

DSPを用いた雑音抑圧法のための 有色性駆動源カルマンフィルタアルゴリズム

Kalman Filter Algorithm with Colored Driving Source for Noise Suppression Method using DSP

寺島 大雅*

Hiromasa TERASHIMA

名取 隆廣*

Takahiro NATORI

田邊 造*

Nari TANABE

古川 利博†

Toshihiro FURUKAWA

1 はじめに

音声信号に環境雑音加わった劣化音声信号からクリアな音声信号を抽出する手法は、携帯電話のスマートフォンや高度道路交通システムのカーナビゲーションに代表される音声符号化や音声認識等の分野にとって、必要不可欠な技術である [1].

劣化音声信号からクリアな音声信号を抽出する手法は、近年さまざまな手法が提案されている [2],[3],[4]. 本論文では、モバイル性が高くローパワーな携帯電話を考慮することにより、単一マイクロフォンを用いた雑音抑圧法について議論する.

ANC(Adaptive Noise Canceller) は単一マイクロフォンを用いた代表的な雑音軽減手法である [5]. しかしながら、この手法はピッチ検出が必要となるため、ピッチ検出法によって雑音抑圧能力が左右される問題が存在している.

この問題を解決する手法として、線形予測を用いた手法が提案されている [2]. この手法は、シンプルであるため実用的な反面、音声信号に依存するパラメータを必要とする問題点が存在している. それゆえ、この手法の雑音抑圧能力はまだ改善の余地といえる.

一方、上記の手法とは全く異なる手法として、カルマンフィルタを用いた手法が提案されている [4]. この手法は、(Step 1) 音声信号を AR システムのパラメータ (以後、AR 係数と称する) を用いてモデル化することで AR 係数を推定する. (Step 2) Step 1 で推定した AR 係数を用いた状態方程式および音声信号に雑音加わった劣化音声信号を用いた観測方程式から構成される状態空間モデルに対して、カルマンフィルタ理論を適用させることより雑音を抑圧している. しかしながら、AR システムの次数決定問題に起因する AR 係数推定精度の劣化が雑音抑圧能力に依存するため、クリアな音声信号を推定することが困難な場合が存在する.

本論文は AR システムのコンセプトを必要としない有色性駆動源カルマンフィルタを用いた雑音抑圧法を提案する. 提案手法は、音声信号のみからなる有色性駆動源を有する状態方程式、および音声信号と雑音からなる観測方程式で構成される状態空間モデルに対して、カルマンフィルタアルゴリズムを用いることで雑音抑圧を実現している.

ここで注意しなければならないことがある. 一般にカルマンフィルタ理論は駆動源が白色信号でかつ、状態量とは無相関という条件のもとで状態空間モデルを構成している. これに

対して、提案手法は状態方程式に有色駆動源 (音声信号) を用いている. この問題について本論文では、有色性駆動源を含む状態空間モデルにカルマンフィルタ理論を適用可能であることについて論理的に証明している.

提案手法の特徴は、AR システムのコンセプトを必要としないことより (1) 少ない演算量でかつ、(2) 高い雑音抑圧能力を有していることである. 提案手法の有効性は、DSP ボードを用いて、(i) 視覚評価 (ii) 主観評価、および (iii) 演算量による処理速度評価より明らかにしている.

2 問題設定

本章は、本論文で扱う劣化音声信号について議論する.

劣化音声信号 $r(n)$ は、クリアな音声信号 $d(n)$ に加法性白色ガウス雑音 $v(n)$ が加わることで次式

$$r(n) = d(n) + v(n) \quad (1)$$

で表わされる. ただし、音声信号 $d(n)$ は雑音 $v(n)$ と無相関とし、雑音の分散値は $E[v(n)^2] = \sigma_v^2$ と仮定する. ここで、 $E[\cdot]$ は期待値である.

本論文の目的は音声信号 $d(n)$ に雑音 $v(n)$ が加わった劣化音声信号 $r(n)$ からクリアな音声信号 $d(n)$ のみを推定することである.

3 従来手法 [4]

本章では、(Step 1) 音声信号に対する AR 係数推定をした後に、(Step 2) 状態方程式と観測方程式からなる状態空間モデルに対してカルマンフィルタ理論を適用する 2 段階処理を用いた雑音抑圧法について簡単に説明する.

3.1 Step 1: AR 係数推定

音声信号 $d(n)$ が過去の音声信号を用いて AR システムでモデル化できると仮定すれば、音声信号 $d(n)$ は

$$d(n) = \sum_{l=1}^{L_c} a_l(n)d(n-l) + e(n) \quad (2)$$

となる. ここで、 $\{a_l(n)\}$ は AR 係数、 L_c は AR 次数、 $e(n)$ は $E[e(n)] = 0$ かつ AWGN(Additive White Gaussian Noise) である駆動源である. AR 係数の算出法はいくつか報告されており [6]、一例として、式 (2) の音声信号 $d(n)$ のかわりに劣化音声信号を用いて Yule-Walker 方程式を解く Levinson-Durbin 法 [10] や Burg 法 [11] がある.

* 諏訪東京理科大学

† 東京理科大学

表1: 従来手法のアルゴリズム

$$\begin{aligned}
& \text{[Initialization]} \quad \hat{x}_c(0|0) = \mathbf{0}, P_c(0|0) = I \\
& r_{\epsilon_c}(n) = E[\epsilon_c(n)\epsilon_c(n)^T] = \sigma_v^2 \\
& R_{\delta_c}[i, j] = \begin{cases} E\left[\left\{r(n) - \sum_{l=0}^{L_c} \alpha_l(n)r(n-l)\right\}^2\right] - \sigma_v^2 & (i, j = 1) \\ 0 & (\text{other}) \end{cases}
\end{aligned}$$

[Iteration]

Step 1

AR parameters $\{\alpha_l(n)\}$ estimation for the clean speech signal $d(n)$ using the linear prediction algorithm.

Step 2

$$1. P_c(n+1|n) = \Phi_c(n+1)P_c(n|n)\Phi_c^T(n+1) + R_{\delta_c}(n+1)$$

$$2. \mathbf{k}_c(n+1) = \{P_c(n+1|n)\mathbf{m}_c\} \cdot \{\mathbf{m}_c^T P_c(n+1|n)\mathbf{m}_c + r_{\epsilon_c}(n+1)\}^{-1}$$

$$3. \hat{x}_c(n+1|n) = \Phi_c(n+1)\hat{x}_c(n|n)$$

$$4. \hat{x}_c(n+1|n+1) = \hat{x}_c(n+1|n) + \mathbf{k}_c(n+1) \cdot \{y_c(n+1) - \mathbf{m}_c^T \hat{x}_c(n+1|n)\}$$

$$5. P_c(n+1|n+1) = \{I - \mathbf{k}_c(n+1)\mathbf{m}_c^T\} P_c(n+1|n)$$

$$6. n = n+1 \quad \text{go back 1.}$$

3.2 Step 2: 音声信号の推定

次いで、以後の議論のために次式を定義する。

$$\left. \begin{aligned} \mathbf{x}_c(n+1) &= [d(n+1), d(n), \dots, d(n-L_c+2)]^T \\ \boldsymbol{\delta}_c(n+1) &= [e(n+1), 0, \dots, 0]^T \\ \epsilon_c(n+1) &= v(n+1) \end{aligned} \right\} (3)$$

式(2)を考慮した状態方程式と式(1)を考慮した観測方程式からなる状態空間モデルは

[状態方程式]

$$\mathbf{x}_c(n+1) = \Phi_c(n+1)\mathbf{x}_c(n) + \boldsymbol{\delta}_c(n+1) \quad (4)$$

[観測方程式]

$$y_c(n+1) = r(n+1) = \mathbf{m}_c^T \mathbf{x}_c(n+1) + \epsilon_c(n+1) \quad (5)$$

となる。ここで、 $L_c \times L_c$ 行列の状態遷移行列 $\Phi_c(n-1)$ 、 L_c 次元観測ベクトル \mathbf{m}_c は次式となる。

$$\left. \begin{aligned} \Phi_c(n-1) &= \begin{bmatrix} \alpha_1(n+1) & \alpha_2(n+1) & \dots & \dots & \alpha_{L_c}(n+1) \\ 1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix} \\ \mathbf{m}_c &= [1, 0, \dots, 0]^T \end{aligned} \right\} (6)$$

従来手法は式(4)の状態方程式と式(5)の観測方程式からなる状態空間モデルに対して、駆動源が白色信号でかつ、状態量である音声信号が雑音と無相関という条件を満たしていることを用いて、カルマンフィルタ理論を適用することで雑音抑圧を実現している。従来手法で用いるアルゴリズムを表1に示す。

しかしながら、従来手法は、AR 次数の決定問題に起因する AR システムのパラメータ推定精度問題によって、雑音抑圧能力が劣化する。それゆえ、AR システムのコンセプトを必要としない雑音抑圧法が望まれる。

表2: 提案手法のアルゴリズム

$$\begin{aligned}
& \text{[Initialization]} \quad \hat{x}_p(0|0) = \mathbf{0}, P_p(0|0) = I, r_{\epsilon_p}(n+1) = \sigma_v^2 \\
& R_{\delta_p}[i, j] = \begin{cases} \frac{1}{L_p-1} \sum_{l=0}^{L_p-1} r^2(n-l) - \sigma_v^2 & (i, j = 1) \\ 0 & (\text{other}) \end{cases}
\end{aligned}$$

[Iteration]

$$1. P_p(n+1|n) = \Phi_p P_p(n|n)\Phi_p^T + R_{\delta_p}(n+1)$$

$$2. \mathbf{k}_p(n+1) = \{P_p(n+1|n)\mathbf{m}_p\} \cdot \{\mathbf{m}_p^T P_p(n+1|n)\mathbf{m}_p + r_{\epsilon_p}(n+1)\}^{-1}$$

$$3. \hat{x}_p(n+1|n) = \Phi_p \hat{x}_p(n|n)$$

$$4. \hat{x}_p(n+1|n+1) = \hat{x}_p(n+1|n) + \mathbf{k}_p(n+1) \cdot \{y_p(n+1) - \mathbf{m}_p^T \hat{x}_p(n+1|n)\}$$

$$5. P_p(n+1|n+1) = \{I - \mathbf{k}_p(n+1)\mathbf{m}_p^T\} P_p(n+1|n)$$

$$6. n = n+1 \quad \text{go back 1.}$$

4 提案手法

本章では従来手法の問題点である AR システムのコンセプトを用いずに、劣化音声信号のみからクリアな音声信号を推定する雑音抑圧法を提案する。

提案手法における L_p 次元の状態ベクトルを

$$\mathbf{x}_p(n+1) = [d(n+1), d(n), \dots, d(n-L_p+2)]^T \quad (7)$$

と定義したとき、AR システムのコンセプトを用いずに音声信号のみで状態方程式を表わすと

[状態方程式]

$$\mathbf{x}_p(n+1) = \Phi_p \mathbf{x}_p(n) + \boldsymbol{\delta}_p(n+1) \quad (8)$$

となる。ここで、 $L_p \times L_p$ 行列の状態遷移行列 Φ_p と L_p 次元駆動源ベクトル $\boldsymbol{\delta}_p(n)$ は次式となる。

$$\left. \begin{aligned} \Phi_p &= \begin{bmatrix} 0 & 0 & \dots & \dots & 0 \\ 1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix} \\ \boldsymbol{\delta}_p(n+1) &= [d(n+1), 0, \dots, 0]^T \end{aligned} \right\} (9)$$

このとき、式(9)の $L_p \times L_p$ 行列の提案手法の状態遷移行列 Φ_p は、0と1のみで構成されているシフト行列となっていることにより、従来手法の状態遷移行列 $\Phi_c(n-1)$ と比較して演算量が軽減可能であることが容易に想像できる。また、駆動源ベクトル $\boldsymbol{\delta}_p(n)$ は式(9)に示すように音声信号 $d(n)$ のみで構成されていることから、有色性駆動源を構成する。

次に、式(1)より音声信号に雑音を加わって構成された劣化音声信号を観測方程式 $y_p(n+1)$ で表わすと

[観測方程式]

$$y_p(n+1) = r(n+1) = \mathbf{m}_p^T \mathbf{x}_p(n+1) + \epsilon_p(n+1) \quad (10)$$

となる。 L_p 次元観測ベクトル \mathbf{m}_p 、雑音 $\epsilon_c(n+1)$ は次式となる。

$$\mathbf{m}_p = [1, 0, \dots, 0]^T, \quad \epsilon_p(n+1) = v(n+1) \quad (11)$$

提案手法は式(8)の状態方程式と式(10)の観測方程式から状態空間モデルを構成し、カルマンフィルタアルゴリズムを用いて雑音抑圧を実行することでクリアな音声信号の推定を行う手法である。提案手法のアルゴリズムを表2に示す。

ここで、状態方程式に用いられている駆動源ベクトル $\delta_p(n)$ に着目すると、提案手法の状態方程式において音声信号であることから有色性駆動源となる。一般にカルマンフィルタ理論は白色性駆動源でかつ、状態量と雑音が無相関であることを適用前提としているため、提案手法のように有色性駆動源を含む状態空間モデルに対してカルマンフィルタ理論への適用条件を満足しない。

そのため、新たに提案した状態空間モデルがカルマンフィルタ理論へ適用可能か検討する必要がある。次節ではこの点について検討を行う。

4.1 提案手法の状態空間モデルをカルマンフィルタに適用した際、状態量に与える影響の検証

本節では、提案手法の状態空間モデルにカルマンフィルタ理論を適用した場合、求めるべき状態量に与える影響を検証する。ただし、以後の議論を簡単にするために次のことを定義する。

$\hat{x}_p(n+1|n)$: 過去の観測信号により求められる状態ベクトル $x_p(n+1)$ の最適推定値

$\tilde{x}_p(n+1|n)$: $\tilde{x}_p(n+1|n) = x_p(n+1) - \hat{x}_p(n+1|n)$ で定義される状態ベクトル $x_p(n+1)$ に対する予測誤差

提案手法の状態空間モデルにカルマンフィルタを適用した際、白色性駆動源の条件が必要なのは表2手順1の $P_p(n+1|n)$ であるため、 $\hat{x}_p(n+1|n) = \Phi_p \hat{x}_p(n+1|n) + \delta_p(n+1)$ を用いて行列 $P_p(n+1|n)$ の導出を行う。行列 $P_p(n+1|n)$ は

$$\begin{aligned} P_p(n+1|n) &= E[\tilde{x}_p(n+1|n)\tilde{x}_p^T(n+1|n)] \\ &= \Phi_p E[\tilde{x}_p(n|n)\tilde{x}_p^T(n|n)]\Phi_p^T + E[\delta_p(n+1)\delta_p^T(n+1)] \\ &\quad + \Phi_p E[\tilde{x}_p(n|n)\delta_p^T(n+1)] + E[\delta_p(n+1)\tilde{x}_p^T(n|n)]\Phi_p^T \\ &= \Phi_p P_p(n|n)\Phi_p^T + R_{\delta_p}(n+1) + Q_p(n+1) + Q_p^T(n+1) \end{aligned} \quad (12)$$

となる。ただし、 $R_{\delta_p}(n+1) = E[\delta_p(n+1)\delta_p^T(n+1)]$ 、 $Q_p(n+1) = \Phi_p E[\tilde{x}_p(n|n)\delta_p^T(n+1)]$ である。

一般にカルマンフィルタ理論では、白色性駆動源の際 $Q_p(n+1) = O$ となるが、提案手法は有色性駆動源であることから $Q_p(n+1) \neq O$ となるため、表2手順2~5に影響を与えることが予想される。よって式(12)の $Q_p(n+1)$ に着目して議論を行う。ただし、以後の議論を簡単にするため次式を定義する。

$$\left. \begin{aligned} \delta_p(n+1) &= \Gamma x_p(n+1) \\ \Phi_p \tilde{x}_p(n|n) &= x_p(n+1) - \hat{x}_p(n+1) - \delta_p(n+1) \\ x_p(n+1) &= \hat{x}_p(n+1|n+1) + \tilde{x}_p(n+1) \\ E[\hat{x}_p(n+1|n)\tilde{x}_p(n+1|n+1)] &= O \end{aligned} \right\} \quad (13)$$

ここで $L_p \times L_p$ 次元の行列 Γ は

$$\Gamma = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix} \quad (14)$$

である。

式(13)と式(14)を用いて $Q_p(n+1)$ を表すと次式のようにになる。

$$\begin{aligned} Q_p(n+1) &= E[\Phi_p \tilde{x}_p(n|n)\tilde{x}_p^T(n+1|n+1)]\Gamma^T \\ &= E[x_p(n+1)x_p^T(n+1)]\Gamma^T - E[\hat{x}_p(n+1|n) \\ &\quad \cdot \hat{x}_p^T(n+1|n+1)]\Gamma^T - \Gamma E[x_p(n+1)x_p^T(n+1)]\Gamma^T \\ &= E[x_p(n+1)x_p^T(n+1)]\Gamma^T - \Phi_p E[\tilde{x}_p(n|n) \\ &\quad \cdot \hat{x}_p^T(n+1|n+1)]\Gamma^T - \Gamma E[x_p(n+1)x_p^T(n+1)]\Gamma^T \end{aligned}$$

ここで、過去の時刻における $P_p(n|n)$ と $\hat{x}_p(n|n)$ の全要素が $Q_p(n+1)$ の影響を受けているものと仮定するとともに、 $Q_p(n+1)$ の影響を含んだ各ベクトル・行列の要素を灰色部、それらの影響がない要素を非灰色部と表すこととする。

これにより、式(4.1)の $Q_p(n+1)$ の要素は次式のようになる。

$$\begin{aligned} Q_p(n+1) &= \begin{bmatrix} E[d(n+1)d(n+1)] \\ E[d(n)d(n)] \\ \vdots \\ \vdots \\ E[d(n-L_p+2)d(n-L_p+2)] \end{bmatrix} \begin{array}{c} O \\ \\ \\ \\ O \end{array} \\ &- \begin{bmatrix} 0 \\ E[\hat{d}(n|n)\hat{d}(n+1|n+1)] \\ E[\hat{d}(n-1|n-1)\hat{d}(n|n)] \\ \vdots \\ E[\hat{d}(n-L_p+2|n-L_p+2) \\ \cdot \hat{d}(n-L_p+3|n-L_p+3)] \end{bmatrix} \begin{array}{c} O \\ \\ \\ \\ O \end{array} \\ &- \begin{bmatrix} E[d(n+1)d(n+1)] \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix} \begin{array}{c} O \\ \\ \\ \\ O \end{array} \end{aligned} \quad (15)$$

$$= \begin{bmatrix} 0 \\ e_2(n) \\ e_3(n) \\ \vdots \\ e_{L_p}(n) \end{bmatrix} \begin{array}{c} O \\ \\ \\ \\ O \end{array} \quad (16)$$

ただし、 $e_i(n) = E[d(n-i+2)d(n+1)] - E[\hat{d}(n-i+2|n)\hat{d}(n+1|n+1)]$ ($1 < i \leq L_p$) とする。従って、 $Q_p(n+1) \neq O$ であることが確認できる。それゆえ、 $Q_p(n+1)$ が手順1~5のどの部分に影響を与えるかについて以下で確認する。

$n+1$ 時刻において過去の $P_p(n|n)$ の全要素が $Q_p(n+1)$ の影響を受けているという仮定より、式(16)が表2手順1の $P_p(n+1|n)$ に与える影響は、

$$\begin{aligned}
 P_p(n+1|n) &= \Phi_p P(n|n) \Phi_p^T + R_{\delta_p}(n+1) + Q_p(n+1) + Q_p^T(n+1) \\
 &= \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & P_p^{L_p-1, L_p-1}(n|n) & \\ 0 & & & \end{bmatrix} \\
 &+ \begin{bmatrix} E[d(n+1)d(n+1)] & & & \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{bmatrix} + \begin{bmatrix} 0 \\ e_r(2) \\ e_r(3) \\ \vdots \\ e_r(L_p) \end{bmatrix} \begin{bmatrix} 0 \\ O \\ O \\ O \\ O \end{bmatrix} \\
 &+ \begin{bmatrix} 0 & e_r(2) & e_r(3) & \cdots & e_r(L_p) \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{bmatrix} \begin{bmatrix} O \\ O \\ O \\ O \\ O \end{bmatrix} \quad (17) \\
 &= \begin{bmatrix} E[d(n+1)d(n+1)] & e_r(2) & e_r(3) & \cdots & e_r(L_p) \\ e_r(2) & & & & \\ e_r(3) & & & & \\ \vdots & & & & \\ e_r(L_p) & & & & \end{bmatrix} \begin{bmatrix} P_p^{L_p-1, L_p-1}(n|n) \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{bmatrix}
 \end{aligned}$$

となる。

同様に、表2手順2の $k_p(n+1)$ が $Q_p(n+1)$ から影響を受ける部分は次式のようになる。

$$\begin{aligned}
 k_p(n+1) &= P_p(n+1|n) m_p \cdot \{m_p^T P_p(n+1|n) m_p + r_{\epsilon_p}(n+1)\}^{-1} \\
 &= \begin{bmatrix} E[d(n+1)d(n+1)] \\ e_r(2) \\ e_r(3) \\ \vdots \\ e_r(L_p) \end{bmatrix} \cdot \{E[d(n+1)d(n+1)] + \sigma_v^2\}^{-1} \\
 &= \begin{bmatrix} k_p(n+1)[1, 1] \\ k_p(n+1) \\ k_p(n+1) \end{bmatrix} \quad (18)
 \end{aligned}$$

さらに、過去の $\hat{x}_p(n|n)$ は全要素が $Q_p(n+1)$ の影響を受けているという仮定より、表2手順3の $\hat{x}_p(n+1|n)$ は

$$\hat{x}_p(n+1|n) = \Phi_p \hat{x}_p(n|n) = \begin{bmatrix} 0 \\ \hat{x}_p(n|n) \\ \hat{x}_p(n|n) \\ \hat{x}_p(n|n) \end{bmatrix} \quad (19)$$

となる。

それゆえ、表2手順4の $\hat{x}_p(n+1|n+1)$ は $Q_p(n+1)$ の影響を受けている式(18)の $k_p(n+1)$ と(19)の $\hat{x}_p(n+1|n)$ が含まれることより、状態量ベクトル $\hat{x}_p(n+1|n+1)$ は次式で表される。

$$\begin{aligned}
 \hat{x}_p(n+1|n+1) &= \hat{x}_p(n+1|n) + k_p(n+1) \cdot \{y_p(n+1) - m_p^T \hat{x}_p(n+1|n)\} \\
 &= \begin{bmatrix} 0 \\ \hat{x}_p(n+1|n) \\ \hat{x}_p(n+1|n) \\ \hat{x}_p(n+1|n) \end{bmatrix} + \begin{bmatrix} k_p(n+1)[1, 1] \\ k_p(n+1) \\ k_p(n+1) \end{bmatrix} \cdot \{y_p(n+1) - m_p^T \hat{x}_p(n+1|n)\}
 \end{aligned}$$

$$\begin{aligned}
 &= \begin{bmatrix} 0 \\ \hat{x}_p(n|n) \\ \hat{x}_p(n|n) \\ \hat{x}_p(n|n) \end{bmatrix} + \begin{bmatrix} k_p(n+1)[1, 1] \\ k_p(n+1) \\ k_p(n+1) \end{bmatrix} \cdot \{y_p(n+1) - 0\} \\
 &= \begin{bmatrix} \hat{x}_p(n+1|n+1)[1, 1] \\ \hat{x}_p(n+1|n+1) \\ \hat{x}_p(n+1|n+1) \end{bmatrix} \quad (20)
 \end{aligned}$$

従って、推定された音声信号は式(19)の $\hat{x}_p(n+1|n)$ の第1要素であることから、有色性駆動源の影響を受けていないことが確認できる。

最後に、表2手順5の $P_p(n+1|n+1)$ は、式(4.1)の手順1と式(18)の手順2の結果を用いて $Q_p(n+1)$ の影響を表すと

$$\begin{aligned}
 P_p(n+1|n+1) &= \{I - k_p(n+1) m_p^T\} P_p(n+1|n) \\
 &= \left\{ \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} - \begin{bmatrix} k_p(n+1)[1, 1] \\ k_p(n+1) \\ k_p(n+1) \end{bmatrix} \begin{bmatrix} 0 \\ O \\ O \end{bmatrix} \right\} \\
 &\cdot \begin{bmatrix} E[d(n+1)d(n+1)] & e_r(2) & e_r(3) & \cdots & e_r(L_p) \\ e_r(2) & & & & \\ e_r(3) & & & & \\ \vdots & & & & \\ e_r(L_p) & & & & \end{bmatrix} \begin{bmatrix} P_p^{L_p-1, L_p-1}(n|n) \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{bmatrix} \\
 &= \begin{bmatrix} \alpha E[d(n+1)d(n+1)] & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{bmatrix} \quad (\alpha = 1 - k_p(n+1)) \quad (21)
 \end{aligned}$$

となる。

以上のことから、本提案手法は $\hat{x}_p(n+1|n+1)$ の第1要素が推定された音声信号であることより、推定された音声信号は有色性駆動源の影響を受けないといえる。このことより、提案手法の状態空間モデルはカルマンフィルタ理論を適応可能であるといえる。

5 実装評価

本章では、提案手法の有効性を確認するために、従来手法[4]と提案手法をDSPボード(TEXAS INSTRUMENTS社製品 TMS320C6713 DSK)へ実装した結果に対して (i) 波形評価、(ii) 主観評価、(iii) 演算量評価を用いて性能比較を行った。また、以下に実装評価で用いた音声信号 $r(n)$ と $v(n)$ を示す。

音声信号 $r(n)$: 成人女性の音声は無音室にて録音を行い、
 音声を16kHzでサンプリングした音声信号
 雑音 $v(n)$: 白色性ガウス雑音

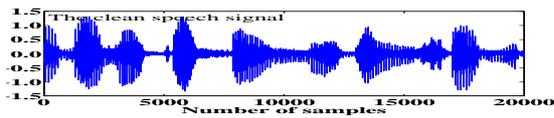


図 1: 音声信号

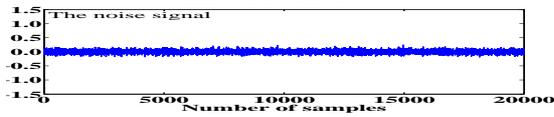


図 2: 雑音信号

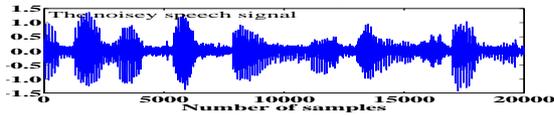


図 3: 観測信号 (音声 + 雑音)

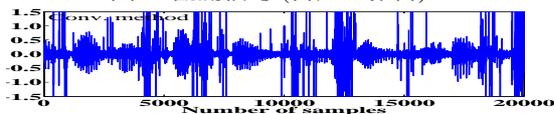


図 4: 従来手法 [4]

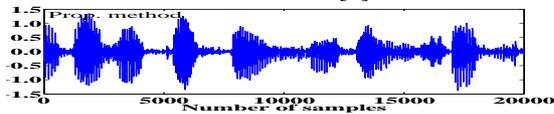


図 5: 提案手法

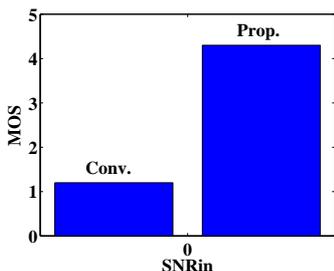


図 6: MOS 聴取り評価表

表 3: MOS 評価基準

基準	評価値
非常に良い	5
良い	4
普通	3
悪い	2
非常に悪い	1

5.1 波形評価

音声信号 (図 1) に環境雑音 (図 2) が加わった劣化音声信号 (図 3) に対して、従来手法と提案手法を用いて雑音抑圧を行った結果を図 4 と図 5 にそれぞれ示す。

従来手法 (図 4) による結果について考察すると、雑音抑圧を行うと同時に音声信号も抑圧している。また、音声信号の無音区間に対応する劣化音声信号の区間、すなわち雑音区間のみで従来手法の雑音抑圧を実行すると著しく性能が劣化している。これは、Step1 の AR 係数推定が雑音に対して処理していることが考えられる。それに対して提案手法 (図 5) は音声信号を抑圧することなく雑音抑圧が出来ていることが確認できる。

このことにより、提案手法は従来手法に比べ高い雑音抑圧能力があり、クリアな音声信号推定が行えていることが確認できる。

5.2 主観評価

本節の主観評価では、MOS (Mean Opinion Score) [4] 評価を用いて聞き取り評価を行った。これは表 3 に示す 5 段階評価基準に従って、50 人 (男性 25 人、女性 25 人) に比較評価してもらい、この結果の平均値を評価結果として図 7 に示す。



図 7: 実装及び DSP ボードブロック図

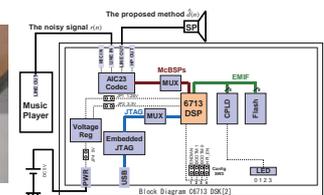


表 4: 乗算回数比較表

Computing or Solving	Multiplications	
	Conv.[4]	Prop.
$1.P(n+1 n)$	$2L_c^2$	0
$2.K(n+1)$	L_c	1
$3.x(n+1 n)$	L_c	0
$4.x(n+1 n+1)$	L_c	1
$5.P(n+1 n+1)$	L_c^2	1
Total	$3L_c^2 + 3L_c$	3

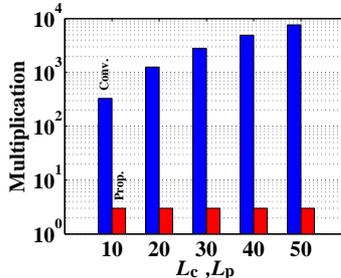


図 8: 従来手法 [4] 及び提案手法の乗算回数

図 7 の MOS 結果より、提案手法は従来手法よりも高い値となった。このことから、提案手法は高音質を維持しながら高い雑音抑圧能力を実現していることがわかる。

5.3 演算量評価

本節の演算量評価では、カルマンフィルタを用いた雑音抑圧法である従来手法と提案手法のカルマンフィルタアルゴリズムの各手順の乗算回数を表 4 に示すとともに、その乗算回数を演算量として比較した結果を図 8 に示す。

図 8 の結果より従来手法は L_c の増加に対して単調増加するが、提案手法は L_p に関係なく常に 3 回と一定である。このことにより、提案手法は高速な雑音抑圧法といえる。

6 まとめ

本論文は、有色性駆動源カルマンフィルタを用いた雑音抑圧法を提案した。

提案手法は音声信号のみからなる有色性駆動源を含む状態方程式と、音声信号と雑音により観測方程式を構成した状態空間モデルに対して、カルマンフィルタ理論を用いることで雑音抑圧を実現している。そのため AR システムのコンセプトを必要としないことにより、(1) 少ない演算量でかつ、(2) 高い雑音抑圧能力を実現した。

提案手法の有効性は DSP ボードを用いて、視覚評価、主観評価、および演算量による処理速度評価より明らかにしている。以上より、提案手法は少ない演算量でかつ、高い雑音抑圧が可能となり実用的な雑音抑圧手法である。

参考文献

- [1] 荒金 陽介, 里村 昌史, 樋山 智, 辻 ゆかり, "自動車内環境における電子メールサポートシステムの構築とその評価" 信学誌 Vol.J87-A No.2 pp.242-252 Feb. 2004.
- [2] 川村 新, "線形予測分析に基づく騒音抑圧法" 電子情報通信学会誌 2002/4 Vol.J85-A No.4 pp.415-423
- [3] R.E.Kalman, "A new approach to Linear Filtering and Prediction Problems," Trans.ASME-journal of Basic Engineering, Vol.82,no.series-D,pp.34-45,1960.
- [4] W.Kim, "Noise Variance Estimation for Kalman Filtering of Noisy Speech" IEICE TRANS.INF.& SYST.,VOL.E84-D,NO.1 JANUARY 2001
- [5] 権川 嘉延, 野村 康雄, "2 次元経路モデルを必要としないアクティブノイズコントロールシステム" 信学誌 Vol.J82-A No.2 pp.209-217 Feb. 1999.
- [6] 江原 義朗, デジタル信号処理, 電機大出版局.
- [7] N. Tanabe, T. Furukawa, S. Tsujii, "Robust Noise Suppression Algorithm with the Kalman Filter Theory for White and Colored Disturbance," IEICE Trans. Fundamentals, vol.E91-A, No.3, March 2008.
- [8] J.Huang and Y.Zhao, "An Energy-Constrained Signal Sub-Space Method for Speech Enhancement and Recognition in White and Colored Noises," Speech Vommunication,vol.26,no.26 pp.165-181,1998.
- [9] M.Kato,A.Sugiyama,and M.Serizawa, "Noise suppression with high speech quality based on weighted noise estimation and MMSESTSA," IEICE Trans.Fundamentals(Japanese Edition),vol.J87-A,no.7,pp.851-860,July 2004.
- [10] N. Levinson, "The Wiener RMS error criterion in iter design and prediction," Journal of Mathematics and Physics, v25, pp.261-278, 1947
- [11] J. P. Burg, "Maximum Entropy Spectral Analysis," Astronomy and Astrophysics Supplement, Vol.15, p.383, 1974
- [12] 谷秋 隆嗣, カルマンフィルタと適応信号処理, コロナ社, 2005.
- [13] 杉山 昭彦, 岸沢 昌宏, 加藤 正徳, 西谷 隆夫, "携帯電話用雑音抑圧技術の標準化," 電子情報通信学会論文誌 A, Vol.J87-A, No.7, pp.913-920, 2002
- [14] 上田 裕市, 阿川 貴之, 青山 正純, 渡邊 亮, "合成フィルタによるデジタル補聴器のための振幅圧縮処理," 電子情報通信学会論文誌 A, Vol.J81-A, No.12, pp.1728-1738
- [15] 藤井 治樹, "ITS への期待とその発展基盤," 電子情報通信学会論文誌 A, Vol.J91-A, No.1, pp.2-10