加算的混合であると仮定すると、

$$y_n = S_n + V_n$$

となる。Zドメインにおいては、等価的に

20

$$Y(z) = S(z) + V(z)$$

が成り立つ。

【0057】

ウィナーフィルタ $B(z)$ を適用することで、フィルタリングによりノイズの多い観測 $Y(z)$ からスピーチ信号 $S(z)$ を再構築することが可能になり、それにより推定スピーチ信号が次式のようになる。

$$\hat{S}(z) := B(z)Y(z) \approx S(z)$$

ウィナーフィルタについての最小二乗平均解は非特許文献 12 に記載の通りであり、

30

【数 1 9】

$$B(z) = \frac{|S(z)|^2}{|S(z)|^2 + |V(z)|^2} \quad (19)$$

ここで、スピーチ信号 S_n 及びノイズ信号 V_n は、それぞれ非相関であると仮定されている。

【0058】

スピーチコーデックにおいて、ノイズの多い信号 y_n のパワースペクトルの推定が線形予測モデル

40

$$|A_y(z)|^{-2}$$

のインパルス応答の形態で利用可能である。換言すれば、

$$|S(z)|^2 + |V(z)|^2 \approx \gamma |A_y(z)|^{-2}$$

であり、ここで、 γ はスケーリング係数である。ノイズの多い線形予測子は、ノイズの多い信号の自己相関行列 R_{yy} から通常どおり計算され得る。

【0059】

更に、ノイズの多い線形予測子はクリーンなスピーチ信号

$$|S(z)|^2$$

のパワースペクトルから、又は等価的にクリーンなスピーチ信号の自己相関行列 R_{ss} から推定されてもよい。強化アルゴリズムは、ノイズ信号が定常であると仮定しており、よって R_{vv} で示すノイズ信号の自己相関は入力信号の非スピーチフレームから推定され得る。その場合、クリーンなスピーチ信号の自己相関行列 R_{ss} は、

$$\hat{R}_{ss} = R_{yy} - R_{vv}$$

で推定され得る。ここで、

$$\hat{R}_{ss}$$

が正定値 (positive definite) であることを確保するという通常の予防措置を講じておくことが有利である。

【 0 0 6 0 】

クリーンなスピーチの推定自己相関行列

$$\hat{R}_{ss}$$

を使用して、対応する線形予測子が決定され得る。ここで、Zドメインにおけるインパルス応答は

$$\hat{A}_s^{-1}(z)$$

である。よって、

$$|S(z)|^2 \approx |\hat{A}_s(z)|^{-2}$$

となり、式 19 は次式のように表すことができる。

【 数 2 0 】

$$B(z) \approx \frac{|\hat{A}_s(z)|^{-2}}{|A_y(z)|^{-2}} = \frac{|A_y(z)|^2}{|\hat{A}_s(z)|^2} \quad (20)$$

【 0 0 6 1 】

換言すれば、FIRモードとIIRモードとのそれぞれにおいて、ノイズの多い信号とクリーン信号との予測子を用いて2回フィルタリングすることにより、クリーン信号のウィナー推定が取得され得る。

【 0 0 6 2 】

畳み込み行列は、 A_s 及び A_y による予測子

$$\hat{A}_s(z) \text{ 及び } A_y(z)$$

を用いたFIRフィルタリングにそれぞれ対応すると言える。同様に、 H_s 及び H_y を予測フィルタリング (IIR) に対応するそれぞれの畳み込み行列であると仮定する。これらの行列を使用することで、従来のCELP符号化は図3bで示すフロー図を用いて表すことができる。ここで、 A_s を用いて入力信号 S_n をフィルタリングして残差を取得し、それを量子化し、 H_s を用いたフィルタリングによって量子化済み信号を再構築することが可能である。

【 0 0 6 3 】

強化と符号化とを結合する従来の手法を図3aに示す。ここでは、符号化の前の前処理ブロックとしてウィナーフィルタリングが適用される。

【 0 0 6 4 】

最後に、提案手法においては、ウィナーフィルタリングがCELP方式のスピーチコーデックと結合されている。図3aのカスケード手法と図3bの結合型手法とを比較した場合

10

20

30

40

50

、追加的なオーバーラップ加算 (OLA) の窓掛けスキームを省略できることが明らかである。更に、符号器における入力フィルタ A_s は H_s を取り消す。従って、図 3 c に示すように、フィルタ結合

$$A_y^2 H_s$$

を用いて劣化した入力信号 y をフィルタリングすることで、推定されたクリーン残差信号

$$\tilde{e} = A_y^2 H_s y$$

が得られる。従って、誤差の最小化は次式となる。

【数 2 1】

$$\min_{\hat{e}} \| \mathbf{W} \mathbf{H}_s (\hat{e} - \tilde{e}) \|^2 \quad (21)$$

10

【 0 0 6 5 】

このように、この手法はクリーンな推定と量子化済み信号との間の距離を一体的に最小化するので、知覚的ドメインにおける干渉と量子化ノイズとの一体的最小化が実行可能になる。

【 0 0 6 6 】

結合型のスピーチ符号化及び強化の手法の性能は、主観的手段と客観的手段との両方を用いて評価された。新たな方法の性能と切り離すため、単純な CELP コーデックが使用され、そこでは残差信号だけが量子化されたが、長期予測 (LTP) の遅延とゲイン、線形予測符号化 (LPC) 及びゲインファクタは量子化されなかった。残差はペア毎の反復方法を使用して量子化され、そこでは、非特許文献 17 に記載されるように、2 個のパルスが全ての位置に連続的に加えられた。更に、推定アルゴリズムの影響を避けるため、クリーンなスピーチ信号の相関行列 R_{ss} は、全てのシミュレートされたシナリオにおいて既知であると仮定された。スピーチ及びノイズ信号が非相関であると仮定して、

20

$$R_{ss} = R_{yy} - R_{vv}$$

が成り立つ。如何なる実用的なアプリケーションにおいても、ノイズ相関行列 R_{vv} 又は代替的にクリーンなスピーチの相関行列 R_{ss} は、獲得されたマイクホン信号から推定されなければならない。一般的な手法は、干渉が定常であると仮定して、スピーチの切れ目におけるノイズ相関行列を推定することである。

30

【 0 0 6 7 】

評価されたシナリオは、所望のクリーンなスピーチ信号と追加的な干渉との混合から成っていた。2 つのタイプの干渉が考慮された。即ち、定常な白色ノイズと、非特許文献 18 の都市化音響風景ライブラリ (Civilisation Soundscape Library) からの車両ノイズの録音のセグメントと、である。残差のベクトル量子化は、AMR-WB コーデック (非特許文献 6) の 7.2 kbit/s 及び 13.2 kbit/s の全体ビットレートにそれぞれ対応する、 2.8 kbit/s 及び 7.2 kbit/s のビットレートを用いて実行された。 12.8 kHz のサンプリングレートが全てのシミュレーションに使用された。

40

【 0 0 6 8 】

強化されかつ符号化された信号は、主観的手段および客観的手段の両方を用いて評価された。よって、式 23 及び式 22 に定義されるようにリスニングテストが行われ、知覚的な大きさの信号対ノイズ比 (SNR) が計算された。結合型の強化処理はフィルタの位相に対して影響を与えないので、この知覚的な大きさ SNR が使用された。合成フィルタ及び再構成フィルタは双方とも、予測フィルタの設計に起因して、最小位相フィルタの制限に拘束されているので、フィルタの位相に対して影響を与えない。

【 0 0 6 9 】

フーリエ変換を演算子

$$\mathcal{F}(\cdot)$$

50

として定義すると、知覚的ドメインにおける再構築されたクリーン基準および推定クリーン信号の絶対値のスペクトル値は次式で示される。

【数 2 2】

$$S = |\mathcal{F}(\mathbf{W}\mathbf{H}_s\mathbf{e}_k)| \quad \text{及び} \quad \hat{S} = |\mathcal{F}(\mathbf{W}\mathbf{H}_s\hat{\mathbf{e}}_k)| \quad (22)$$

【 0 0 7 0 】

修正された知覚的な信号対ノイズ比 (P S N R) の定義は次式となる。

【数 2 3】

$$\text{PSNR}_{\text{ABS}} = 10 \log_{10} \frac{\|S\|^2}{\|\hat{S} - S\|^2} \quad (23)$$

10

【 0 0 7 1 】

主観的評価のために、U S A C の標準 (非特許文献 8) に使用されたテストセットからスピーチ項目、即ち、上述のように白色ノイズおよび車両ノイズによって汚損された項目が使用された。S T A X (登録商標) の静電型ヘッドホンを防音環境において使用し、14 人の被験者について、非特許文献 19 の隠れ基準及びアンカー付き多刺激検法 (Multiple Stimuli with Hidden Reference and Anchor: M U S H R A) のリスニングテストを行った。そのリスニングテストの結果を図 6 に示し、図 7 には平均および 95 % の信頼区間 (confidence intervals) を示す差分 M U S H R A スコアを表す。

20

【 0 0 7 2 】

図 6 の絶対値の M U S H R A テスト結果は、隠れ基準が常に正確に 100 ポイントへと割り当てられたことを示す。オリジナルのノイズの多い混合は、各項目について最低の平均スコアを与えられたので、全ての強化方法が知覚的品質を向上させたことが分かる。低いビットレートについての平均スコアは、カスケード方式と比較して、全項目にわたる平均について 6.4 M U S H R A ポイントの統計的に有意な改善を示している。高いビットレートについては、全項目にわたる平均は改善が見られるが、しかし統計的に有意とは言えない。

【 0 0 7 3 】

結合型の方法と前強化の方法とのより詳細な比較を得るために、差分 M U S H R A スコアが図 7 に示されている。ここでは、前強化の方法と結合型の方法との間の差が各リスナーおよび項目について計算されている。差分結果は、低いビットレートについての統計的に有意な改善を示すことで、絶対値の M U S H R A スコアを証明しているが、他方、高いビットレートについての改善は統計的に有意でない。

30

【 0 0 7 4 】

換言すれば、結合型のスピーチ強化及び符号化のための方法が示され、その方法は全体的な干渉および量子化ノイズの最小化を可能にする。対照的に、従来の手法は、カスケード型の処理ステップにおいて強化および符号化を適用している。両方の処理ステップを一体化することは、繰り返しの窓掛け及びフィルタリング操作が省略できるため、演算量から見ても魅力的である。

40

【 0 0 7 5 】

C E L P タイプのスピーチコーデックは、非常に低い遅延となるように設計されており、従って、処理窓が将来の処理窓とオーバーラップすることを回避している。対照的に、周波数ドメインで適用された従来強化方法は、オーバーラップ加算窓に依存し、それがオーバーラップ長に応じた追加的な遅延を導入してしまう。結合型の手法は、オーバーラップ加算の窓掛けを必要とせず、スピーチコーデック (非特許文献 15) で適用されるような窓掛けスキームを使用し、よってアルゴリズム上の遅延の増大を回避している。

【 0 0 7 6 】

本提案方法に関して知られている課題は、信号位相が無傷で残る従来スペクトル・ウィナーフィルタリングとは異なり、提案方法は、位相を修正する時間ドメインフィルタを適

50

用するということである。そのような位相修正は、適切な全帯域通過フィルタの適用により容易に対処され得る。しかしながら、位相修正に起因する如何なる知覚的劣化も気づかれてこなかったもので、そのような全帯域通過フィルタは演算量を低く保持するために省略された。しかしながら、客観的評価において、方法同士の公正な比較ができるように、知覚的大きさSNRが測定されたことに注目されたい。この客観的手段は、本提案方法がカスケード処理よりも平均で3dB良好であることを示している。

【0077】

提案方法の性能優位性は、MUSHRリスニングテストの結果によって更に確認され、そのテストは平均で6.4ポイントの改善を示している。これらの結果は、結合型の強化及び符号化の適用が、品質と演算量との両方において全体システムにとって有益である一方で、CELPスピーチコーデックの低いアルゴリズム上の遅延を維持していることを証明している。

10

【0078】

図8は、線形予測符号化を使用して低減された背景ノイズを有するオーディオ信号を符号化する方法800の概略ブロック図を示す。この方法800は、オーディオ信号の背景ノイズ表現を推定するステップS802と、オーディオ信号表現からオーディオ信号の推定された背景ノイズ表現を差し引くことで、背景ノイズ低減済みオーディオ信号表現を生成するステップS804と、オーディオ信号表現に対し線形予測分析を行って線形予測フィルタ係数の第1セットを取得し、かつ背景ノイズ低減済みオーディオ信号表現に対し線形予測分析を行って線形予測フィルタ係数の第2セットを取得するステップS806と、取得されたLPC係数の第1セットと取得されたLPC係数の第2セットとによって時間ドメインフィルタのカスケードを制御して、オーディオ信号から残差信号を取得するステップS808と、を含む。

20

【0079】

この明細書において、線上の信号は時には線の参照番号で呼ばれ、又は時には線に帰属する参照番号そのもので示されている。従って、この表記法では、ある信号を有する線が信号それ自体を示している。線は、回路接続された実装における物理的な線であり得る。しかし、コンピュータ実装においては物理的な線は存在せず、線によって表現された信号は、1つの計算モジュールから他の計算モジュールへと伝送される。

【0080】

本発明はこれまでブロック図の文脈で説明されており、その中で各ブロックは実際の又は理論上のハードウェア要素を表していたが、本発明はコンピュータ実装された方法によってもまた実現され得る。後者の場合には、各ブロックは対応する方法ステップを表し、これらのステップは、対応する理論上または物理的なハードウェアのブロックによって実行される機能を表している。

30

【0081】

これまで幾つかの態様を装置の文脈で示してきたが、これらの態様は対応する方法の説明をも表しており、1つのブロック又は装置が1つの方法ステップ又は方法ステップの特徴に対応することは明らかである。同様に、方法ステップを説明する文脈で示した態様もまた、対応する装置の対応するブロックもしくは項目又は特徴を表している。方法ステップの幾つか又は全ては、例えばマイクロプロセッサ、プログラム可能なコンピュータ又は電子回路など、ハードウェア装置により（ハードウェア装置を使用して）実行されてもよい。幾つかの実施形態において、最も重要な方法ステップの1つ以上が、そのような装置によって実行されてもよい。

40

【0082】

本発明の、伝送され又は符号化された信号は、デジタル記憶媒体上に記憶されることができ、又は、例えばインターネットのような無線伝送媒体または有線伝送媒体などの伝送媒体上で伝送されることができる。

【0083】

所定の実施要件にもよるが、本発明の実施形態は、ハードウェア又はソフトウェアにおい

50

て実施可能である。この実施は、その中に格納される電子的に読み取り可能な制御信号を有し、本発明の各方法が実行されるようにプログラム可能なコンピュータシステムと協働する（又は協働可能な）、デジタル記憶媒体、例えばフレキシブルディスク、DVD、ブルーレイ、CD、ROM、PROM、EPROM、EEPROM、フラッシュメモリなどのデジタル記憶媒体を使用して実行することができる。従って、デジタル記憶媒体はコンピュータ読み取り可能であり得る。

【0084】

本発明に従う幾つかの実施形態は、上述した方法の1つを実行するようプログラム可能なコンピュータシステムと協働可能で、電子的に読み取り可能な制御信号を有するデータキャリアを含む。

10

【0085】

一般的に、本発明の実施例は、プログラムコードを有するコンピュータプログラム製品として構成することができ、そのプログラムコードは当該コンピュータプログラム製品がコンピュータ上で作動するときに、本発明の方法の一つを実行するよう作動可能である。そのプログラムコードは例えば機械読み取り可能なキャリアに記憶されていても良い。

【0086】

本発明の他の実施形態は、上述した方法の1つを実行するための、機械読み取り可能なキャリアに格納されたコンピュータプログラムを含む。

【0087】

換言すれば、本発明の方法のある実施形態は、そのコンピュータプログラムがコンピュータ上で作動するときに、上述した方法の1つを実行するためのプログラムコードを有するコンピュータプログラムである。

20

【0088】

本発明の他の実施形態は、上述した方法の1つを実行するために記録されたコンピュータプログラムを含む、データキャリア（又はデジタル記憶媒体、又はコンピュータ読み取り可能な媒体などの非一時的記憶媒体）である。そのデータキャリア、デジタル記憶媒体又は記録された媒体は、典型的には有形及び/又は非一時的である。

【0089】

本発明の他の実施形態は、上述した方法の1つを実行するためのコンピュータプログラムを表現するデータストリーム又は信号列である。そのデータストリーム又は信号列は、例えばインターネットのようなデータ通信接続を介して伝送されるよう構成されても良い。

30

【0090】

他の実施形態は、上述した方法の1つを実行するよう構成又は適応された、例えばコンピュータ又はプログラム可能な論理デバイスのような処理手段を含む。

【0091】

他の実施形態は、上述した方法の1つを実行するためのコンピュータプログラムがインストールされたコンピュータを含む。

【0092】

本発明に係るさらなる実施形態は、上述した方法の1つを実行するためのコンピュータプログラムを受信器へ（例えば電子的又は光学的に）伝送するよう構成された装置又はシステムを含む。受信器は、例えばコンピュータ、モバイル装置、メモリ装置等であってもよい。この装置又はシステムは、例えばコンピュータプログラムを受信器へと送信するためのファイルサーバを含み得る。

40

【0093】

幾つかの実施形態においては、（例えば書換え可能ゲートアレイのような）プログラム可能な論理デバイスが、上述した方法の幾つか又は全ての機能を実行するために使用されても良い。幾つかの実施形態では、書換え可能ゲートアレイは、上述した方法の1つを実行するためにマイクロプロセッサと協働しても良い。一般的に、そのような方法は、好適には任意のハードウェア装置によって実行される。

【0094】

50

上述した実施形態は、本発明の原理を単に例示的に示したに過ぎない。本明細書に記載した構成及び詳細について修正及び変更が可能であることは、当業者にとって明らかである。従って、本発明は、本明細書に実施形態の説明及び解説の目的で提示した具体的詳細によって限定されるものではなく、添付した特許請求の範囲によってのみ限定されるべきである。

【 図 1 】

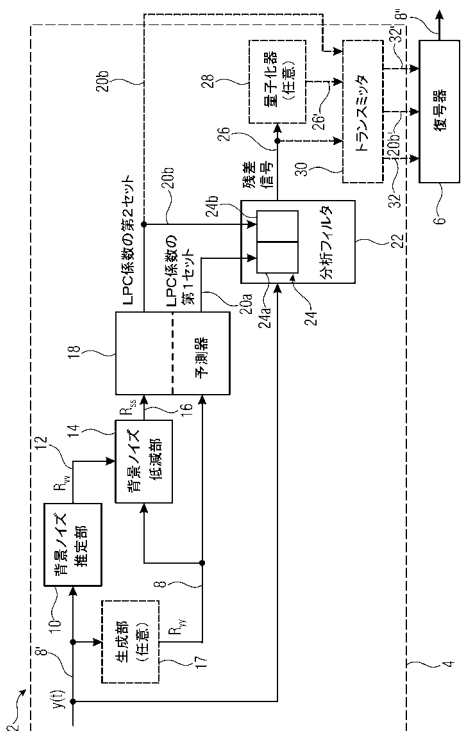


Fig. 1

【 図 2 】

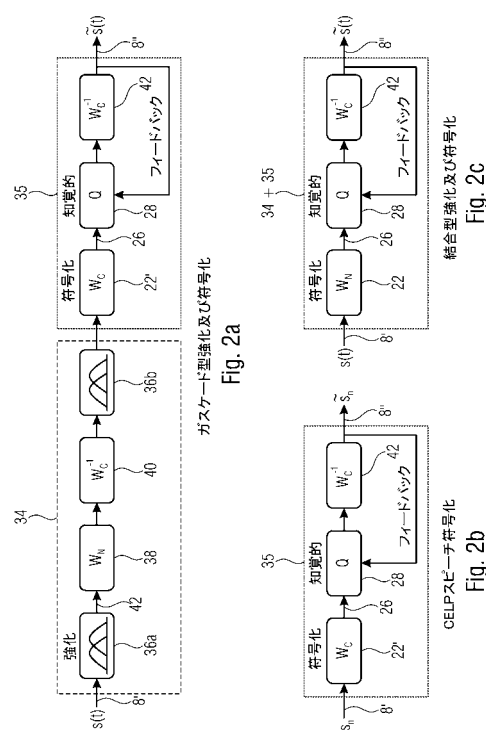
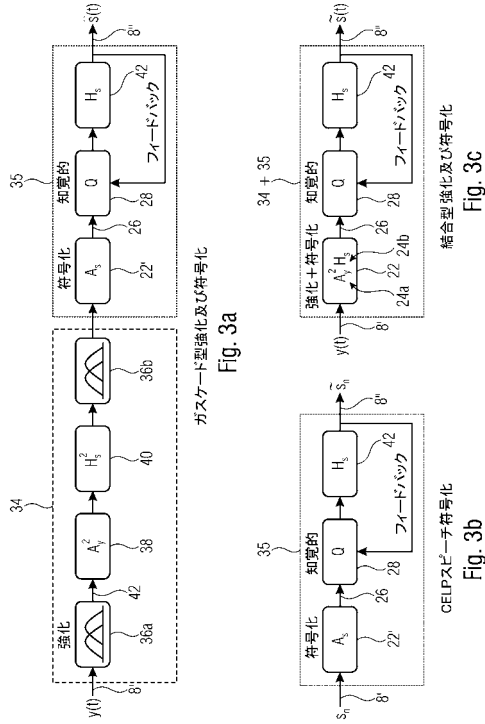


Fig. 2a

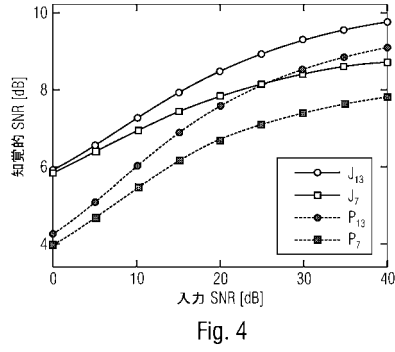
Fig. 2c

Fig. 2b

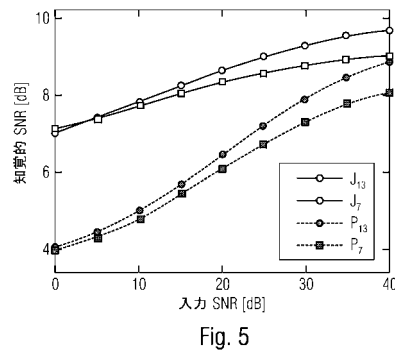
【 図 3 】



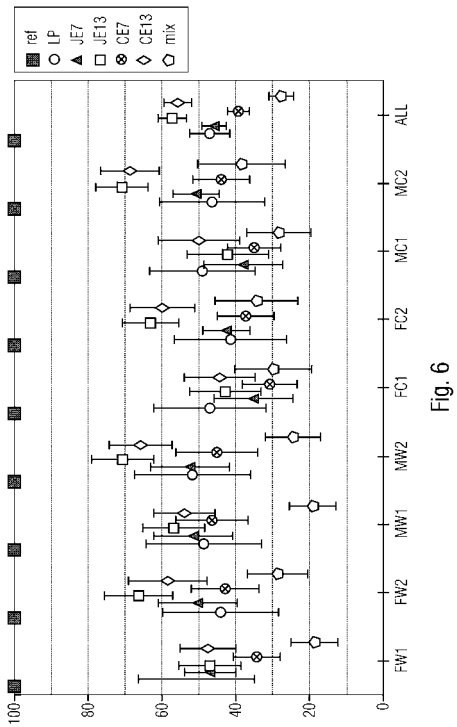
【 図 4 】



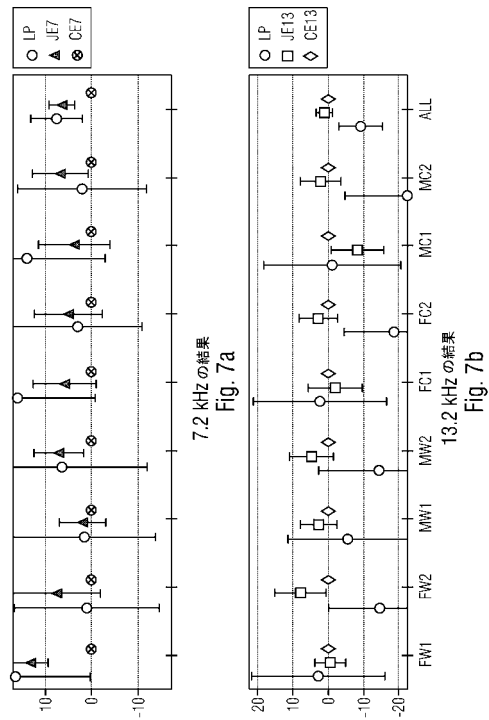
【 図 5 】



【 図 6 】



【 図 7 】



【 図 8 】

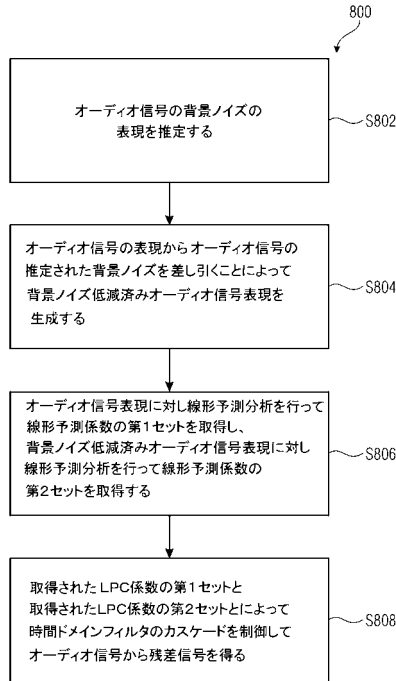


Fig. 8

【 手続 補正書 】

【 提出日 】 平成30年4月17日 (2018.4.17)

【 手続 補正 1 】

【 補正対象書類名 】 特許請求の範囲

【 補正対象項目名 】 全文

【 補正方法 】 変更

【 補正の内容 】

【 特許請求の範囲 】

【 請求項 1 】

線形予測符号化を使用して低減された背景ノイズを有するオーディオ信号 (8 ') を符号化する符号器 (4) であって、
 オーディオ信号 (8 ') の背景ノイズ表現 (1 2) を推定するよう構成された背景ノイズ推定部 (1 0) と、
 オーディオ信号表現 (8) から前記オーディオ信号 (8 ') の推定された背景ノイズ表現 (1 2) を差し引くことによって、背景ノイズ低減済みオーディオ信号表現 (1 6) を生成するよう構成された背景ノイズ低減部 (1 4) と、
 前記オーディオ信号表現 (8) に対し線形予測分析を行って線形予測フィルタ (L P C) 係数の第 1 セット (2 0 a) を取得し、前記背景ノイズ低減済みオーディオ信号表現 (1 6) に対し線形予測分析を行って線形予測フィルタ (L P C) 係数の第 2 セット (2 0 b) を取得するよう構成された予測部 (1 8) と、
 取得された L P C 係数の第 1 セット (2 0 a) と取得された L P C 係数の第 2 セット (2 0 b) とによって制御される時間ドメインフィルタのカスケード (2 4 , 2 4 a , 2 4 b) から構成され、前記オーディオ信号 (8 ') から残差信号 (2 6) を得る分析フィルタ (2 2) と、
 を含む符号器 (4) 。

【請求項 2】

前記時間ドメインフィルタのカスケード(24)は、前記取得されたLPC係数の第1セット(20a)を使用する2回の線形予測フィルタ(24a)と、前記取得されたLPC係数の第2セット(20b)を使用する1回の他の線形予測フィルタの逆(24b)とを含む、請求項1に記載の符号器(4)。

【請求項 3】

前記時間ドメインフィルタのカスケード(24)はウィナーフィルタである、請求項1又は2に記載の符号器(4)。

【請求項 4】

前記背景ノイズ推定部(10)は、前記オーディオ信号(8')の背景ノイズ表現(12)として、前記背景ノイズの自己相関を推定するよう構成され、
前記背景ノイズ低減部(14)は、オーディオ信号の自己相関(8)から前記背景ノイズの自己相関(12)を差し引くことにより、前記背景ノイズ低減済みオーディオ信号表現(16)を生成するよう構成され、前記オーディオ信号の自己相関(8)は前記オーディオ信号表現であり、前記背景ノイズ低減済みオーディオ信号表現(16)は背景ノイズ低減済みオーディオ信号の自己相関である、請求項1乃至3のいずれかに記載の符号器(4)。

【請求項 5】

前記表現は自己相関である、請求項1乃至4のいずれかに記載の符号器(4)。

【請求項 6】

前記LPC係数の第2セット(20b)を伝送するよう構成されたトランスミッタ(30)をさらに含む、請求項1乃至5のいずれかに記載の符号器(4)。

【請求項 7】

前記残差信号(26)を伝送するよう構成されたトランスミッタをさらに含む、請求項1乃至6のいずれかに記載の符号器(4)。

【請求項 8】

前記残差信号(26)を伝送の前に量子化及び/又は符号化するよう構成された量子化部(28)をさらに含む、請求項1乃至7のいずれかに記載の符号器(4)。

【請求項 9】

前記LPC係数の第2セット(20b)を伝送の前に量子化及び/又は符号化するよう構成された量子化部(28)をさらに含む、請求項1乃至8のいずれかに記載の符号器(4)。

【請求項 10】

前記量子化部は、符号励振線形予測(CELP)、エントロピー符号化、又は変換符号化励振(TCX)を使用するよう構成された、請求項8又は9に記載の符号器。

【請求項 11】

請求項1乃至10のいずれかに記載の符号器(4)と、
符号化されたオーディオ信号を復号化するよう構成された復号器(6)と、
を含むシステム(2)。

【請求項 12】

線形予測符号化を使用して低減された背景ノイズを有するオーディオ信号を符号化する方法(800)であって、
オーディオ信号の背景ノイズ表現を推定するステップ(S802)と、
オーディオ信号表現から前記オーディオ信号の推定された背景ノイズ表現を差し引くことによって、背景ノイズ低減済みオーディオ信号表現を生成するステップ(S804)と、
前記オーディオ信号表現に対し線形予測分析を行って線形予測フィルタ(LPC)係数の第1セットを取得し、前記背景ノイズ低減済みオーディオ信号表現に対し線形予測分析を行って線形予測フィルタ(LPC)係数の第2セットを取得するステップ(S806)と、
取得されたLPC係数の第1セットと取得されたLPC係数の第2セットとによって時間

ドメインフィルタのカスケードを制御して、前記オーディオ信号から残差信号を得るステップ(S 808)と、
を含む方法。

【請求項13】

請求項12に記載の方法を実行するための、プログラムコードを有するコンピュータプログラム。

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】0023

【補正方法】変更

【補正の内容】

【0023】

実施形態によれば、残差信号は、残差信号及び/又はLPC係数の第2セット20bを伝送の前に量子化及び/又は符号化するよう構成された、量子化部28へと提供されてもよい。量子化部は、例えば変換符号化励振(TCX)、符号励振線形予測(CELP)又は例えばエントロピー符号化などのロスレス符号化などを実行してもよい。

【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】0024

【補正方法】変更

【補正の内容】

【0024】

更なる実施形態によれば、残差信号の符号化は、量子化部28における符号化の代替例として、トランスミッタ30において実行されてもよい。このように、トランスミッタは、例えば変換符号化励振(TCX)、符号励振線形予測(CELP)又は例えばエントロピー符号化などのロスレス符号化を実行して、残差信号を符号化する。更に、トランスミッタは、LPC係数の第2セットを伝送するよう構成されてもよい。任意選択的な受信機は復号器6である。このように、トランスミッタ30は、残差信号26又は量子化済み残差信号26'を受信してもよい。一実施例によれば、少なくとも量子化済み残差信号が量子化部内で既に符号化されていない場合、トランスミッタは残差信号又は量子化済み残差信号を符号化してもよい。残差信号又は代替的に量子化済み残差信号を任意選択的に符号化した後、トランスミッタへと提供されたそれぞれの信号は、符号化済み残差信号32又は符号化されかつ量子化された残差信号32'として伝送される。更に、トランスミッタは、LPC係数の第2セット20bを受け取り、任意選択的に例えば残差信号を符号化するために使用されたのと同じ符号化方法を用いてそれを符号化して、その符号化されたLPC係数の第2セット20b'を、LPC係数の第1セットを伝送せずに、例えば復号器6へと更に伝送してもよい。換言すれば、LPC係数の第1セット20aは伝送される必要がない。

【手続補正4】

【補正対象書類名】明細書

【補正対象項目名】0028

【補正方法】変更

【補正の内容】

【0028】

図2aと図3aとは、カスケード型の強化および符号化を実行する信号処理連鎖の強化部分と符号化部分とを示す。強化部分34は周波数ドメインで作動してもよく、ブロック36a及び36bは、例えばMDC Tを使用する時間-周波数変換および例えばIMDC Tを使用する周波数-時間変換を実行してもよく、又は任意の他の適切な変換を使用して時間-周波数変換および周波数-時間変換を実行してもよい。フィルタ38と40とは、周波数変換されたオーディオ信号42の背景ノイズ低減を実行してもよい。ここで、背景ノ

イズの周波数部分は、オーディオ信号 8' の周波数スペクトルに対するそれらの影響力を低減することでフィルタリングされてもよい。よって、周波数 - 時間変換部 36 b は、周波数ドメインから時間ドメインへの逆変換を実行してもよい。強化部分 34 において背景ノイズ低減が実行された後で、符号化部分 35 は低減された背景ノイズを有するオーディオ信号の符号化を実行してもよい。従って、分析フィルタ 22' は適切な LPC 係数を用いて残差信号 26 を計算する。残差信号は量子化されて合成フィルタ 44 へと提供されてもよく、その合成フィルタは図 2 a と図 3 a の場合には分析フィルタ 22' の逆である。合成フィルタ 42 は図 2 a と図 3 a の場合には分析フィルタ 22' の逆であるため、残差信号 26 を決定するために使用された LPC 係数が復号器へと伝送されて、復号化済みオーディオ信号 8'' を決定する。

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No PCT/EP2016/072701

A. CLASSIFICATION OF SUBJECT MATTER INV. G10L19/06 G10L21/0208 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G10L		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 1 944 761 A1 (SIEMENS NETWORKS GMBH & CO KG [DE]) 16 July 2008 (2008-07-16) paragraph [0005] - paragraph [0007] paragraph [0012] - paragraph [0022] paragraph [0028] - paragraph [0050] paragraph [0060] - paragraph [0066] ----- -/--	1-13
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.		
<input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents :		
"A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed		
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
24 October 2016		02/11/2016
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Virette, David

1

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2016/072701

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>SRIRAM SRINIVASAN ET AL: "Codebook Driven Short-Term Predictor Parameter Estimation for Speech Enhancement", IEEE TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, USA, vol. 14, no. 1, January 2006 (2006-01), pages 163-176, XP002551735, ISSN: 1558-7916, DOI: 10.1109/TSA.2005.854113 page 163, paragraph I. - page 166, paragraph II.</p> <p>-----</p>	1-13
A	<p>US 6 263 307 B1 (ARSLAN LEVENT M [US] ET AL) 17 July 2001 (2001-07-17) column 1, line 57 - column 2, line 22 column 4, line 20 - line 47 column 11, line 14 - column 13, line 46; figures 4,9b column 14, line 7 - column 15, line 53; figure 5</p> <p>-----</p>	1-13
A	<p>Emmanuel Thepie Fapi ET AL: "Noise Reduction within Network through Modification of LPC Parameters", 7th International ITG Conference on Source and Channel Coding (SCC), 2008, 14 January 2008 (2008-01-14), pages 1-6, XP055312348, Retrieved from the Internet: URL:http://ieeexplore.ieee.org/ielx5/5755489/5755490/05755780.pdf?tp=&arnumber=5755780&isnumber=5755490 [retrieved on 2016-10-19] page 1, paragraph 2. - page 4, paragraph 5.; figures 1,2,3</p> <p>-----</p>	1-13

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2016/072701

Patent document cited in search report	Publication date	Patent family member(s)	Publication date	
EP 1944761	A1	16-07-2008	EP 1944761 A1	16-07-2008
			WO 2008086920 A1	24-07-2008

US 6263307	B1	17-07-2001	US RE43191 E	14-02-2012
			US 6263307 B1	17-07-2001

 フロントページの続き

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

(72)発明者 ベックストレム, トム

フィンランド共和国 00100 ヘルシンキ ヒータニーメンカツ 1 イー 13

(72)発明者 ヨキーネン, エンマ

フィンランド共和国 00370 ヘルシンキ コナランティエー 18 シー 43