

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2020年8月13日(13.08.2020)



(10) 国際公開番号

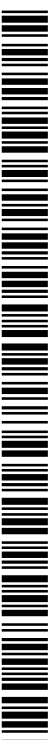
WO 2020/162188 A1

- (51) 国際特許分類:
G10L 21/0272 (2013.01)
- (21) 国際出願番号: PCT/JP2020/002205
- (22) 国際出願日: 2020年1月23日(23.01.2020)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2019-018424 2019年2月5日(05.02.2019) JP
- (71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).
- (72) 発明者: 佐藤 遼太郎 (SATO, Ryotaro); 〒1808585 東京都武蔵野市緑町三丁目9番11号 N T T 知的財産センタ内 Tokyo (JP). 丹羽健太(NIWA, Kenta); 〒1808585 東京都武蔵野市緑町三丁目9番11号 N T T 知的財産センタ内 Tokyo (JP). 原田 登(HARADA, Noboru); 〒1808585 東京都武蔵野市緑町三丁目9番11号 N T T 知的財産センタ内 Tokyo (JP).
- (74) 代理人: 中尾 直樹, 外 (NAKAO, Naoki et al.); 〒1600022 東京都新宿区新宿三丁目1番22号 新宿NSOビル6階 Tokyo (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): 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, JO, KE, KG, KH, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(54) **Title:** LATENT VARIABLE OPTIMIZATION DEVICE, FILTER COEFFICIENT OPTIMIZATION DEVICE, LATENT VARIABLE OPTIMIZATION METHOD, FILTER COEFFICIENT OPTIMIZATION METHOD, AND PROGRAM

(54) 発明の名称: 潜在変数最適化装置、フィルタ係数最適化装置、潜在変数最適化方法、フィルタ係数最適化方法、プログラム

(57) **Abstract:** Provided is a technology which optimizes a latent variable by using a cost function obtained by taking into simultaneous consideration of a plurality of probabilistic assumptions. In this latent variable optimization device including an optimization unit which optimizes a latent variable $\sim w^*$, the optimization unit takes $v_j (1 \leq j \leq J)$ as an auxiliary variable of the latent variable $\sim w^*$ expressed as $v_j = D_j \sim w^* + b_j$ by using a matrix D_j and a vector b_j , and cost terms of the auxiliary variable v_j are expressed by using a probability distribution of the auxiliary variable v_j , which is a log-concave distribution. The optimization unit optimizes the latent variable $\sim w^*$ by solving a minimization problem of the cost function including the sum of cost terms of the auxiliary variable v_j .



WO 2020/162188 A1

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (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), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告 (条約第21条(3))

(57) 要約: 複数の確率的仮定を同時に考慮したコスト関数を用いて潜在変数を最適化する技術を提供する。潜在変数 $\sim w^*$ を最適化する最適化部を含む潜在変数最適化装置であって、最適化部は、 $v_j (1 \leq j \leq J)$ を行列 D_j とベクトル b_j を用いて $v_j = D_j \sim w^* + b_j$ と表される潜在変数 $\sim w^*$ の補助変数とし、補助変数 v_j のコスト項は、log-concaveである補助変数 v_j の確率分布を用いて表されるものであり、最適化部は、補助変数 v_j のコスト項の和を含むコスト関数の最小化問題を解くことにより潜在変数 $\sim w^*$ を最適化する。

明 細 書

発明の名称：

潜在変数最適化装置、フィルタ係数最適化装置、潜在変数最適化方法、フィルタ係数最適化方法、プログラム

技術分野

[0001] 本発明は、目的音強調におけるフィルタ係数など最適化の対象となるモデルの潜在変数を最適化する技術に関する。

背景技術

[0002] 特定の方角から到来する音のみを強調し他方向の雑音を抑圧する信号処理手法として、マイクロホンアレイを用いたビームフォーミングがよく知られている。この手法は電話会議システム、自動車内のコミュニケーションシステム、スマートスピーカー等で実用化されている。ビームフォーミングに関する従来手法の多くは、何らかの制約のもとでコスト関数最小化問題を解くことで最適なフィルタを導出していた。例えば、非特許文献1に記載のMVDRビームフォーマは、出力信号のパワーをコスト関数としてこれを目的音源方角に対する無歪み特性制約のもと最小化することで得られる。また、最尤（ML）ビームフォーマは、出力信号に含まれる雑音のパワーをコスト関数としてこれを最小化することで導出される。その他、ビームフォーマの性能を向上させるために、付加的な制約やコスト項をコスト関数に追加する試みもこれまで行われてきた。

先行技術文献

非特許文献

[0003] 非特許文献1：J. Capon, “High-resolution frequency-wavenumber spectrum analysis”, Proceedings of the IEEE, vol.57, no.8, pp.1408-1418, Aug. 1969.

発明の概要

発明が解決しようとする課題

[0004] ビームフォーマを現実の状況に適用する際、複数の特性を同時に持たせることができれば、応用上有用と考えられる。例えば、音声に対して高い強調性能を保ちつつ、低遅延特性も両立したビームフォーマが必要とされるような場面があるだろう。ビームフォーマの特性への要請は、理論的にはビームフォーマのフィルタ係数から定義される補助変数に対する確率的仮定という形式でモデル化できる。例えば、もし強調したい音が人間の音声であるという事前知識があったとすれば、推定される信号は時間周波数領域でラプラス分布等のスパース性の高い分布に従うと仮定するのが妥当であろう。また、フィルタ係数は周波数方向について連続かつなめらかに変化するのが自然であるということは経験的に知られているが、従来手法ではフィルタ係数の周波数方向に関する特性を考慮していなかったため、空間相関行列がランク落ちするような周波数ビンで解が不安定になり、この特性が満たされないという状況がみられた。もし、なめらかさを考慮した設計を行うことができれば、遅延の少ないフィルタが得られるという効果も期待される。以上のような仮定をフィルタ係数の推定に同時に組み込むことができれば、目的音強調だけでなく多様な特性を持ったビームフォーマが構成できると期待される。

[0005] しかしながら、これまではコスト関数の最適化に関する数理的手法の検討が十分ではなく、特に複数の確率的仮定を同時に考慮したコスト関数の最適化に関する検討は行われてこなかった。

[0006] そこで本発明では、複数の確率的仮定を同時に考慮したコスト関数を用いて潜在変数を最適化する技術を提供することを目的とする。

課題を解決するための手段

[0007] 本発明の一態様は、潜在変数 $\tilde{w}^*$ を最適化する最適化部を含む潜在変数最適化装置であって、 v_j ($1 \leq j \leq J$) を行列 D_j とベクトル b_j を用いて $v_j = D_j \tilde{w}^* + b_j$ と表される潜在変数 $\tilde{w}^*$ の補助変数とし、補助変数 v_j のコスト項は、log-concaveである補助変数 v_j の確率分布を用いて表されるものであり、前記最適化部は、補助変数 v_j のコスト項の和を含むコスト関数の最小化問題を解くことにより潜在変

数 $\tilde{w}^*$ を最適化する。

発明の効果

[0008] 本発明によれば、複数の確率的仮定を同時に考慮したコスト関数を用いて潜在変数を最適化することが可能となる。

図面の簡単な説明

[0009] [図1]潜在変数最適化アルゴリズムを示す図である。

[図2]潜在変数最適化装置100（フィルタ係数最適化装置100）の構成を示すブロック図である。

[図3]潜在変数最適化装置100（フィルタ係数最適化装置100）の動作を示すフローチャートである。

[図4]最適化部120の構成を示すブロック図である。

[図5]最適化部120の動作を示すフローチャートである。

発明を実施するための形態

[0010] 以下、本発明の実施の形態について、詳細に説明する。なお、同じ機能を有する構成部には同じ番号を付し、重複説明を省略する。

[0011] 各実施形態の説明に先立って、この明細書における表記方法について説明する。

[0012] $_$ （アンダースコア）は下付き添字を表す。例えば、 x^{y_z} は y_z が x に対する上付き添字であり、 x_{y_z} は y_z が x に対する下付き添字であることを表す。

[0013] また、ある文字 x に対する $\hat{x}$ や $\tilde{x}$ のような上付き添え字の“ $\hat{}$ ”や“ $\tilde{}$ ”は、本来“ x ”の真上に記載されるべきであるが、明細書の記載表記の制約上、 $\hat{x}$ や $\tilde{x}$ と記載しているものである。

[0014] <技術的背景>

本発明の実施形態では、フィルタ係数自身やフィルタ係数から定まる補助変数に対する確率的仮定に基づいて設計したコスト関数を用いて、フィルタ係数を最適化（学習）する。ここで、補助変数はフィルタ係数のアフィン変換として表現されるものに限定されるが、例えば推定された目的音、出力信号に含まれる雑音、隣接周波数ビン間のフィルタ係数の差分などはこの範疇

に含まれる。フィルタ係数やその補助変数に対して仮定される確率分布がすべてlog-concave（対数凹）であれば、フィルタ係数や補助変数を互いに独立とみなした際の同時分布もまたlog-concaveとなるため、負の対数尤度は凸関数となり、コスト関数最適化問題は線形関係式で制約された補助変数に関する凸関数の最適化問題に帰着される。この最適化問題は、例えば、交互方向乗数法(Alternating Direction Method of Multipliers, ADMM)を用いて解くことができ、最適なフィルタ係数を効率的に計算できる。

[0015] 以下、上記説明した本発明の実施形態の原理について詳しく説明していく。まず、最初に、確率に基づく最適化の観点からビームフォーミングの問題を定式化し、従来のビームフォーミングの最適化問題がこの定式化の枠で記述できることを説明する。

[0016] 《ビームフォーミングの問題の定式化》

ここでは、記号・ノーテーションを定義し、問題を定式化する。まず、ビームフォーミングの問題を数理的に記述するための種々の定義を行う。

[0017] 空間中に単一の目的音源と複数個の妨害音源があり、M個の無指向性マイクロホンアレイでこれらの混合音を収録し、観測されたMチャンネル信号をビームフォーミングフィルタに通すことで、特定の方角から到来する目的音を強調するという状況を考える。この状況を記述するモデルを導入するため、まず変数を定義する。以下の議論は、基本的に短時間フーリエ変換(short-time Fourier transform, STFT)ドメインで行う。

[0018] $a_f \in \mathbb{C}^M$ ($f=1, \dots, F$)を周波数ビン f における目的音源からマイクロホンアレイへの伝達関数、 $a_{ikf} \in \mathbb{C}^M$ を周波数ビン f における k 番目の妨害音源からマイクロホンアレイへの伝達関数、 $s_{f,t} \in \mathbb{C}$ を周波数ビン f 、時間フレーム t ($t=1, \dots, T$)における目的音信号、 $n_{ikf,t} \in \mathbb{C}^M$ を周波数ビン f 、時間フレーム t における k 番目の妨害音信号とする。これらの記号を用いて、マイクロホンアレイが観測する信号 $z_{f,t}$ は、瞬時混合仮定のもと

[0019] [数1]

$$z_{f,t} = s_{f,t} a_f + \sum_k n_{ikf,t} a_{ikf} + n_{bf,t} \quad \cdots (1)$$

[0020] と表される。ここで、 $n_{bf,t}$ は、特定の妨害音源由来と仮定していない雑音信号（例えば、マイクの性能自体に起因する雑音）を表す。

[0021] 我々が求めたいのは、観測信号 $z_{f,t}$ から目的音信号 $s_{f,t}$ の精度よい推定値 $y_{f,t}$ を与えるような線形フィルタである。以下では、この線形フィルタのフィルタ係数を $w_f \in \mathbb{C}^M$ とする。推定値 $y_{f,t}$ の時間フレームを表す添字 t を省略して、推定値 y_f と表すことにすると、 z_f , y_f , w_f の関係は

[0022] [数2]

$$y_f = w_f^H z_f \quad \cdots (2)$$

[0023] で与えられる。ここで、 H は複素共役転置を表す。

[0024] ここで、フィルタ係数と観測音に依存する変数を導入する。フィルタを用いることで、観測音から目的音を抽出できるならば、目的音を観測音から差し引くことで妨害音源などに起因する非目的音も推定できるはずである。そこで、観測音に含まれる非目的音の推定値 $e_f \in \mathbb{C}^M$ を

[0025] [数3]

$$e_f = z_f - y_f h_f = z_f - (w_f^H z_f) h_f \quad \cdots (3)$$

[0026] で定義する。ここで、 $h_f \in \mathbb{C}^M$ は目的音源方向のアレイマニフォールドベクトルである。本来、式(1)のようなモデルのもとではアレイマニフォールドベクトル h_f ではなく伝達関数 a_f を適用するのが望ましいが、実用上传達関数を常に正確に知るのは難しい。そこで、式(3)の定義ではアレイマニフォールドベクトル h_f を用いる。なお、ビームフォーマが適切に目的音を抽出できていれば、非目的音の推定値 e_f は主に妨害音と背景雑音によって構成されると期待される。

[0027] ここで更に、目的音と非目的音の推定値はともにフィルタ係数のアフィン変換として表現できることに着目し、フィルタ係数を用いた変換式として表現する。

[0028] なお、以後、表記を簡潔にするため、周波数ビン f に関する添え字を持つ任意の変数 x_f に関して、すべての周波数ビンに関する情報をまとめたものを $\tilde{x} = (x_1^T, \dots, x_F^T)^T$ で表すこととする。

[0029] また、時間フレーム t について行列 F_t と G_t を

[0030] [数4]

$$F_t = \text{diag}[z_{1,t}, z_{2,t}, \dots, z_{F,t}]^T \in C^{F \times MF} \quad \dots(4)$$

$$G_t = \text{diag}[h_1 z_{1,t}^T, h_2 z_{2,t}^T, \dots, h_F z_{F,t}^T] \in C^{MF \times MF} \quad \dots(5)$$

[0031] で定義する。

[0032] すると、時間フレーム t におけるビームフォーマの出力である目的音の推定値 $\tilde{y}_t$ （以下、推定目的音という）と雑音の推定値 $\tilde{e}_t$ （以下、推定雑音、推定非目的音という）はそれぞれ以下のフィルタ係数 $\tilde{w}^*$ のアフィン変換の形式で表現される。

[0033] [数5]

$$\tilde{y}_t = F_t \tilde{w}^* \quad \dots(6)$$

$$\tilde{e}_t = -G_t \tilde{w}^* + \tilde{z}_t \quad \dots(7)$$

[0034] （ただし、 $*$ は複素共役）

[0035] 《従来のビームフォーミングの最適化問題》

まず、従来のビームフォーミングの最適化問題を、確率モデルの観点から定義されたコスト関数の最小化問題として記述する。

[0036] $\tilde{y}$, $\tilde{e}$, $\tilde{w}^*$ が確率変数であると解釈し、これらの確率分布 $P_y(\tilde{y})$, $P_e(\tilde{e})$, $P_w(\tilde{w}^*)$ は既に分かっているものと仮定する。このうち、確率分布 $P_y(\tilde{y})$, $P_e(\tilde{e})$ については、音の統計的な性質が反映されているものと期待される。一方、確率分布 $P_w(\tilde{w}^*)$ は、目的音源方角に対する周波数応答の仮定を表現するためにしばしば用いられるものである。これらの仮定のもと、観測音の時系列 $\{\tilde{z}_t\}_{t=1}^T$ に対する確率変数 $\tilde{w}^*$ の尤度関数は

[0037] [数6]

$$L(\tilde{w}^*; \{\tilde{z}_t\}) \propto \prod_{t=1}^T P_y(\tilde{y}_t; \tilde{w}^*) P_e(\tilde{e}_t; \tilde{w}^*) \cdot P_w(\tilde{w}^*) \quad \dots(8)$$

[0038] と表される。ただし、 $\tilde{y}_t$, $\tilde{e}_t$ は式(6)、式(7)で表される $\tilde{w}^*$ のアフィン変換によって定まる。この尤度を $\tilde{w}^*$ に関して最大化することで、確率モデルの意味

で最適なフィルタが導出できる。尤度最大化は負の対数尤度最小化と等価であるので、解くべきは

[0039] [数7]

$$\min_{\tilde{\mathbf{w}}^*} -\log L(\tilde{\mathbf{w}}^*; \{\tilde{\mathbf{z}}_f\}) \quad \cdots (9)$$

[0040] という形式の問題になる。

[0041] 様々な従来のビームフォーミングの最適化問題が式(9)に基づく定式化として解釈できる。以下では、具体例として非特許文献1のMVDRビームフォーマの最適化問題について説明する。

[0042] [フィルタ設計フェーズ：フィルタ係数 $\tilde{\mathbf{w}}^*$ のコスト関数]

周波数ビン f で観測音 $\mathbf{z}_f$ の空間相関行列の推定値 $\mathbf{R}_f = \mathbf{E}_{\mathbf{z}_f}[\mathbf{z}_f \mathbf{z}_f^H]$ ($f=1, \dots, F$)が既知であり、また観測音に含まれる推定非目的音 $\tilde{\mathbf{e}}_t$ が正規分布 $N(0, \mathbf{R}_f)$ に従う（つまり、 $\mathbf{e}_{f,t} \sim N(0, \mathbf{R}_f)$ ）と仮定する。

[0043] このとき、尤度関数の非目的音に対する項 $\prod_t P_e(\tilde{\mathbf{e}}_t; \tilde{\mathbf{w}}^*)$ は

[0044] [数8]

$$\prod_{t=1}^T P_e(\tilde{\mathbf{e}}_t; \tilde{\mathbf{w}}^*) \propto \prod_{t=1}^T \prod_{f=1}^F \exp(-\mathbf{e}_{f,t}^H \mathbf{R}_f^{-1} \mathbf{e}_{f,t}) \quad \cdots (10)$$

$$= \exp \left(- \sum_{t,f} (\mathbf{z}_{f,t} - (\mathbf{w}_f^H \mathbf{z}_{f,t}) \mathbf{h}_f)^H \mathbf{R}_f^{-1} (\mathbf{z}_{f,t} - (\mathbf{w}_f^H \mathbf{z}_{f,t}) \mathbf{h}_f) \right) \cdots (11)$$

[0045] と表される。他の項（ $\prod_t P_y(\tilde{\mathbf{y}}_t; \tilde{\mathbf{w}}^*)$ や $P_w(\tilde{\mathbf{w}}^*)$ ）にはここでは特に確率分布の仮定をおかない。

[0046] [フィルタ設計フェーズ：フィルタ係数 $\tilde{\mathbf{w}}^*$ の制約条件]

フィルタ係数 $\tilde{\mathbf{w}}^* = (\mathbf{w}_1^{*T}, \dots, \mathbf{w}_F^{*T})$ の各 $\mathbf{w}_f^*$ には目的音源方向に対する無歪み制約 $\mathbf{w}_f^H \mathbf{h}_f = 1$ を課す。

[0047] [フィルタ設計フェーズ：最適化問題]

これらの仮定のもと、単純な平方完成によって式(9)に基づく最適化問題は周波数ビン f について

[0048]

[数9]

$$\min_{w_f} (w_f - \gamma_f R_f^{-1} h_f)^H R_f (w_f - \gamma_f R_f^{-1} h_f) \quad \cdots (12)$$

$$s.t. \quad w_f^H h_f = 1$$

[0049] (ただし、 $\gamma_f = (h_f^H R_f^{-1} h_f)^{-1}$) という形に帰着できる。

[0050] [フィルタ設計フェーズ：コスト関数最適化]

式(12)の問題の解は、よく知られたMVDRビームフォーマ(すなわち、 $\gamma_f R_f^{-1} h_f$)である。

[0051] 次に、以上の手続きで得られたフィルタ係数を使用して実際にビームフォーマを動作させる際の計算(フィルタ使用フェーズ)について説明する。

[0052] [フィルタ使用フェーズ：フィルタ再設計]

ビームフォーミングの処理では、観測音をフレーム毎に区切り、各フレームに対して離散フーリエ変換を行う必要があるが、リアルタイムでビームフォーミングを行う状況ではフレーム長が長いと遅延が大きくなってしまう。そこで、低遅延のフィルタを再設計する。まず、フィルタ設計フェーズで設計したフィルタ係数 $\tilde{w}^*$ に対して逆フーリエ変換を行い、フィルタを時間領域の表現に戻すことで、各マイクロホン $m(m=1, \dots, M)$ のインパルス応答 $w_m[i]$ を得る。そして、入力として与えられた指定フレーム長 N_{tap} をもとに、各インパルス応答 $w_m[i]$ のうち最初の $N_{\text{tap}}/2$ 成分からなるベクトル $w'_{m1}[i]$ と最後の $N_{\text{tap}}/2$ 成分からなるベクトル $w'_{m2}[i]$ のみを取り出し(つまり、残りの要素はすべて無視し)、長さが N_{tap} に短縮された新たなインパルス応答

[0053] [数10]

$$w_m''[i] = [w'_{m2}^T, w'_{m1}^T]^T \quad \cdots (13)$$

[0054] を導入する。このインパルス応答 $w_m''[i]$ を再度離散フーリエ変換することで、要素数が N_{tap} に削減された(再設計された)フィルタ係数 $\tilde{w}^*$ が計算される。

[0055] [フィルタ使用フェーズ：離散フーリエ変換(DFT)]

次に、ビームフォーミング処理対象となる観測音を時間方向に N_{tap} サンプルずつ切り出して、切り出した各区間(フレーム)に対して離散フーリエ変換

を施し、STFT領域での観測音 $\tilde{z}$ を出力する。

[0056] [フィルタ使用フェーズ：畳み込み]

ここでは、STFT領域での観測音 $\tilde{z}$ とフィルタ係数 $\tilde{w}'^*$ を入力とし、式(2)のたたみ込みを行い、STFT領域での推定目的音 $\tilde{y}$ を出力する。

[0057] [フィルタ使用フェーズ：逆離散フーリエ変換（逆DFT）]

最後に、STFT領域での推定目的音 $\tilde{y}$ に逆離散フーリエ変換を施し、ビームフォーミング処理を施した時間領域波形、つまり時系列の推定目的音を得る。

[0058] 式(9)は、変数 $\tilde{w}^*$ に関する最適化問題という定式化になっており、上述のMVDRビームフォーマのような比較的単純な例に対しては容易に解くことができるが、コスト関数が複雑な式になる場合、同様に最適化を行うのは一般に困難である。このことから、従来の手法は2つの問題点を抱えていることがわかる。一つ目として、これまでの手法では目的音 $\tilde{y}$ や非目的音 $\tilde{e}$ に対して仮定される確率分布が正規分布のように単純なものに限定されてしまいがちであったが、正規分布は必ずしも音源分布の記述として適切ではない点がある。二つ目として、フィルタ係数 $\tilde{w}^*$ に対する制約として用いることができるコスト関数も制限されていて、特に多様な確率的仮定を同時に考慮するのは困難である点がある。これまでもフィルタ係数 $\tilde{w}^*$ に対して付加的な制約の導入を検討したものもあったが、一般に多様な確率的仮定を同時に考慮することで構成されるような複雑なコスト関数の最小化は極めて難しい問題であった。特に、低遅延・安定性・高い雑音抑圧性能を同時に達成するようなビームフォーマを構成するのは困難であった。

[0059] 上記問題を解決するために、式(6)の推定目的音 $\tilde{y}_t$ や式(7)の推定雑音 $\tilde{e}_t$ のような補助変数に対しても確率的仮定をおき、コスト関数を各補助関数に関するコスト項の和として表現することを考える。以下では、この考えに基づくコスト関数の設計方法について説明する。

[0060] 《複数の特性を考慮したコスト関数に基づく最適化問題》

ビームフォーミングの最適化問題のための、新たなコスト関数は、さまざま

まな凸関数の項の和として表現する。また、各項の引数は、フィルタ係数 $\tilde{w}^*$ や新たに導入される（推定目的音 $\tilde{y}_t$ や推定雑音 $\tilde{e}_t$ のような）フィルタ係数 $\tilde{w}^*$ のアフィン変換として定義できる補助変数とする。。これらの要請を満たすコスト関数は、いずれも最適化問題が解きやすい。言い換えれば、これらの要請を満たす範囲内で自由にコスト関数を設計できる。以下、詳しく説明する。

[0061] [フィルタ設計フェーズ：フィルタ係数 $\tilde{w}^*$ の補助変数]

J を任意の自然数とし、補助変数 v_j ($j=1, \dots, J$)を導入する。補助変数 v_j としてはフィルタ係数 $\tilde{w}^*$ と線形の関係、つまり、線形関係式 $v_j = D_j \tilde{w}^* + b_j$ を満たすものを用いる。各補助変数 v_j とそれがみたす関係式は式(6)、式(7)の一般化になっており、この意味で線形関係式を満たすとの制約は従来の手法を包含するものである。

[0062] 簡単のため、 $\hat{v} = (v_1^T, \dots, v_J^T)^T$, $\hat{D} = (D_1^T, \dots, D_J^T)^T$, $\hat{b} = (b_1^T, \dots, b_J^T)^T$ のような表記を以下では採用する。

[0063] [フィルタ設計フェーズ：フィルタ係数 $\tilde{w}^*$ のコスト項・補助変数 v_j のコスト項]

コスト関数 L は、フィルタ係数 $\tilde{w}^*$ のコスト項 L_0 と補助変数 v_j のコスト項 L_j ($j=1, \dots, J$)を用いて

[0064] [数11]

$$L(\tilde{w}^*, \hat{v}) = L_0(\tilde{w}^*) + \sum_{j=1}^J L_j(v_j) \quad \cdots (14)$$

[0065] (ただし、 L_j ($j=0, \dots, J$)は凸関数) という形式で表す

[0066] 凸関数の和は凸関数であるから、コスト関数 L も凸関数となる。補助変数 v_j のコスト項が凸関数であるという制約は一見突飛かもしれないが、これは実は式(6)の推定目的音 $\tilde{y}_t$ や式(7)の推定雑音 $\tilde{e}_t$ などの補助変数の確率分布にlog-concaveなものを採用していることに他ならない。ここで、確率分布がlog-concaveであるとはその確率密度関数の対数の-1倍(negative log)が凸関数であるという意味である。正規分布やラプラス分布など音源の確率モデルの記

述に一般的に用いられる確率分布の多くもこの性質を満たしている。式(14)のコスト項 L_j は補助変数 v_j の確率密度関数の対数の-1倍と解釈できるため、コスト項の凸性はlog-concaveな確率分布のみを考える限り自動的に満たされる性質である。

[0067] [フィルタ設計フェーズ：最適化問題]

以上の検討から、我々が解くべき問題は

[0068] [数12]

$$\min_{\hat{v}, \tilde{w}^*} L_0(\tilde{w}^*) + \sum_{j=1}^J L_j(v_j) \quad s.t. \quad \hat{v} = \hat{D}\tilde{w}^* + \hat{b} \quad \dots(15)$$

[0069] という典型的な線形制約付き凸最適化問題に帰着できる。式(15)の問題は、フィルタ係数に関する項と補助変数に関する項に分割し、それぞれについて交互に最適化を行うことで解くことができる。この問題を解く具体的なアルゴリズムとして様々なものが知られており、一例として交互方向乗数法（ADM）を採用したアルゴリズムがある（当該アルゴリズムについて後述する）。

[0070] 続いて、式(15)の問題定式化において、補助変数に課す確率的仮定として多様なものが使用可能であることを示すため、高い雑音抑圧性能を保ちつつ低遅延かつ音声の強調に適したフィルタを設計する例について説明する。

[0071] 《複数の特性を考慮したビームフォーマの具体的設計例》

ここでは問題設定として実践的な状況を仮定し、式(15)の枠組みでコスト関数を具体的に設計する例を示す。具体的には、複数の妨害音が鳴っている環境下で、既知の位置から発せられた音声をストリーミング配信する、という状況を仮定する。なお、妨害音源は周波数ビン毎に複素正規分布に従う雑音を発しているものとする。本状況では、目的音源が音声であるという情報を反映した上で、遅延が少なくかつ高い強調性能を保つようなビームフォーマが要望されるだろう。

[0072] [フィルタ設計フェーズ：フィルタ係数 $\tilde{w}^*$ のコスト項]

上記状況では特にフィルタ係数 $\tilde{w}^*$ におくべき制約は存在しないので、フィルタ係数 $\tilde{w}^*$ に対するコスト項 L_0 は考えないことにする。つまり、 $L_0(\tilde{w}^*)=0$ と

する。

[0073] [フィルタ設計フェーズ：フィルタ係数 $\tilde{w}^*$ の補助変数、補助変数のコスト項]

続いて、音源に関して既知の情報やビームフォーマに要望する特性の一つ一つについて検討し、補助変数とそのコスト項の設計を行っていく。

[0074] まず、目的音の分布を考える。上記仮定した状況では目的音は音声であるという情報が既知である。音声はスパース性という性質を持つことが知られているため、推定目的音がスパースな確率分布に従うという仮定を考慮したコスト項を設計することで、この既知の情報が活かせると考えられる。そこで、補助変数として推定目的音 $\tilde{y}_t$ を採用する。補助変数 $\tilde{y}_t$ の定義は式(6)と同一である。そして、補助変数 $\tilde{y}_t$ が、次式のラプラス分布に従うという仮定をおく。

[0075] [数13]

$$P(y_{f,t}) \propto \exp(-\beta |y_{f,t}|) \quad \cdots (16)$$

[0076] ここで $\beta (>0)$ は分布の形状を定める定数パラメータである。ラプラス分布はスパースな変数の分布の表現にしばしば用いられるものであり、上記仮定した状況では適切であると考えられる。式(16)の仮定のもと、補助変数 $\tilde{y}$ のコスト項 L_y はラプラス分布の対数の-1倍である

[0077] [数14]

$$L_y(\tilde{y}) = \sum_{t=1}^T \sum_{f=1}^F \beta |y_{f,t}| \quad \cdots (17)$$

[0078] という形になる。ラプラス分布はlog-concaveであるので、コスト項 L_y は凸関数になり、式(15)の枠組みのもとで扱える。

[0079] 次に、非目的音に関しても何らかの確率分布を仮定し、同様に補助変数・コスト項の導入を行う。観測音に含まれる非目的音の推定量として、式(7)で定義される推定非目的音 $\tilde{e}_t$ を補助変数として導入する。上記仮定した状況では、主に妨害音からなる非目的音は正規分布に従うと仮定する。つまり、補助変数 $\tilde{e}_t$ は

[0080] [数15]

$$P(e_{f,t}) \propto \exp(-e_{f,t}^H R_f^{-1} e_{f,t}) \quad \cdots (18)$$

[0081] という確率分布に従って出力されるとみなす。ここで、 R_f は非目的音に関する空間相関行列であり、観測データから見積もることができる。式(18)の仮定を、補助変数 $\tilde{e}$ に対するコスト項の形式に書き換えると、

[0082] [数16]

$$L_e(\tilde{e}) = \sum_{t=1}^T \sum_{f=1}^F e_{f,t}^H R_f^{-1} e_{f,t} \quad \cdots (19)$$

[0083] という形になる。正規分布はlog-concaveであるので、このコスト項 L_e も凸関数である。

[0084] ここで更なる補助変数とコスト項を導入することで、ビームフォーマに低遅延性を持たせることを目指す。そのためフィルタ係数 $\tilde{w}^*$ にどのようなコスト項を課せば低遅延なフィルタが実現可能か検討する。従来の広帯域ビームフォーマは周波数ビン毎に独立してフィルタ係数を導出しており、隣接周波数ビン間の関連性は考慮していなかった。しかし、周波数ビン方向に不連続やなめらかでないという周波数特性は、時間領域において裾の長いインパルス応答の原因となる。また、位相遅れを引き起こす群遅延も抑制されることが望ましい。このような特性を持たないフィルタ係数を解として得るために、フィルタ係数の周波数ビン方向に関する差分を新たな補助変数として導入し、これらの補助変数（のノルム）を小さくするようなコスト項を課するのが有効であると考えられる。具体的には、新たに

[0085] [数17]

$$\eta_f = w_f^* - 2w_{f+1}^* + w_{f+2}^* \quad (f=1, \dots, F-2) \quad \cdots (20)$$

[0086] という $F-2$ 個の補助変数 η_f を定義する。式(20)は、 η_f がフィルタの振幅・位相特性の周波数方向に関する2階微分の情報を含むように意図されている。式(20)を用いて、補助変数 η_f に対するコスト項 $L_{\eta_f}(\eta_f)$ を

[0087]

[数18]

$$L_{\eta_f}(\eta_f) = \lambda \|\eta_f\|_2 \quad \cdots (21)$$

[0088] で定義する。

[0089] [フィルタ設計フェーズ：コスト関数最適化]

以上のように、式(18)、式(16)、式(20)のような補助変数に関する仮定をおいたことで、コスト関数Lは、各コスト項の和

[0090] [数19]

$$L(\tilde{\mathbf{w}}^*, \hat{\mathbf{v}}) = \sum_{t=1}^T \sum_{f=1}^F (e_{f,t}^H R_f^{-1} e_{f,t} + \beta |y_{f,t}|) + \sum_{f=1}^{F-2} \lambda \|\eta_f\|_2 \quad \cdots (22)$$

$$\hat{\mathbf{v}} = [e_{1,1}, \cdots, e_{F,T}, y_{1,1}, \cdots, y_{F,T}, \eta_1, \cdots, \eta_{F-2}] \quad \cdots (23)$$

[0091] となる。このコスト関数Lに出現する $2FT+F-2$ 個の補助変数はみなフィルタ係数 $\tilde{\mathbf{w}}^*$ のアフィン変換として表されるので、式(22)の最小化問題は式(15)の具体例である。

[0092] ここまではビームフォーミングを対象として最適化問題の議論をしてきたが、これまで説明した数理的な枠組みはより汎用的な適用範囲を持つものであり、音響処理に限られるものではない。本枠組みの汎用性を端的に示すため、上記枠組みを画像処理に適用した例について説明する。

[0093] 《画像処理における最適化問題の一例》

例えば、同じ形状の物体が大量に写っている画像のように周期的な絵柄の画像（以下、元の画像という）にノイズが重畳された画像が入力として与えられ、当該画像からノイズを除去した画像を得たいという状況を考える。元の画像の各画素の値を表す行列を $\mathbf{S}=[S_{x,y}]_{1 \leq x \leq X, 1 \leq y \leq Y}$ 、各画素に加わるノイズを表す行列を $\mathbf{N}$ と表す。ノイズの値は画素毎に独立に平均0・分散1の正規分布に従って生成されるものとする。我々が観測できるのはノイズを含んだ画像 $\mathbf{Y}=\mathbf{S}+\mathbf{N}$ である。このとき、画像Yから元の画像Sを精度よく推定する問題を考えるため、行列Sを式(15)における $\tilde{\mathbf{w}}^*$ とみなし、行列Sや行列Sのアフィン変換によって定まる補助変数に関するコスト項を構成していく。

[0094] [フィルタ設計フェーズ：行列Sのコスト項]

まず、推定結果として得られた画像は元の画像と概ね一致してほしいので、行列Sに関するコスト項として、入力画像Yの各画素との二乗誤差を課すことにする。行列Sに関するコスト項を具体的に書くと、次式になる。

[0095] [数20]

$$L_S(S) = \sum_{x,y} |S_{x,y} - Y_{x,y}|^2 \quad \dots(24)$$

[0096] 式(24)のコスト項 L_S は凸関数である。

[0097] [フィルタ設計フェーズ：補助変数・補助変数のコスト項]

次に、適切にノイズを除去するための補助変数とそのコスト項を設計する。画像は通常滑らかで、隣接画素間の値の変動は小さいということを我々は経験的に知っている。画素毎に独立なノイズはこの性質に反した不自然な振る舞いを示すため、この不自然さを嫌うようなコスト項を設計することでノイズが除去できると考えられる。そこで、隣接画素間の差分として定義される量 D_1 , D_2 を次式で与えられる補助変数として導入する。

[0098] [数21]

$$D_{1x,y} = S_{x+1,y} - S_{x,y} \quad (1 \leq x \leq X-1, 1 \leq y \leq Y) \quad \dots(25)$$

$$D_{2x,y} = S_{x,y+1} - S_{x,y} \quad (1 \leq x \leq X, 1 \leq y \leq Y-1) \quad \dots(26)$$

[0099] そして、滑らかでより自然性の高い画像ならば補助変数 D_1 , D_2 の絶対値は傾向として小さくなるはずである。そこで、補助変数 D_1 , D_2 に対して次のような凸なコスト項を課す。

[0100] [数22]

$$L_{D1}(D_1) = \sum_{x,y} |D_{1x,y}| \quad \dots(27)$$

$$L_{D2}(D_2) = \sum_{x,y} |D_{2x,y}| \quad \dots(28)$$

[0101] これらのコスト項 L_{D1} , L_{D2} は、ノイズ除去の意味合いを持つコスト項である。

[0102] ここでは更に、元の画像が周期的な構造を有するという事前知識を持っている状況を仮定し、このような事前知識を活かせる補助変数やコスト項を設計する。周期的な画像では、画像を2次元フーリエ変換して得られた空間周波数スペクトルがスパースな構造を持つと期待される。2次元フーリエ変換はアフィン変換として記述できるので、空間周波数スペクトルを補助変数として採用することにし、これらの補助変数をスパースに誘導するようなコスト項を設計すれば我々の目的が達成されと考えられる。具体的には、画像の2次元フーリエ変換 $R=[R_{k,j}]$ を補助変数として導入する。これらは離散フーリエ変換行列 $W_{k,j}$ によって

[0103] [数23]

$$R_{k,j} = \sum_{x,y} W_{k,j,x,y} S_{x,y} \quad \cdots (29)$$

[0104] という式で定義でき、確かに行列 S のアフィン変換になっている。コスト項としては、

[0105] [数24]

$$L_R(R) = \sum_{k,j} |R_{k,j}| \quad \cdots (30)$$

[0106] という形の凸関数を仮定する。

[0107] [フィルタ設計フェーズ：コスト関数最適化]

以上のコスト項の設計により、最適化すべきコスト関数 L は

[0108] [数25]

$$L(S, D_1, D_2, R) = L_S(S) + L_{D1}(D_1) + L_{D2}(D_2) + L_R(R) \cdots (31)$$

[0109] という形になる。式(31)の変数のうち行列 S が推定対象の変数であり、その他の変数は行列 S の補助変数である。

[0110] 以上の議論から、画像処理の場合においても式(15)の枠組みでコスト関数が設計できることがわかる。

[0111] 《ADMMに基づく最適化アルゴリズム》

図1は、式(15)で表される線形制約付き凸最適化問題を実際に解くための

反復アルゴリズムを示す図である。当該アルゴリズムは式(15)の問題を効率的に解くアルゴリズムの一つとして知られるADMMに基づくものである。ADMMは元の問題の双対問題に対して最適化を行うようなアルゴリズムであり、補助変数 v_j と同じ次元を持つ双対変数 u_j を用いる。

[0112] 以下では、ビームフォーマの事例の一つの例として構成したコスト関数(22)に対して図1のアルゴリズムを適用した例について説明する。ここでは、図1の式より具体的な反復更新式を導出する。

[0113] まず、変数 $\tilde{w}^*$ の更新則に関してはコスト項 L_0 が式(22)では消えていることから、図1のステップ3の式は

[0114] [数26]

$$\tilde{w}^* \leftarrow (\hat{D}^H \hat{D})^{-1} \hat{D}^H (\hat{v} - \hat{u} + \hat{b}) \quad \dots(32)$$

[0115] という形式に帰着される。この式に現れる行列 $\hat{D}^H \hat{D} = \sum_j D_j^H D_j$ は

[0116] [数27]

$$\hat{D}^H \hat{D} = \begin{bmatrix} A_1 + I_M & -2I_M & I_M & \cdots & 0 \\ -2I_M & A_3 + 5I_M & -4I_M & \cdots & 0 \\ I_M & -4I_M & A_5 + 6I_M & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A_F + I_M \end{bmatrix} \dots(33)$$

$$A_f = (1 + \|h_f\|_2^2) \sum_t z_{f,t}^* z_{f,t}^T \quad \dots(34)$$

[0117] という形式のブロック帯行列となるため、行列 $\hat{D}^H \hat{D}$ のコレスキー分解を行うことにより更新の際必要になる $(\hat{D}^H \hat{D})^{-1}$ による乗算の計算を効率化できる。

[0118] 続いて、補助変数の更新則を求める。この更新則は各コスト項の近接作用素として記述される。ここで、関数 f の近接作用素 prox_f は $\text{prox}_f(x) = \arg \min_y f(y) + \frac{1}{2} \|x - y\|_2^2$ という形で定義される。この形とコスト項を見比べると、補助変数 $y_{f,t}$ と補助変数 η_f の更新則は l^2 ノルムの近接作用素

[0119] [数28]

$$\text{prox}_{\lambda \|\cdot\|_2}(z) = \max(\|z\|_2 - \lambda, 0) \frac{z}{\|z\|_2} \quad \dots(35)$$

[0120] で表されることがわかる。一方、補助変数 $e_{f,t}$ に関するコスト項は単純な二次形式であるので、 $e_{f,t}$ の更新式は定義から解析的に容易に導ける。結局、補助変数の更新則は

[0121] [数29]

$$\begin{aligned} e_{f,t} &\leftarrow \frac{\gamma}{2} R_f (I + \frac{\gamma}{2} R_f)^{-1} (z_{f,t} - (w_f^H z_{f,t}) h_f + u_{e,f,t}) \\ y_{f,t} &\leftarrow \max \left(1 - \frac{\beta}{\gamma \|w_f^H z_{f,t} + u_{y,f,t}\|_2}, 0 \right) (w_f^H z_{f,t} + u_{y,f,t}) \quad \dots(36) \\ \eta_f &\leftarrow \max \left(1 - \frac{\lambda}{\gamma \|w_f^* - 2w_{f+1}^* + w_{f+2}^* + u_{\eta,f}\|_2}, 0 \right) (w_f^* - 2w_{f+1}^* + w_{f+2}^* + u_{\eta,f}) \end{aligned}$$

[0122] という形になる。ここで、 I は単位行列を表す。

[0123] 《効果》

本発明の実施形態の原理は、ビームフォーマのフィルタ係数導出をコスト関数最適化問題として解釈した上で、フィルタ係数やその補助変数に対して個別のコスト項に基づく制約を課すことで、所望する複数の特性を兼ね備えたビームフォーマを設計するものである。

[0124] 従来法では、事前知識や所望する特性などの様々な要因を考慮した複雑なコスト関数を使用した設計を行うことができなかった。一方、本発明の実施形態の原理に従えば、補助変数という形で新たな変数を複数導入し、これらに関して個別にコスト項を設計するという枠組みでコスト関数を構成する。各コスト項は確率的仮定であるという意味を持っており、特にlog-concaveな確率的仮定を課す場合、線形制約付き凸最適化問題に帰着され、様々な数理的手法で比較的容易に最適化問題を解くことができる。これにより、複数の仮定を同時に考慮したフィルタ設計が可能となる。

[0125] <第1実施形態>

以下、図2～図3を参照して潜在変数最適化装置100を説明する。図2は、潜在変数最適化装置100の構成を示すブロック図である。図3は、潜在変数最適化装置100の動作を示すフローチャートである。図2に示すよ

うに潜在変数最適化装置 100 は、セットアップデータ計算部 110 と、最適化部 120 と、記録部 190 を含む。記録部 190 は、潜在変数最適化装置 100 の処理に必要な情報を適宜記録する構成部である。記録部 190 は、例えば、最適化対象となる潜在変数を記録する。

[0126] 潜在変数最適化装置 100 は、最適化用データを用いて、最適化の対象となるモデルの潜在変数 $\tilde{w}^*$ を最適化する。ここで、モデルとは、入力データを入力とし、出力データを出力とする関数（例えば、観測音を入力データとし、目的音を出力データとするビームフォーマのフィルタ）のことであり、最適化用データとは、潜在変数の最適化に用いる入力データ、または、潜在変数の最適化に用いる入力データと出力データの組のことをいう。

[0127] 図 3 に従い潜在変数最適化装置 100 の動作について説明する。

[0128] S 110 において、セットアップデータ計算部 110 は、最適化用データを用いて、潜在変数 $\tilde{w}^*$ を最適化する際に用いるセットアップデータを計算する。例えば、潜在変数 $\tilde{w}^*$ を最適化するために用いるコスト関数 L

[0129] [数30]

$$L(\tilde{w}^*, v_1, \dots, v_J) = L_0(\tilde{w}^*) + \sum_{j=1}^J L_j(v_j) \quad \dots (*)$$

[0130] （ただし、 $v_j (1 \leq j \leq J)$ は行列 D_j とベクトル b_j を用いて $v_j = D_j \tilde{w}^* + b_j$ と表される潜在変数 $\tilde{w}^*$ の補助変数、 L_0 は潜在変数 $\tilde{w}^*$ のコスト項、 $L_j (1 \leq j \leq J)$ は補助変数 v_j のコスト項）で用いる $D_j (1 \leq j \leq J)$ 、 $b_j (1 \leq j \leq J)$ 、コスト項 $L_i (0 \leq i \leq J)$ に含まれるパラメータがセットアップデータの一例である。なお、コスト項 $L_i (0 \leq i \leq J)$ は凸関数とするのが好ましい。

[0131] 例えば、補助変数 $v_j (1 \leq j \leq J)$ のコスト項は、log-concaveである補助変数 v_j の確率分布を用いて表されるものとする、コスト項 $L_i (1 \leq i \leq J)$ は凸関数となる。また、例えば、潜在変数 $\tilde{w}^*$ のコスト項 $L_0 = 0$ としてもよく、コスト関数 L は補助変数 $v_j (1 \leq j \leq J)$ のコスト項の和を含むものであればよい。

[0132] S 120 において、最適化部 120 は、コスト関数 L の最小化問題を解くことにより潜在変数 $\tilde{w}^*$ を最適化する。以下、図 4 ～図 5 を参照して最適化部 1

20について説明する。図4は、最適化部120の構成を示すブロック図である。図5は、最適化部120の動作を示すフローチャートである。図4に示すように最適化部120は、初期化部121、潜在変数更新部122と、補助変数更新部123と、双対変数更新部124と、カウンタ更新部125と、終了条件判定部126を含む。

[0133] 図5に従い最適化部120の動作について説明する。

[0134] S121において、初期化部121は、カウンタ n を初期化する。具体的には、 $n=1$ とする。また、初期化部121は、補助変数 $\hat{v}=(v_1^T, \dots, v_J^T)^T$ 、双対変数 $\hat{u}=(u_1^T, \dots, u_J^T)^T$ を初期化する。さらに、初期化部121は、 γ にも初期値となる定数を設定する。

[0135] S122において、潜在変数更新部122は、現時点で得られている補助変数 $\hat{v}$ 、双対変数 $\hat{u}$ の値を用いて、次式により、潜在変数 $\tilde{w}^*$ を更新する。

[0136] [数31]

$$\tilde{w}^* \leftarrow \arg \min_{\tilde{w}^*} L_0(\tilde{w}^*) + \frac{\gamma}{2} \left\| \hat{D} \tilde{w}^* - \hat{v} + \hat{u} + \hat{b} \right\|_2^2$$

[0137] ここで、 $\hat{D}=(D_1^T, \dots, D_J^T)^T$ 、 $\hat{b}=(b_1^T, \dots, b_J^T)^T$ である。

[0138] S123において、補助変数更新部123は、現時点で得られている潜在変数 $\tilde{w}^*$ 、双対変数 u_j の値を用いて、次式により、補助変数 $v_j (1 \leq j \leq J)$ を更新する。

[0139] [数32]

$$v_j \leftarrow \arg \min_{v_j} \frac{1}{\gamma} L_j(v_j) + \frac{1}{2} \left\| D_j \tilde{w}^* + u_j + b_j - v_j \right\|_2^2$$

[0140] S124において、双対変数更新部124は、現時点で得られている潜在変数 $\tilde{w}^*$ 、補助変数 v_j 、双対変数 u_j の値を用いて、次式により、双対変数 $u_j (1 \leq j \leq J)$ を更新する。

[0141] [数33]

$$u_j \leftarrow u_j + D_j \tilde{w}^* - v_j + b_j$$

[0142] S125において、カウンタ更新部125は、カウンタ n を1だけインクリ

メントする。具体的には、 $n \leftarrow n+1$ とする。

[0143] S 1 2 6において、終了条件判定部 1 2 6は、カウンタ n が所定の更新回数 $N_{\text{iteration}}$ ($N_{\text{iteration}}$ は1以上の整数であり、例えば10万)に達した場合（つまり、 $n > N_{\text{iteration}}$ となり、終了条件が満たされた場合）は、そのときの潜在変数の値 $\tilde{w}^*$ を出力して、処理を終了する。それ以外の場合、S 1 2 2の処理に戻る。つまり、最適化部 1 2 0は、S 1 2 2～S 1 2 6の処理を繰り返す。

[0144] なお、式(*)で定義されるコスト関数 L を用いる場合、 J が2以上であっても最適化することができる。

[0145] また、式(*)のように、コスト関数 L は、補助変数 v_j のコスト項をlog-concaveである補助変数 v_j の確率分布を用いて表されるものとして、補助変数 v_j のコスト項の和を含むものであればよく、例えば、コスト関数 L は、補助変数 v_j のコスト項をlog-concaveである補助変数 v_j の確率分布を用いて表されるものとして、補助変数 v_j のコスト項とlog-concaveな確率分布に基づいて定まるコスト項との和として表されるものであってもよい。

[0146] 本実施形態の発明によれば、潜在変数及び潜在変数から定まる補助変数に対する確率的仮定に基づくコスト関数を用いて潜在変数を最適化することが可能となる。

[0147] [適用例]

ここでは、潜在変数最適化装置 1 0 0を音源強調に用いるビームフォーマのフィルタ係数の最適化に適用した例について説明する。そこで、以下では潜在変数最適化装置 1 0 0のことをフィルタ係数最適化装置 1 0 0と呼ぶことにする。フィルタ係数最適化装置 1 0 0の最適化対象はビームフォーマのフィルタ係数となる。フィルタ係数最適化装置 1 0 0の構成は図 2 に示す通りである。

[0148] 以下、図 3 に従いフィルタ係数最適化装置 1 0 0の動作について説明する。

[0149] S 1 1 0において、セットアップデータ計算部 1 1 0は、最適化用データを用いて、フィルタ係数 $\tilde{w}^* = (w_1^{*T}, \dots, w_F^{*T})$ （ただし、 $w_f^* (1 \leq f \leq F)$ は周波数

ビン f のフィルタ係数)を最適化する際に用いるセットアップデータを計算する。例えば、フィルタ係数 $\tilde{w}^*$ を最適化するために用いるコスト関数 L

[0150] [数34]

$$L(\tilde{w}^*, \hat{v}) = \sum_{t=1}^T \sum_{f=1}^F (e_{f,t}^H R_f^{-1} e_{f,t} + \beta |y_{f,t}|) + \sum_{f=1}^{F-2} \lambda \|\eta_f\|_2 \cdots (**)$$

$$\hat{v} = [e_{1,1}, \cdots, e_{F,T}, y_{1,1}, \cdots, y_{F,T}, \eta_1, \cdots, \eta_{F-2}]$$

[0151] (ただし、 $e_{f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$)は時間フレーム t における周波数ビン f の推定非目的音を表すフィルタ係数 $\tilde{w}^*$ の補助変数、 $y_{f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$)は時間フレーム t における周波数ビン f の推定目的音を表すフィルタ係数 $\tilde{w}^*$ の補助変数、 η_f ($1 \leq f \leq F-2$)は $\eta_f = w_f^* - 2w_{f+1}^* + w_{f+2}^*$ で定義されるフィルタ係数 $\tilde{w}^*$ の補助変数、 R_f ($1 \leq f \leq F$)は周波数ビン f の非目的音に関する空間相関行列、 β (>0)は所定の定数、 λ は所定の定数)で用いるコスト項 $L_{e,f,t}(e_{f,t}) = e_{f,t}^H R_f^{-1} e_{f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$)、コスト項 $L_{y,f,t}(y_{f,t}) = \beta |y_{f,t}|$ ($1 \leq f \leq F$, $1 \leq t \leq T$)、コスト項 $L_{\eta,f}(\eta_f) = \lambda \|\eta_f\|_2$ ($1 \leq f \leq F-2$)に含まれるパラメータがセットアップデータの一例である。なお、コスト項 $L_{e,f,t}(e_{f,t})$ ($1 \leq f \leq F$, $1 \leq t \leq T$)、 $L_{y,f,t}(y_{f,t})$ ($1 \leq f \leq F$, $1 \leq t \leq T$)、 $L_{\eta,f}(\eta_f)$ ($1 \leq f \leq F-2$)はいずれも凸関数である。

[0152] ただし、コスト項 $L_{e,f,t}(e_{f,t})$, $L_{y,f,t}(y_{f,t})$, $L_{\eta,f}(\eta_f)$ は上記コスト項に限るものではなく、例えば、補助変数 $e_{f,t}$, $y_{f,t}$, η_f のコスト項は、それぞれlog-concaveである補助変数 $e_{f,t}$, $y_{f,t}$, η_f の確率分布を用いて表されるものであればよい。

[0153] なお、上記コスト関数 L の定義式では、フィルタ係数 $\tilde{w}^*$ のコスト項 $L_0=0$ となっている。

[0154] S120において、最適化部120は、コスト関数 L の最小化問題を解くことによりフィルタ係数 $\tilde{w}^*$ を最適化する。以下、図4～図5を参照して最適化部120について説明する。図4は、最適化部120の構成を示すブロック図である。図5は、最適化部120の動作を示すフローチャートである。ここでは、最適化部120に含まれる潜在変数更新部122のことをフィルタ係数更新部122と呼ぶことにする。

[0155] 以下、図5に従い最適化部120の動作について説明する。

[0156] S121において、初期化部121は、カウンタ n を初期化する。具体的には、 $n=1$ とする。また、初期化部121は、補助変数 $\hat{v}=[e_{1,1}, \dots, e_{F,T}, y_{1,1}, \dots, y_{F,T}, \eta_1, \dots, \eta_{F-2}]$ 、双対変数 $\hat{u}=[u_{e,1,1}, \dots, u_{e,F,T}, u_{y,1,1}, \dots, u_{y,F,T}, u_{\eta_1,1}, \dots, u_{\eta_{F-2},1}]$ （ただし、 $u_{e,f,t}$ ($1 \leq f \leq F, 1 \leq t \leq T$) は補助変数 $e_{f,t}$ の双対変数、 $u_{y,f,t}$ ($1 \leq f \leq F, 1 \leq t \leq T$) は補助変数 $y_{f,t}$ の双対変数、 u_{η_f} ($1 \leq f \leq F-2$) は補助変数 η_f の双対変数とする）を初期化する。さらに、初期化部121は、 γ にも初期値となる定数を設定する。

[0157] S122において、フィルタ係数更新部122は、現時点で得られている補助変数 $\hat{v}$ 、双対変数 $\hat{u}$ の値を用いて、次式により、フィルタ係数 $\tilde{w}^*$ を更新する。

[0158] [数35]

$$\tilde{w}^* \leftarrow (\hat{D}^H \hat{D})^{-1} \hat{D}^H (\hat{v} - \hat{u} - \hat{b})$$

[0159] ここで、 $\hat{D}$ 、 $\hat{b}$ は次式により与えられる。

[0160] [数36]

$$\hat{D} = \begin{bmatrix} -h_1 z_{1,1}^T & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -h_1 z_{1,T}^T & 0 & 0 & \dots & 0 \\ 0 & -h_2 z_{2,1}^T & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -h_F z_{F,T}^T \\ z_{1,1}^T & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ z_{1,T}^T & 0 & 0 & \dots & 0 \\ 0 & z_{2,1}^T & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & z_{F,T}^T \\ 1 & -2 & 1 & \dots & 0 \\ 0 & 1 & -2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix},$$

$$\hat{b} = [z_{1,1}^T, \dots, z_{1,T}^T, z_{2,1}^T, \dots, z_{F,T}^T, 0, \dots, 0]^T$$

[0161] S 1 2 3において、補助変数更新部 1 2 3は、現時点で得られている潜在変数 $\tilde{w}^*$ 、双対変数 $u_{e,f,t}$, $u_{y,f,t}$, $u_{\eta,f}$ の値を用いて、次式により、補助変数 $e_{f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$)、補助変数 $y_{f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$)、補助変数 η_f ($1 \leq f \leq F-2$)を更新する。

[0162] [数37]

$$e_{f,t} \leftarrow \frac{\gamma}{2} R_f (I + \frac{\gamma}{2} R_f)^{-1} (z_{f,t} - (w_f^H z_{f,t}) h_f + u_{e,f,t})$$

$$y_{f,t} \leftarrow \max \left(1 - \frac{\beta}{\gamma \|w_f^H z_{f,t} + u_{y,f,t}\|}, 0 \right) (w_f^H z_{f,t} + u_{y,f,t})$$

$$\eta_f \leftarrow \max \left(1 - \frac{\lambda}{\gamma \|w_f^* - 2w_{f+1}^* + w_{f+2}^* + u_{\eta,f}\|}, 0 \right) (w_f^* - 2w_{f+1}^* + w_{f+2}^* + u_{\eta,f})$$

[0163] (ただし、 $z_{f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$)は時間フレーム t における周波数ビン f の観測音、 h_f ($1 \leq f \leq F$)は周波数ビン f におけるビーム方向のアレイマニフォールドベクトル)

[0164] S 1 2 4において、双対変数更新部 1 2 4は、現時点で得られている潜在変数 $\tilde{w}^*$ 、補助変数 $e_{f,t}$, $y_{f,t}$, η_f の値を用いて、次式により、双対変数 $u_{e,f,t}$, $u_{y,f,t}$, $u_{\eta,f}$ を更新する。

[0165] [数38]

$$u_{e,f,t} \leftarrow u_{e,f,t} + (z_{f,t} - h_f (w_f^H z_{f,t})) - e_{f,t}$$

$$u_{y,f,t} \leftarrow u_{y,f,t} + w_f^H z_{f,t} - y_{f,t}$$

$$u_{\eta,f} \leftarrow u_{\eta,f} + (w_{f+2}^* - 2w_{f+1}^* + w_f^*) - \eta_f$$

[0166] S 1 2 5において、カウンタ更新部 1 2 5は、カウンタ n を1だけインクリメントする。具体的には、 $n \leftarrow n+1$ とする。

[0167] S 1 2 6において、終了条件判定部 1 2 6は、カウンタ n が所定の更新回数 $N_{\text{iteration}}$ ($N_{\text{iteration}}$ は1以上の整数であり、例えば10万)に達した場合 (つまり、 $n > N_{\text{iteration}}$ となり、終了条件が満たされた場合) は、そのときのフィルタ係数の値 $\tilde{w}^*$ を出力して、処理を終了する。それ以外の場合、S 1 2 2の処理に戻る

る。つまり、最適化部120は、S122～S126の処理を繰り返す。

[0168] なお、式(**)のように、コスト関数 L は、補助変数 $e_{f,t}$, $y_{f,t}$, η_f のコスト項をそれぞれlog-concaveである補助変数 $e_{f,t}$, $y_{f,t}$, η_f の確率分布を用いて表されるものとして、補助変数 $e_{f,t}$, $y_{f,t}$, η_f のコスト項の和を含むものであればよく、例えば、コスト関数 L は、補助変数 $e_{f,t}$, $y_{f,t}$, η_f のコスト項をそれぞれlog-concaveである補助変数 $e_{f,t}$, $y_{f,t}$, η_f の確率分布を用いて表されるものとして、補助変数 $e_{f,t}$, $y_{f,t}$, η_f のコスト項とlog-concaveな確率分布に基づいて定まるコスト項との和として表されるものであってもよい。

[0169] <補記>

本発明の装置は、例えば単一のハードウェアエンティティとして、キーボードなどが接続可能な入力部、液晶ディスプレイなどが接続可能な出力部、ハードウェアエンティティの外部に通信可能な通信装置（例えば通信ケーブル）が接続可能な通信部、CPU（Central Processing Unit、キャッシュメモリやレジスタなどを備えていてもよい）、メモリであるRAMやROM、ハードディスクである外部記憶装置並びにこれらの入力部、出力部、通信部、CPU、RAM、ROM、外部記憶装置の間のデータのやり取りが可能なように接続するバスを有している。また必要に応じて、ハードウェアエンティティに、CD-ROMなどの記録媒体を読み書きできる装置（ドライブ）などを設けることとしてもよい。このようなハードウェア資源を備えた物理的実体としては、汎用コンピュータなどがある。

[0170] ハードウェアエンティティの外部記憶装置には、上述の機能を実現するために必要となるプログラムおよびこのプログラムの処理において必要となるデータなどが記憶されている（外部記憶装置に限らず、例えばプログラムを読み出し専用記憶装置であるROMに記憶しておくこととしてもよい）。また、これらのプログラムの処理によって得られるデータなどは、RAMや外部記憶装置などに適宜に記憶される。

[0171] ハードウェアエンティティでは、外部記憶装置（あるいはROMなど）に記憶された各プログラムとこの各プログラムの処理に必要なデータが必要に

応じてメモリに読み込まれて、適宜にCPUで解釈実行・処理される。その結果、CPUが所定の機能（上記、…部、…手段などと表した各構成要件）を実現する。

[0172] 本発明は上述の実施形態に限定されるものではなく、本発明の趣旨を逸脱しない範囲で適宜変更が可能である。また、上記実施形態において説明した処理は、記載の順に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的にあるいは個別に実行されるとしてもよい。

[0173] 既述のように、上記実施形態において説明したハードウェアエンティティ（本発明の装置）における処理機能をコンピュータによって実現する場合、ハードウェアエンティティが有すべき機能の処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、上記ハードウェアエンティティにおける処理機能がコンピュータ上で実現される。

[0174] この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよい。具体的には、例えば、磁気記録装置として、ハードディスク装置、フレキシブルディスク、磁気テープ等を、光ディスクとして、DVD（Digital Versatile Disc）、DVD-RAM（Random Access Memory）、CD-ROM（Compact Disc Read Only Memory）、CD-R（Recordable）／RW（ReWritable）等を、光磁気記録媒体として、MO（Magneto-Optical disc）等を、半導体メモリとしてEEPROM（Electrically Erasable and Programmable-Read Only Memory）等を用いることができる。

[0175] また、このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納し

ておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

[0176] このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶装置に格納する。そして、処理の実行時、このコンピュータは、自己の記憶装置に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。また、このプログラムの別の実行形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP (Application Service Provider) 型のサービスによって、上述の処理を実行する構成としてもよい。なお、本形態におけるプログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの（コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等）を含むものとする。

[0177] また、この形態では、コンピュータ上で所定のプログラムを実行させることにより、ハードウェアエンティティを構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。

[0178] 上述の本発明の実施形態の記載は、例証と記載の目的で提示されたものである。網羅的であるという意思はなく、開示された厳密な形式に発明を限定する意思もない。変形やバリエーションは上述の教示から可能である。実施形態は、本発明の原理の最も良い例証を提供するために、そして、この分野の当業者が、熟考された実際の使用に適するように本発明を色々な実施形態で、また、色々な変形を付加して利用できるようにするために、選ばれて表

現されたものである。すべてのそのような変形やバリエーションは、公正に合法的に公平に与えられる幅にしたがって解釈された添付の請求項によって定められた本発明のスコープ内である。

請求の範囲

[請求項1] 潜在変数 $\tilde{w}^*$ を最適化する最適化部を含む潜在変数最適化装置であって、

$v_j (1 \leq j \leq J)$ を行列 D_j とベクトル b_j を用いて $v_j = D_j \tilde{w}^* + b_j$ と表される潜在変数 $\tilde{w}^*$ の補助変数とし、

補助変数 v_j のコスト項は、log-concaveである補助変数 v_j の確率分布を用いて表されるものであり、

前記最適化部は、

補助変数 v_j のコスト項の和を含むコスト関数の最小化問題を解くことにより潜在変数 $\tilde{w}^*$ を最適化する

潜在変数最適化装置。

[請求項2] 請求項1に記載の潜在変数最適化装置であって、

前記コスト関数は

[数39]

$$L(\tilde{w}^*, v_1, \dots, v_J) = L_0(\tilde{w}^*) + \sum_{j=1}^J L_j(v_j)$$

(ただし、 L_0 は潜在変数 $\tilde{w}^*$ のコスト項、 $L_j (1 \leq j \leq J)$ は補助変数 v_j のコスト項)で定義される

潜在変数最適化装置。

[請求項3] 請求項2に記載の潜在変数最適化装置であって、

$u_j (1 \leq j \leq J)$ を補助変数 v_j の双対変数、 $\hat{v} = (v_1^T, \dots, v_J^T)^T$ 、 $\hat{D} = (D_1^T, \dots, D_J^T)^T$ 、 $\hat{b} = (b_1^T, \dots, b_J^T)^T$ 、 $\hat{u} = (u_1^T, \dots, u_J^T)^T$ 、 γ を所定の定数とし、

前記最適化部は、

次式により、潜在変数 $\tilde{w}^*$ を更新する潜在変数更新部と、

[数40]

$$\tilde{w}^* \leftarrow \arg \min_{\tilde{w}^*} L_0(\tilde{w}^*) + \frac{\gamma}{2} \left\| \hat{D} \tilde{w}^* - \hat{v} + \hat{u} + \hat{b} \right\|_2^2$$

次式により、補助変数 v_j ($1 \leq j \leq J$)を更新する補助変数更新部と、
[数41]

$$v_j \leftarrow \arg \min_{v_j} \frac{1}{\gamma} L_j(v_j) + \frac{1}{2} \|D_j \tilde{w}^* + u_j + b_j - v_j\|_2^2$$

次式により、双対変数 u_j ($1 \leq j \leq J$)を更新する双対変数更新部と
[数42]

$$u_j \leftarrow u_j + D_j \tilde{w}^* - v_j + b_j$$

を含む潜在変数最適化装置。

[請求項4]

ビームフォーマのフィルタ係数 $\tilde{w}^* = (w_1^{*T}, \dots, w_F^{*T})$ (ただし、 w_f^* ($1 \leq f \leq F$)は周波数ビン f のフィルタ係数) を最適化する最適化部を含む
フィルタ係数最適化装置であって、

$e_{f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$)を時間フレーム t における周波数ビン f の推定非目的音を表すフィルタ係数 $\tilde{w}^*$ の補助変数、 $y_{f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$)を時間フレーム t における周波数ビン f の推定目的音を表すフィルタ係数 $\tilde{w}^*$ の補助変数、 η_f ($1 \leq f \leq F-2$)を $\eta_f = w_f^* - 2w_{f+1}^* + w_{f+2}^*$ で定義されるフィルタ係数 $\tilde{w}^*$ の補助変数とし、

補助変数 $e_{f,t}$, $y_{f,t}$, η_f のコスト項は、それぞれlog-concaveである
補助変数 $e_{f,t}$, $y_{f,t}$, η_f の確率分布を用いて表されるものであり、

前記最適化部は、

補助変数 $e_{f,t}$, $y_{f,t}$, η_f のコスト項の和を含むコスト関数の最小化問題を解くことによりフィルタ係数 $\tilde{w}^*$ を最適化する

フィルタ係数最適化装置。

[請求項5]

請求項4に記載のフィルタ係数最適化装置であって、

前記コスト関数は

[数43]

$$L(\tilde{w}^*, \hat{v}) = \sum_{t=1}^T \sum_{f=1}^F (e_{f,t}^H R_f^{-1} e_{f,t} + \beta |y_{f,t}|) + \sum_{f=1}^{F-2} \lambda \|\eta_f\|_2$$

$$\hat{v} = [e_{1,1}, \dots, e_{F,T}, y_{1,1}, \dots, y_{F,T}, \eta_1, \dots, \eta_{F-2}]$$

(ただし、 $R_f (1 \leq f \leq F)$ は周波数ビン f の非目的音に関する空間相関行列、 $\beta (>0)$ は所定の定数、 λ は所定の定数)で定義される

フィルタ係数最適化装置。

[請求項6]

請求項5に記載のフィルタ係数最適化装置であって、

$u_{e,f,t} (1 \leq f \leq F, 1 \leq t \leq T)$ を補助変数 $e_{f,t}$ の双対変数、 $u_{y,f,t} (1 \leq f \leq F, 1 \leq t \leq T)$ を補助変数 $y_{f,t}$ の双対変数、 $u_{\eta,f} (1 \leq f \leq F-2)$ を補助変数 η_f の双対変数、 $\hat{u} = [u_{e,1,1}, \dots, u_{e,F,T}, u_{y,1,1}, \dots, u_{y,F,T}, u_{\eta,1}, \dots, u_{\eta,F-2}]$ 、 γ を所定の定数とし、

前記最適化部は、

次式により、フィルタ係数 $\hat{w}^*$ を更新するフィルタ係数更新部と、

[数44]

$$\hat{w}^* \leftarrow (\hat{D}^H \hat{D})^{-1} \hat{D}^H (\hat{v} - \hat{u} - \hat{b})$$

(ただし、 $\hat{D}$ 、 $\hat{b}$ は次式により与えられ、

[数45]

$$\hat{D} = \begin{bmatrix} -h_1 z_{1,1}^T & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -h_1 z_{1,T}^T & 0 & 0 & \dots & 0 \\ 0 & -h_2 z_{2,1}^T & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -h_F z_{F,T}^T \\ z_{1,1}^T & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ z_{1,T}^T & 0 & 0 & \dots & 0 \\ 0 & z_{2,1}^T & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & z_{F,T}^T \\ 1 & -2 & 1 & \dots & 0 \\ 0 & 1 & -2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix},$$

$$\hat{b} = [z_{1,1}^T, \dots, z_{1,T}^T, z_{2,1}^T, \dots, z_{F,T}^T, 0, \dots, 0]^T$$

$z_{f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$) は時間フレーム t における周波数ビン f の観測音、 h_f ($1 \leq f \leq F$) は周波数ビン f におけるビーム方向のアレイマニフォールドベクトルである。))

次式により、補助変数 $e_{f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$)、補助変数 $y_{f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$)、補助変数 η_f ($1 \leq f \leq F-2$) を更新する補助変数更新部と、

[数46]

$$\begin{aligned} e_{f,t} &\leftarrow \frac{\gamma}{2} R_f (I + \frac{\gamma}{2} R_f)^{-1} (z_{f,t} - (w_f^H z_{f,t}) h_f + u_{e,f,t}) \\ y_{f,t} &\leftarrow \max \left(1 - \frac{\beta}{\gamma \|w_f^H z_{f,t} + u_{y,f,t}\|_2}, 0 \right) (w_f^H z_{f,t} + u_{y,f,t}) \\ \eta_f &\leftarrow \max \left(1 - \frac{\lambda}{\gamma \|w_f^* - 2w_{f+1}^* + w_{f+2}^* + u_{\eta,f}\|_2}, 0 \right) (w_f^* - 2w_{f+1}^* + w_{f+2}^* + u_{\eta,f}) \end{aligned}$$

次式により、双対変数 $u_{e,f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$)、双対変数 $u_{y,f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$)、双対変数 $u_{\eta,f}$ ($1 \leq f \leq F-2$) を更新する双対変数更新部と

[数47]

$$\begin{aligned} u_{e,f,t} &\leftarrow u_{e,f,t} + (z_{f,t} - h_f (w_f^H z_{f,t})) - e_{f,t} \\ u_{y,f,t} &\leftarrow u_{y,f,t} + w_f^H z_{f,t} - y_{f,t} \\ u_{\eta,f} &\leftarrow u_{\eta,f} + (w_{f+2}^* - 2w_{f+1}^* + w_f^*) - \eta_f \end{aligned}$$

を含むフィルタ係数最適化装置。

[請求項7]

ビームフォーマのフィルタ係数 $\tilde{w}^* = (w_1^{*T}, \dots, w_F^{*T})$ (ただし、 w_f^* ($1 \leq f \leq F$) は周波数ビン f のフィルタ係数) を最適化する最適化部を含むフィルタ係数最適化装置であって、

η_f ($1 \leq f \leq F-2$) を $\eta_f = w_f^* - 2w_{f+1}^* + w_{f+2}^*$ で定義されるフィルタ係数 $\tilde{w}^*$ の補助変数とし、

補助変数 η_f のコスト項は、log-concaveである補助変数 η_f の確率分

布を用いて表されるものであり、

前記最適化部は、

補助変数 η_f のコスト項の和を含むコスト関数の最小化問題を解くことによりフィルタ係数 $\tilde{w}^*$ を最適化する
フィルタ係数最適化装置。

[請求項8] 潜在変数最適化装置が、潜在変数 $\tilde{w}^*$ を最適化する最適化ステップを実行する潜在変数最適化方法であって、

$v_j (1 \leq j \leq J)$ を行列 D_j とベクトル b_j を用いて $v_j = D_j \tilde{w}^* + b_j$ と表される潜在変数 $\tilde{w}^*$ の補助変数とし、

補助変数 v_j のコスト項は、log-concaveである補助変数 v_j の確率分布を用いて表されるものであり、

前記最適化ステップは、

補助変数 v_j のコスト項の和を含むコスト関数の最小化問題を解くことにより潜在変数 $\tilde{w}^*$ を最適化する
潜在変数最適化方法。

[請求項9] 請求項8に記載の潜在変数最適化方法であって、

前記コスト関数は

[数48]

$$L(\tilde{w}^*, v_1, \dots, v_J) = L_0(\tilde{w}^*) + \sum_{j=1}^J L_j(v_j)$$

(ただし、 L_0 は潜在変数 $\tilde{w}^*$ のコスト項、 $L_j (1 \leq j \leq J)$ は補助変数 v_j のコスト項) で定義される

潜在変数最適化方法。

[請求項10] フィルタ係数最適化装置が、ビームフォーマのフィルタ係数 $\tilde{w}^* = (w_1^*, \dots, w_F^*)$ (ただし、 $w_f^* (1 \leq f \leq F)$ は周波数ビン f のフィルタ係数) を最適化する最適化ステップを実行するフィルタ係数最適化方法であって、

$e_{f,t} (1 \leq f \leq F, 1 \leq t \leq T)$ を時間フレーム t における周波数ビン f の推

定非目的音を表すフィルタ係数 $\tilde{w}^*$ の補助変数、 $y_{f,t}$ ($1 \leq f \leq F$, $1 \leq t \leq T$) を時間フレーム t における周波数ビン f の推定目的音を表すフィルタ係数 $\tilde{w}^*$ の補助変数、 η_f ($1 \leq f \leq F-2$) を $\eta_f = w_f^* - 2w_{f+1}^* + w_{f+2}^*$ で定義されるフィルタ係数 $\tilde{w}^*$ の補助変数とし、

補助変数 $e_{f,t}$, $y_{f,t}$, η_f のコスト項は、それぞれlog-concaveである補助変数 $e_{f,t}$, $y_{f,t}$, η_f の確率分布を用いて表されるものであり、

前記最適化ステップは、

補助変数 $e_{f,t}$, $y_{f,t}$, η_f のコスト項の和を含むコスト関数の最小化問題を解くことによりフィルタ係数 $\tilde{w}^*$ を最適化する

フィルタ係数最適化方法。

[請求項11] 請求項10に記載のフィルタ係数最適化方法であって、
前記コスト関数は

[数49]

$$L(\tilde{w}^*, \hat{v}) = \sum_{t=1}^T \sum_{f=1}^F (e_{f,t}^H R_f^{-1} e_{f,t} + \beta |y_{f,t}|) + \sum_{f=1}^{F-2} \lambda \|\eta_f\|_2$$

$$\hat{v} = [e_{1,1}, \dots, e_{F,T}, y_{1,1}, \dots, y_{F,T}, \eta_1, \dots, \eta_{F-2}]$$

(ただし、 R_f ($1 \leq f \leq F$) は周波数ビン f の非目的音に関する空間相関行列、 β (>0) は所定の定数、 λ は所定の定数) で定義される

フィルタ係数最適化方法。

[請求項12] フィルタ係数最適化装置が、ビームフォーマのフィルタ係数 $\tilde{w}^* = (w_1^{*T}, \dots, w_F^{*T})$ (ただし、 w_f^* ($1 \leq f \leq F$) は周波数ビン f のフィルタ係数) を最適化する最適化ステップを実行するフィルタ係数最適化方法であって、

η_f ($1 \leq f \leq F-2$) を $\eta_f = w_f^* - 2w_{f+1}^* + w_{f+2}^*$ で定義されるフィルタ係数 $\tilde{w}^*$ の補助変数とし、

補助変数 η_f のコスト項は、log-concaveである補助変数 η_f の確率分布を用いて表されるものであり、

前記最適化ステップは、

補助変数 η_f のコスト項の和を含むコスト関数の最小化問題を解く
ことによりフィルタ係数 $\tilde{w}^*$ を最適化する

フィルタ係数最適化方法。

[請求項13] 請求項1ないし3のいずれか1項に記載の潜在変数最適化装置、請求項4ないし7のいずれか1項に記載のフィルタ係数最適化装置のいずれかとしてコンピュータを機能させるためのプログラム。

[図1]

ADMMに基づく潜在変数最適化

```

1: Initialize  $\hat{v}, \hat{u}, \gamma$ 
2: for  $n=1, \dots, N_{\text{iteration}}$  do
3:  $\tilde{w}^* \leftarrow \operatorname{argmin}_{\tilde{w}} L_0(\tilde{w}) + \gamma/2 \|\hat{D}\tilde{w} - \hat{v} + \hat{u} + \hat{b}\|_2^2$ 
4: for  $j=1, \dots, J$  do
5:  $v_j \leftarrow \operatorname{prox}_{L_j(\cdot)/\gamma}(D_j \tilde{w}^* + u_j + b_j)$ 
6:  $u_j \leftarrow u_j + D_j \tilde{w}^* - v_j + b_j$ 
7: end for
8: end for
9: return  $\tilde{w}^*$ 

```

図 1

[図2]

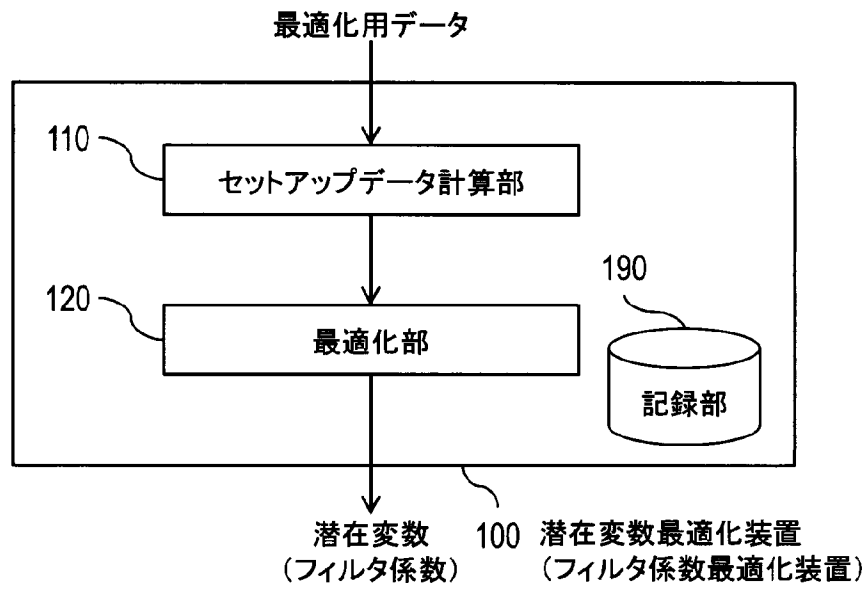


図2

[図3]

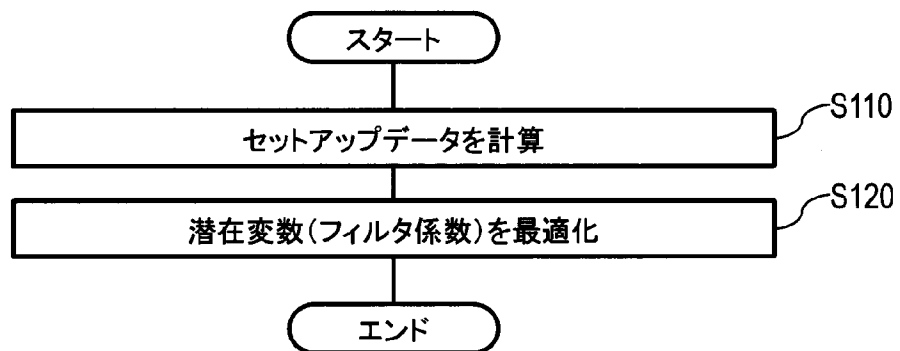


図3

[図4]

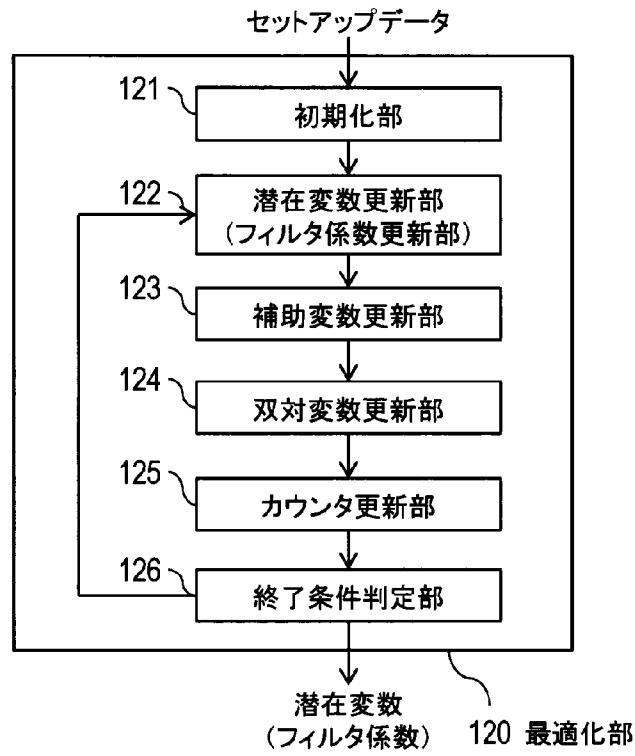


図4

[図5]

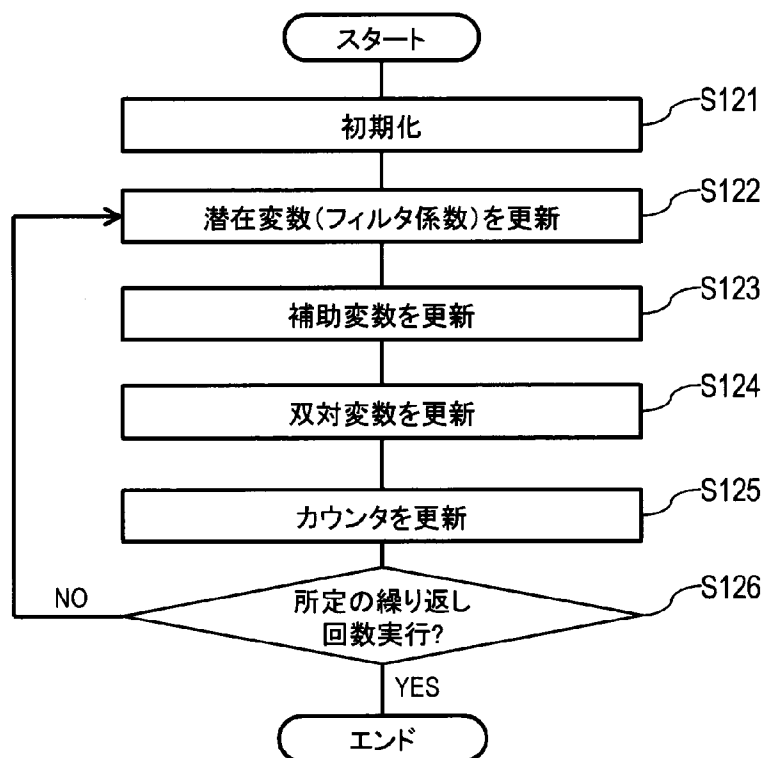


図5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/002205

A. CLASSIFICATION OF SUBJECT MATTER

G10L 21/0272 (2013.01) i

FI: G10L21/0272 100B; G10L21/0272 100A

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)

G10L21/0272

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2020
Registered utility model specifications of Japan	1996-2020
Published registered utility model applications of Japan	1994-2020

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	US 9668066 B1 (CEDAR AUDIO LTD.) 30.05.2017 (2017-05-30) column 7, line 48 to column 8, line 55	1-2, 8-9, 13 3-7, 10-12



Further documents are listed in the continuation of Box C.



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

06 April 2020 (06.04.2020)

Date of mailing of the international search report

14 April 2020 (14.04.2020)

Name and mailing address of the ISA/

Japan Patent Office

3-4-3, Kasumigaseki, Chiyoda-ku,

Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2020/002205

Patent Documents
referred in the
Report

Publication
Date

Patent Family

Publication
Date

US 9668066 B1

30 May 2017

(Family: none)

A. 発明の属する分野の分類（国際特許分類（IPC）） G10L 21/0272(2013.01)i FI: G10L21/0272 100B; G10L21/0272 100A														
B. 調査を行った分野 調査を行った最小限資料（国際特許分類（IPC）） G10L21/0272 最小限資料以外の資料で調査を行った分野に含まれるもの <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922 - 1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971 - 2020年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996 - 2020年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994 - 2020年</td> </tr> </table> 国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）			日本国実用新案公報	1922 - 1996年	日本国公開実用新案公報	1971 - 2020年	日本国実用新案登録公報	1996 - 2020年	日本国登録実用新案公報	1994 - 2020年				
日本国実用新案公報	1922 - 1996年													
日本国公開実用新案公報	1971 - 2020年													
日本国実用新案登録公報	1996 - 2020年													
日本国登録実用新案公報	1994 - 2020年													
C. 関連すると認められる文献														
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号												
X	US 9668066 B1 (CEDAR AUDIO LTD.) 30.05.2017 (2017 - 05 - 30) 第7欄第4行-第8欄第5行	1-2, 8-9, 13												
A		3-7, 10-12												
☐ C欄の続きにも文献が列挙されている。 ☒ パテントファミリーに関する別紙を参照。														
<table border="0"> <tr> <td>* 引用文献のカテゴリー</td> <td>“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの</td> </tr> <tr> <td>“A” 特に関連のある文献ではなく、一般的技術水準を示すもの</td> <td>“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの</td> </tr> <tr> <td>“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの</td> <td>“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの</td> </tr> <tr> <td>“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）</td> <td>“&” 同一パテントファミリー文献</td> </tr> <tr> <td>“O” 口頭による開示、使用、展示等に言及する文献</td> <td></td> </tr> <tr> <td>“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献</td> <td></td> </tr> </table>			* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの	“A” 特に関連のある文献ではなく、一般的技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの	“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの	“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献	“O” 口頭による開示、使用、展示等に言及する文献		“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献	
* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの													
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの													
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの													
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献													
“O” 口頭による開示、使用、展示等に言及する文献														
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献														
国際調査を完了した日 06.04.2020	国際調査報告の発送日 14.04.2020													
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） 上田 雄 5Z 5095 電話番号 03-3581-1101 内線 3591													

国際調査報告
パテントファミリーに関する情報

国際出願番号

PCT/JP2020/002205

引用文献			公表日	パテントファミリー文献	公表日
US	9668066	B1	30.05.2017	(ファミリーなし)	