

演算量の軽減を考慮した実用的な雑音抑圧法

A Practical Kalman Filter based Noise Suppression Algorithm
with Reduced Computational Complexity

小野塚 康浩* 田畑 雅崇* 田邊 造* 松江 英明* 古川 利博†

1 はじめに

雑音の影響により劣化した音声信号 (劣化音声信号) からクリアな音声信号を推定する単一マイクロフォンを用いた雑音抑圧法 [1] は、携帯電話やカーナビゲーション等のための音声符号化や音声認識の分野において必要不可欠な技術である。

その中でも、カルマンフィルタ [2] を用いた雑音抑圧法 [3] は他の雑音抑圧法よりも高い雑音抑圧能力を持つことが知られている。この手法は、(Step1) 劣化音声信号から音声信号に対する AR システムのパラメータ (AR 係数) を推定し、(Step2) Step1 で推定された AR 係数を用いた状態空間モデルに対してカルマンフィルタ理論により雑音抑圧を実行する 2 段階処理である。しかしながら、この手法は高い雑音抑圧能力はあるが Step1 の AR 次数の決定が困難となることもあるため、AR 係数推定精度が Step2 の雑音抑圧能力に影響を与えてしまう問題がある。従って、AR システムのコンセプトに依存しない雑音抑圧法が求められる。

この問題を解決する手法として、著者らは AR システムのコンセプトを必要としないカルマンフィルタのみを用いた雑音抑圧法を提案した [4]。この手法は高い雑音抑圧能力を得ることが可能であるが、逆行列を必要とするため実用面において改善の余地がある。

本論文は、文献 [4] を拡張した雑音抑圧法を提案する。提案手法の特徴は、文献 [4] の状態空間モデルを再構成することで、音質を犠牲にすることなく少ない演算量で高い雑音抑圧能力を得ることが可能なことである。提案手法の有効性は、客観評価と主観評価および演算量評価から確認している。また、有色性駆動源を持つ新たに提案した状態空間モデルに関しても定量的に確認している。

2 問題設定

クリアな音声信号 $d(n)$ に雑音 $v(n)$ が加わり劣化したと仮定すると、劣化音声信号 $d(n)$ は次式で与えられる。

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

本論文は劣化音声信号 $r(n)$ から、音声信号 $d(n)$ を推定する雑音抑圧法について議論するものである。本論文では、近年様々な雑音分散推定法 [5] が提案されているという理由からここでは雑音分散値 σ_v^2 を既知とし特に議論しない。

* 諏訪東京理科大学

† 東京理科大学

表 1: 提案手法のアルゴリズム

[Initialization]
$\hat{x}_p(0 0) = \mathbf{0}, P_p(0 0) = I, r_{\epsilon_p}(n+1) = \sigma_v^2$
$R_{\delta_p}[i, j] = \begin{cases} \frac{1}{L_p-1} \sum_{l=1}^{L_p-1} r^2(n-l) - \sigma_v^2 & (i, j = 1) \\ 0 & (other) \end{cases}$
[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. $\hat{\mathbf{x}}_p(n+1 n) = \Phi_p \hat{\mathbf{x}}_p(n n)$
4. $\hat{\mathbf{x}}_p(n+1 n+1) = \hat{\mathbf{x}}_p(n+1 n) + \mathbf{k}_p(n+1) \cdot \{y_p(n+1) - \mathbf{m}_p^T \hat{\mathbf{x}}_p(n+1 n)\}$
5. $P_p(n+1 n+1) = \{I - \mathbf{k}_p(n+1) \mathbf{m}_p^T\} P_p(n+1 n)$
6. $n+1 \leftarrow n$ go back 1.

3 提案手法

本章は、文献 [4] を拡張することで有色性駆動源を持つカルマンフィルタに基づく実用的な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 $\mathbf{x}_p(n+1) = [d(n+1), d(n), \dots, d(n-L_p+2)]^T$ と定義し、駆動源に音声信号 $d(n)$ のみを用いた状態方程式は

$$[\text{状態方程式}] \quad \mathbf{x}_p(n+1) = \Phi_p \mathbf{x}_p(n) + \delta_p(n+1) \quad (2)$$

となる。ただし $L_p \times L_p$ 次元の状態遷移行列 Φ_p と L_p 次元の駆動源ベクトル δ_p は次式で表される。

$$\Phi_p = \left. \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \right\} \quad (3)$$

$$\delta_p(n+1) = [d(n+1), 0, \dots, 0]^T$$

次いで演算量を軽減するために観測信号 $y_p(n) = r(n)$ と定義し、式 (1) に基づく観測方程式は

$$[\text{観測方程式}] \quad y_p(n+1) = \mathbf{m}_p^T \mathbf{x}_p(n+1) + \epsilon_p(n+1) \quad (4)$$

と書き直される。ただし L_p 次元の観測ベクトル $\mathbf{m}_p = [1, 0, \dots, 0]^T$ とし、雑音 $\epsilon_p(n+1) = v(n+1)$ とする。

提案手法は、式 (2) の状態方程式と式 (4) の観測方程式からなる状態空間モデルより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を推定する手法である。提案手法のアルゴリズムは表 1 に示す。

ここで状態空間モデルを構成する駆動源ベクトル $\delta_p(n)$ に着目すると、提案手法の駆動源ベクトル $\delta_p(n)$ は式 (3) に示

すように音声信号 $d(n)$ で構成されている。すなわち、駆動源が有色性信号である。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という条件で適用可能なのに対して、提案手法は先に述べた条件を満たさない。それゆえ、詳細は後に述べるが、状態ベクトルと駆動源ベクトルとの間に相関が生じてしまう。従って、提案手法の状態空間モデルがカルマンフィルタに適用可能か検討する必要がある。

次章では提案手法の状態空間モデルに対してカルマンフィルタ理論を適用可能かについて議論する。

4 提案手法の状態空間モデルにカルマンフィルタを適用可能についての検討

本章では、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。

$\hat{\mathbf{x}}_p(n+1|n)$: 過去の観測信号 $y_p(1), y_p(2), \dots, y_p(n)$ により求められる状態ベクトル $\mathbf{x}_p(n+1)$ の最適推定値

$\tilde{\mathbf{x}}_p(n+1|n)$: $\tilde{\mathbf{x}}_p(n+1|n) = \mathbf{x}_p(n+1) - \hat{\mathbf{x}}_p(n+1|n)$ で定義される状態ベクトル $\mathbf{x}_p(n+1)$ 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1手順1の $P_p(n+1|n)$ である。それゆえ、 $\hat{\mathbf{x}}_p(n+1|n) = \Phi_p \hat{\mathbf{x}}_p(n+1|n) + \delta_p(n+1)$ [4][5] を用いて行列 $P_p(n+1|n)$ を導出すると次式のようになる。

$$\begin{aligned} P_p(n+1|n) &= E[\tilde{\mathbf{x}}_p(n+1|n)\tilde{\mathbf{x}}_p^T(n+1|n)] \\ &= E[\{\Phi_p \tilde{\mathbf{x}}_p(n|n) + \delta_p(n+1)\}\{\Phi_p \tilde{\mathbf{x}}_p(n|n) + \delta_p(n+1)\}^T] \\ &= \Phi_p E[\tilde{\mathbf{x}}_p(n|n)\tilde{\mathbf{x}}_p^T(n|n)]\Phi_p^T + E[\delta_p(n+1)\delta_p^T(n+1)] \\ &\quad + \Phi_p E[\tilde{\mathbf{x}}_p(n|n)\delta_p^T(n+1)] + E[\delta_p(n+1)\tilde{\mathbf{x}}_p^T(n|n)]\Phi_p^T \\ &= \Phi_p P_p(n|n)\Phi_p^T + R_{\delta_p}(n+1) + Q_p(n+1) + Q_p^T(n+1) \end{aligned} \quad (5)$$

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

一般にカルマンフィルタ理論では、白色性駆動源のとき $Q_p(n+1) = O$ となる。一方、提案手法は有色性駆動源であることから $Q_p(n+1) \neq O$ である。従って、有色性駆動ベクトルは直接的又は間接的に表1手順2~5に影響を与えることが予想される。そこで、次に式(5)の $Q_p(n+1)$ に着目して議論する。ただし、以後の議論を簡単にするため次式を定義する [4]。

$$\left. \begin{aligned} \delta_p(n+1) &= \Gamma \mathbf{x}_p(n+1) \\ \Phi_p \tilde{\mathbf{x}}_p(n|n) &= \mathbf{x}_p(n+1) - \hat{\mathbf{x}}_p(n+1) - \delta_p(n+1) \\ \mathbf{x}_p(n+1) &= \hat{\mathbf{x}}_p(n+1|n+1) + \tilde{\mathbf{x}}_p(n+1) \\ E[\hat{\mathbf{x}}_p(n+1|n)\tilde{\mathbf{x}}_p^T(n+1|n+1)] &= O \end{aligned} \right\} \quad (6)$$

ここで $L_p \times L_p$ 次元の行列 Γ は

$$\Gamma = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix} \quad (7)$$

である。

式(6)と式(7)を用いて $Q_p(n+1)$ を表すと次式のようになる。

$$\begin{aligned} Q_p(n+1) &= E[\Phi_p \tilde{\mathbf{x}}_p(n|n)\mathbf{x}_p^T(n+1)]\Gamma^T \\ &= E[\{\mathbf{x}_p(n+1) - \hat{\mathbf{x}}_p(n+1|n) - \delta_p(n+1)\}\mathbf{x}_p^T(n+1)]\Gamma^T \\ &= E[\mathbf{x}_p(n+1)\mathbf{x}_p^T(n+1)]\Gamma^T \\ &\quad - E[\hat{\mathbf{x}}_p(n+1|n)\{\mathbf{x}_p(n+1|n+1) + \tilde{\mathbf{x}}_p(n+1|n+1)\}^T]\Gamma^T \\ &\quad - \Gamma E[\mathbf{x}_p(n+1)\mathbf{x}_p^T(n+1)]\Gamma^T \\ &= E[\mathbf{x}_p(n+1)\mathbf{x}_p^T(n+1)]\Gamma^T - E[\hat{\mathbf{x}}_p(n+1|n)\hat{\mathbf{x}}_p^T(n+1|n+1)]\Gamma^T \\ &\quad - E[\hat{\mathbf{x}}_p(n+1|n)\tilde{\mathbf{x}}_p^T(n+1|n+1)]\Gamma^T - \Gamma E[\mathbf{x}_p(n+1)\mathbf{x}_p^T(n+1)]\Gamma^T \\ &= E[\mathbf{x}_p(n+1)\mathbf{x}_p^T(n+1)]\Gamma^T - E[\hat{\mathbf{x}}_p(n+1|n)\hat{\mathbf{x}}_p^T(n+1|n+1)]\Gamma^T \\ &\quad - \Gamma E[\mathbf{x}_p(n+1)\mathbf{x}_p^T(n+1)]\Gamma^T \\ &= E[\mathbf{x}_p(n+1)\mathbf{x}_p^T(n+1)]\Gamma^T - \Phi_p E[\hat{\mathbf{x}}_p(n|n)\hat{\mathbf{x}}_p^T(n+1|n+1)]\Gamma^T \\ &\quad - \Gamma E[\mathbf{x}_p(n+1)\mathbf{x}_p^T(n+1)]\Gamma^T \end{aligned} \quad (8)$$

ここで、過去の時刻における $P_p(n|n)$ と $\hat{\mathbf{x}}_p(n|n)$ の全要素が $Q_p(n+1)$ の影響を受けているものと仮定し、 $Q_p(n+1)$ 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せば、式(8)の $Q_p(n+1)$ の要素は次式のようになる。

$$\begin{aligned} Q_p(n+1) &= \begin{bmatrix} E[d(n+1)d(n+1)] \\ E[d(n)d(n)] \\ \vdots \\ \vdots \\ E[d(n-L_p+2)d(n-L_p+2)] \end{bmatrix} O \\ &- \begin{bmatrix} 0 \\ E[\hat{d}(n|n)\hat{d}(n+1|n+1)] \\ E[\hat{d}(n-1|n-1)\hat{d}(n|n)] \\ \vdots \\ E[\hat{d}(n-L_p+2|n-L_p+2)\hat{d}(n-L_p+3|n-L_p+3)] \end{bmatrix} O \\ &- \begin{bmatrix} E[d(n+1)d(n+1)] \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix} O = \begin{bmatrix} 0 \\ e_2(n) \\ e_3(n) \\ \vdots \\ e_{L_p}(n) \end{bmatrix} O \end{aligned} \quad (9)$$

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(9) $Q_p(n+1)$ が表1手順1の $P_p(n+1|n)$ に与える影響は、

$$\begin{aligned}
 P_p(n+1|n) &= \Phi_p P(n|n) \Phi_p^T + R_{\delta_p}(n+1) + Q_p(n+1) + Q_p^T(n+1) \\
 &= \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & P_p^{L_p-1, L_p-1}(n|n) & \\ 0 & & & \end{bmatrix} \\
 &+ \begin{bmatrix} E[d(n+1)d(n+1)] & & & \\ & 0 & & \\ & \vdots & & \\ & 0 & & \end{bmatrix} \begin{bmatrix} O \\ O \\ O \\ O \end{bmatrix} + \begin{bmatrix} 0 \\ e_r(2) \\ e_r(3) \\ \vdots \\ e_r(L_p) \end{bmatrix} \begin{bmatrix} O \\ O \\ O \\ O \\ O \end{bmatrix} \\
 &+ \begin{bmatrix} 0 & e_r(2) & e_r(3) & \cdots & e_r(L_p) \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{bmatrix} \begin{bmatrix} O \\ O \\ O \\ O \\ O \end{bmatrix} \quad (10)
 \end{aligned}$$

$$= \begin{bmatrix} E[d(n+1)d(n+1)] & e_r(2) & e_r(3) & \cdots & e_r(L_p) \\ e_r(2) & & & & \\ e_r(3) & & & & \\ \vdots & & & & \\ e_r(L_p) & & & & \end{bmatrix} \begin{bmatrix} P_p^{L_p-1, L_p-1}(n|n) \\ \\ \\ \\ \end{bmatrix}$$

となる .

同様に、表 1 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ようになる .

$$\begin{aligned}
 k_p(n+1) &= P_p(n+1|n) \mathbf{m}_p \cdot \{ \mathbf{m}_p^T P_p(n+1|n) \mathbf{m}_p + r_{e_p}(n+1) \}^{-1} \\
 &= \begin{bmatrix} E[d(n+1)d(n+1)] \\ e_r(2) \\ e_r(3) \\ \vdots \\ e_r(L_p) \end{bmatrix} \cdot \{ E[d(n+1)d(n+1)] + \sigma_v^2 \}^{-1} \\
 &= \begin{bmatrix} k_p(n+1)[1, 1] \\ k_p(n+1)[2, 1] \\ \vdots \\ k_p(n+1)[L_p, 1] \end{bmatrix} \quad (11)
 \end{aligned}$$

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 1 手順 3 の $\hat{x}_p(n+1|n)$ は

$$\hat{x}_p(n+1|n) = \Phi_p \hat{x}_p(n|n) = \begin{bmatrix} 0 \\ \hat{x}_p(n|n)[1, 1] \\ \hat{x}_p(n|n)[2, 1] \\ \vdots \\ \hat{x}_p(n|n)[L_p-1, 1] \end{bmatrix} \quad (12)$$

となる .

それゆえ、表 1 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 (11) の $k_p(n+1)$ と (12) の $\hat{x}_p(n+1|n)$ が含まれることより、求める状態量ベクトル $\hat{x}_p(n+1|n+1)$ は次式で表される .

$$\begin{aligned}
 \hat{x}_p(n+1|n+1) &= \hat{x}_p(n+1|n) + \mathbf{k}_p(n+1) \cdot \{ y_p(n+1) - \mathbf{m}_p^T \hat{x}_p(n+1|n) \}
 \end{aligned}$$

$$\begin{aligned}
 &= \begin{bmatrix} 0 \\ \hat{x}_p(n|n)[1, 1] \\ \hat{x}_p(n|n)[2, 1] \\ \vdots \\ \hat{x}_p(n|n)[L_p-1, 1] \end{bmatrix} + \begin{bmatrix} k_p(n+1)[1, 1] \\ k_p(n+1)[2, 1] \\ \vdots \\ k_p(n+1)[L_p, 1] \end{bmatrix} \cdot \{ y_p(n+1) - 0 \} \\
 &= \begin{bmatrix} \hat{x}_p(n+1|n+1)[1, 1] \\ \hat{x}_p(n+1|n+1)[2, 1] \\ \vdots \\ \hat{x}_p(n+1|n+1)[L_p, 1] \end{bmatrix} \quad (13)
 \end{aligned}$$

従って、推定された音声信号は式 (12) 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 1 手順 5 の $P_p(n+1|n+1)$ は、式 (4) の手順 1 と式 (11) の手順 2 の結果を用いて $Q_p(n+1)$ の影響を表すと

$$\begin{aligned}
 P_p(n+1|n+1) &= \{ I - \mathbf{k}_p(n+1) \mathbf{m}_p^T \} P_p(n+1|n) \\
 &= \left\{ \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} - \begin{bmatrix} k_p(n+1)[1, 1] \\ k_p(n+1)[2, 1] \\ \vdots \\ \vdots \\ k_p(n+1)[L_p, 1] \end{bmatrix} \begin{bmatrix} O \\ O \\ O \\ O \\ O \end{bmatrix} \right\} \\
 &= \begin{bmatrix} E[d(n+1)d(n+1)] & e_r(2) & e_r(3) & \cdots & e_r(L_p) \\ e_r(2) & & & & \\ e_r(3) & & & & \\ \vdots & & & & \\ e_r(L_p) & & & & \end{bmatrix} \cdot \begin{bmatrix} P_p^{L_p-1, L_p-1}(n|n) \\ \\ \\ \\ \end{bmatrix} \\
 &= \begin{bmatrix} \alpha E[d(n+1)d(n+1)] & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{bmatrix} \quad (\alpha = 1 - \mathbf{k}_p(n+1)[1, 1]) \quad (14)
 \end{aligned}$$

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により、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、新たに提案した有色性駆動源を持つ状態空間モデルの有効性が確認できる .

5 計算機シミュレーション

提案手法の有効性を明らかにするために、文献 [1]、文献 [3]、文献 [4] と提案手法の性能比較をおこなう .

5.1 シミュレーション条件

以下の条件下でシミュレーションを実行する

- (A) 音声信号 $d(n)$: 成人男性と成人女性の音声は無音室にて録音した音声を 16kHz でサンプリングした音声信号 .

表 2: 有色雑音に対する雑音抑圧能力 SNR_{out}

SNR _{in} [dB]	L_{c1}, L_{c2} L_{c3}, L_p	Conv.1[3]	Conv.2[1]	Conv.3[4]	Prop.
0	20	-0.032	1.66	1.71	2.70
	30	-0.012	1.66	2.71	3.32
	40	-0.014	1.61	3.36	3.78
	50	-0.010	1.61	4.07	4.32
5	20	-0.013	6.12	6.25	6.89
	30	-0.020	6.12	7.20	7.72
	40	-0.004	6.13	7.93	8.23
	50	-0.023	6.14	8.43	8.59

表 4: 各手順における乗算回数

Computing or Solving	Multiplications		
	conv.1[3]	conv.3[4]	Prop.1
手順 1	$2L_{c1}^2$	0	0
手順 2	L_{c1}	$L_{c2} - 1$	1
手順 3	L_{c1}	0	0
手順 4	L_{c1}	$L_{c2} - 1$	1
手順 5	L_{c1}^2	$L_{c2} - 1$	1
Total	$3L_{c1}^2 + 3L_{c1}$	$3L_{c2} - 3$	3

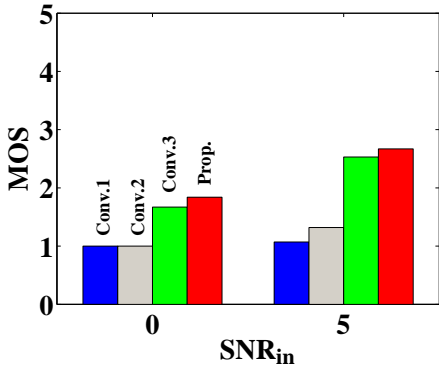


表 3: MOS 評価基準

評価値	基準	基準の説明
5	とても良い	まったく差が無い
4	良い	わずかに差がある
3	普通	差がある
2	劣る	かなり差がある
1	悪い	非常に差がある

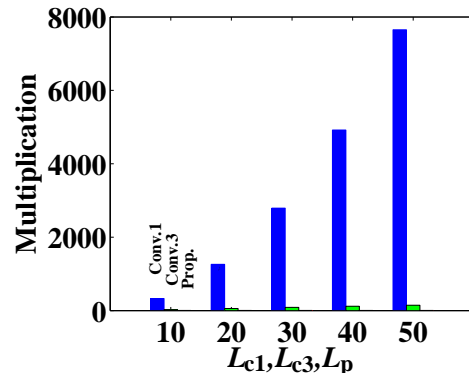


図 1: パブル雑音における MOS 評価

図 2: 表 4 に基づいた演算量比較 (総乗算回数)

(B) 雑音 $v(n)$: 分散 σ_v^2 のパブル雑音 [7] とし, 信号対雑音比が $\text{SNR}_{in} = 10 \log_{10} \left[\frac{1}{N} \sum_{n=1}^N d^2(n) \right]$ [dB] となるような雑音. ただし, 本論文では雑音の分散値は既知としている.

なお, 以下の結論で用いる L_{c1}, L_{c2} は文献 [3] と文献 [1] の AR 係数の次数もしくは状態ベクトルのサイズであり, L_{c3}, L_p は文献 [4] と提案手法の状態ベクトルのサイズである.

5.2 シミュレーション結果

上記で述べたシミュレーション条件に従い, 文献 [3] を従来手法 1, 文献 [1] を従来手法 2, 文献 [4] を従来手法 3 とし提案手法との客観評価, 主観評価および演算量評価の観点から性能比較を行った. それぞれの結果は, 表 2, 図 1 および図 2 に示す.

5.2.1 客観評価

客観評価として, クリアな音声信号 $d(n)$ と推定された音声信号 $\hat{d}(n)$ との差を用いて雑音抑圧能力を $\text{SNR}_{out} = 10 \log_{10} \frac{\sum_{n=1}^N \hat{d}(n)^2}{\sum_{n=1}^N \{d(n) - \hat{d}(n)\}^2}$ [dB] で評価した結果を表 2 に示す.

表 2 から, 提案手法は各従来手法 1~3 よりも高い SNR_{out} を得ていることより, 提案手法の有効性が確認できる. ここで L に着目すると, 従来手法 1 の L_{c1} と従来手法 2 の L_{c2} は大きさを変えても SNR_{out} はほぼ同じ値を取るが, カルマンフィルタのみを用いた従来手法 3 の L_{c3} と提案手法の L_p は L の値が大きいくほど SNR_{out} が大きくなる. これは, 従来手法 1 と従来手法 2 の AR 次数の決定が困難であることが影響していると考えられる.

5.2.2 主観評価

主観評価では, MOS (Mean Opinion Score) を用いて評価する. これは, 表 3 に示す 5 段階評価を用い, 推定信号 $\hat{d}(n)$ を複数人 (25 人) に聞き取り評価を調査してもらい, その平均値を評価量としたもので結果を図 1 に示す.

図 1 から, 0[dB], 5[dB] どちらの場合においても提案手法は従来手法より客観評価と同様に高い評価を得ており, 提案手法の有効性が確認できる.

以上のことより, 提案手法は実環境に近い有色性雑音でも音声信号 $d(n)$ の音質を犠牲にすることなく実用的な雑音抑圧法といえる.

5.2.3 演算量評価

カルマンフィルタを用いた雑音抑圧手法である従来手法 1[3] および従来手法 3[4] 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 に示し, その総乗算回数を演算量として比較した結果を図 2 に示す. 表 4 と図 2 より, 提案手法は他の従来手法 1, 従来手法 2 よりも少ない演算量で高い雑音抑圧能力を得ている. また, 提案手法は L_p のサイズに関係なく乗算回数が 3 回である. これにより, 提案手法は大幅な演算量の軽減を実現している.

6 まとめ

本論文は文献 [4] を拡張することで, 演算量が少ない実用的な雑音抑圧法を提案した. また, 新たに設定した有色性駆動源を持つ提案手法の状態空間モデルがカルマンフィルタ理論に適用可能であることについても定量的に解析した. 提案手法の有効性は, 客観評価と主観評価および演算量評価の観点から従来手法と比較することで, 提案手法が実用的な雑音抑圧法であることを明らかにした.

今後の課題として, 雑音分散 σ_v^2 の推定精度が雑音抑圧能力に与える影響を解析することなどがあげられる.

参考文献

- [1] 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
- [2] 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.
- [3] W.Kim, "Noise Variance Estimation for Kalman Filtering of Noisy Speech" IEICE TRANS. INF. & SYST., VOL.E84-D, NO.1 JANUARY 2001
- [4] 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.
- [5] 谷萩 隆嗣, カルマンフィルタと適応信号処理, コロナ社, 2005.
- [6] J.Huang and Y.Zhao, "An Energy-Constrained Signal Sub-Space Method for Speech Enhancement and Recognition in White and Colored Noises," Speech Communication, vol.26, no.26 pp.165-181, 1998.
- [7] M.Kato, A.Sugiyama, and M.Serizawa, "Noise suppression with high speech quality based on weighted noise estimation and MMSE-TSA," IEICE Trans. Fundamentals (Japanese Edition), vol.J87-A, no.7, pp.851-860, July 2004.