

カルマンフィルタによる単一マイク形ノイズキャンセラにおける パラメータ推定精度と収束特性の関係

Effects of Estimation Error of Parameters on Convergence Property
in A Noise Canceller

Consisting of Kalman Filter and A Single Microphone

辻 祐亮† 中山 謙二‡ 平野 晃宏‡
Yusuke Tsuji Kenji Nakayama Akihiro Hirano

†金沢大学大学院 自然科学研究科 電子情報システム専攻
Graduate School of Nature Science & Technology, Kanazawa Univ.

‡金沢大学 工学部 情報システム工学科
Dept. of Information and Systems Eng., Faculty of Eng., Kanazawa Univ.
E-mail: nakayama@t.kanazawa-u.ac.jp

あらまし

カルマンフィルタを用いて単一マイク形ノイズキャンセラを構成する場合に、カルマンフィルタのパラメータの推定精度と、ノイズキャンセラとしての収束特性の関係を理論的、およびシミュレーションにより解析する。

カルマンフィルタを用いるためには、状態遷移行列 $A(n+1, n)$ 、音声の生成に用いる入力信号の共分散 $Q(n)$ 、観測過程における雑音の共分散 $R(n)$ という3つのパラメータが必要である。これらのパラメータは実環境では雑音が混在した信号を用いて推定するため精度が落ち、カルマンフィルタの精度も劣化する。そこで、それぞれのパラメータの推定精度が、ノイズキャンセラにどのような影響を与えるのかについて解析を行なった。その結果、ノイズキャンセラの収束特性において、 $Q(n)$ の誤差による劣化が最も大きく、次に $A(n+1, n)$ の誤差の影響が大きく、 $R(n)$ の誤差の影響はほとんどないことが明らかになった。

Abstract

In this paper, effects of estimation error of parameters used in a noise canceller consists of the Kalman filter and a single microphone are analyzed in theoretically and numerically.

In this type of noise cancellers, three kinds of parameters should be estimated. They include a state transition matrix A , covariance matrices of the signal

Q and the noise R . These parameters are estimated using the noisy signal. The estimation error degrade noise cancellation performance. Effect of the estimation error in Q is most severe, and that of R is almost negligible. Through several simulations, these properties are confirmed. Some methods are introduced in order to suppress these effects. SNR is improved by around 3dB.

1 まえがき

現在、携帯電話などの移動体通信機器は、一人一台持つのが当たり前の世の中になっている。そのため利用者が増加し、街頭や車内など雑音の大きい場所で使用される機会が増えている。そのような場所では、雑音が利用者の会話中にマイクに拾われると、相手側では会話聞きづらくなり快適な通話の妨げとなる。そこで雑音を除去するためにノイズキャンセラが必要となる。その方式には単一マイクの方式、つまり音声と雑音が同時に一つのマイクから入力される場合と、複数のマイクを用いる方式、つまり音声と雑音がそれぞれ別々に入力される場合とがある。携帯電話などを対象とした場合には、小型化・低消費電力化の点から、単一マイク方式の方が望ましい。

また現在すでに様々なノイズキャンセラの方式が提案されているが、ETSIが定める3GPPにおける携帯電話用雑音抑圧技術の標準化において、絶対的に優位なアルゴリズムが存在せず、その性能に関する要求条件とその

評価方法のみを定めている [1][2][3]。

本稿では、単一マイク方式のノイズキャンセラを対象とし、ノイズキャンセラにはモデル推定形のカルマンフィルタを用いる。カルマンフィルタを用いて単一マイク方式のノイズキャンセラを実現するためには3つのパラメータが必要である。そこで、それら3つのパラメータの推定精度とノイズキャンセラとしての収束特性との関係について解析する。

2 カルマンフィルタ

2.1 カルマンフィルタの構成

音声 $x(n)$ の生成過程を状態方程式で示す。

$$x(n+1) = A(n+1, n)x(n) + v_1(n) \quad (1)$$

さらに観測過程を次のように示す。

$$y(n) = C(n)x(n) + v_2(n) \quad (2)$$

ここで、 $A(n+1, n)$ はシステムの時刻 n と $n+1$ を関係づける状態遷移行列、 $C(n)$ は実際に観測されるデータ $y(n)$ に含まれる信号成分を決定する観測行列である。

音声 $x(n)$ に雑音 $v_2(n)$ が混入した観測信号 $y(n)$ から、音声を図1のカルマンフィルタで推定する[7]。カルマンフィルタとは、信号の音声モデルが正確にわかっている場合に、音声の振幅特性を最適に推定することができるフィルタである。

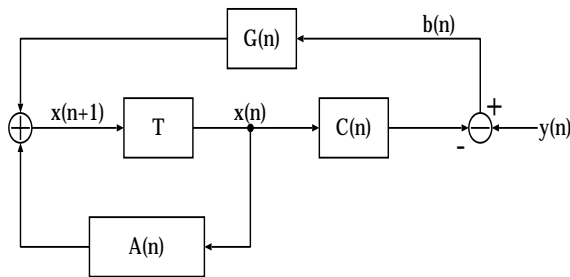


図1: カルマンフィルタの構成

図1より、状態の推定は次式によって更新される。

$$x(n+1|y_n) = A(n+1, n)x(n|y_{n-1}) + G(n)b(n) \quad (3)$$

ここで、カルマンゲイン $G(n)$ 、イノベーション $b(n)$ は次式に示すように巡回的に更新される。

$$S(n) = C(n)E(n, n-1)C^T(n) + R(n) \quad (4)$$

$$G(n) = A(n+1, n)E(n, n-1)C^T(n)S^{-1}(n) \quad (5)$$

$$b(n) = y(n) - C(n)\hat{x}(n|y_{n-1}) \quad (6)$$

$$E(n) = E(n, n-1) - A(n, n+1)G(n)C(n)E(n, n-1) \quad (7)$$

$$E(n+1, n) = A(n+1, n)E(n)A^T(n+1, n) + Q(n) \quad (8)$$

ここで、

$$Q(n) = E[v_1(n)v_1^T(n)] \quad (9)$$

$$R(n) = E[v_2(n)v_2^T(n)] \quad (10)$$

初期条件として、

$$x(1|y_0) = 0 \quad (11)$$

$$E(1, 0) = E[x(1)x^T(1)] \quad (12)$$

式(3)から式(10)までを巡回的に解くことにより、信号 $x(n+1)$ を最小二乗近似の意味で最適に予測するフィルタが形成される。

2.2 カルマンフィルタのパラメータ推定

カルマンフィルタを用いるためには、次の3つのパラメータ推定が必要である。

1. 状態遷移行列 $A(n+1, n)$
2. 音声の生成に用いる入力信号 $v_1(n)$ の共分散行列 $Q(n)$
3. 観測過程における雑音 $v_2(n)$ の共分散行列 $R(n)$

それぞれの推定方法を次に示す。

2.2.1 状態遷移行列 $A(n+1, n)$

状態遷移行列 $A(n+1, n)$ の推定にはLevinson-Durbin Algorithm[7]を用いる。

Levinson-Durbin Algorithmは、 M 次の予測誤差フィルタの係数が与えられれば、 $M+1$ 次の予測誤差フィルタの係数を計算できる。今、音声モデルの次数と同じ次数まで予測誤差フィルタを構成する。本稿では音声信号を20次のARモデルで近似しているため、予測誤差フィルタの次数も20次となる。

2.2.2 共分散行列 $Q(n)$

音声の生成に用いる入力信号 $v_1(n)$ の推定には、音声の生成過程の逆のフィルタ処理を行なうことにより推定する。

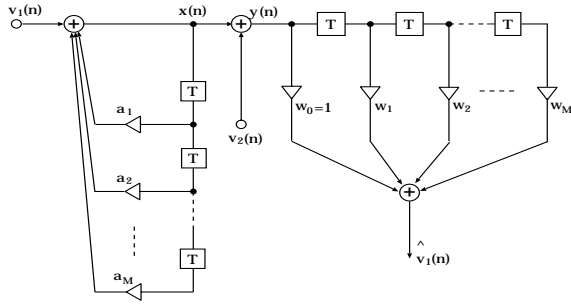


図 2: 音声の生成過程の逆のフィルタ処理

今、 $v_1(n)$ を白色雑音、 $v_2(n)$ をゼロとすると、

$$\frac{\sum_{i=0}^M w_i z^{-i}}{1 - \sum_{i=1}^M a_i z^{-i}} = 1, w_0 = 1 \quad (13)$$

となるため、

$$w_i = -a_i, 1 \leq i \leq M \quad (14)$$

とすることができる。そのため $\hat{v}_1(n)$ は $v_1(n)$ となるはずである。

2.2.3 共分散行列 $R(n)$

観測過程における雑音 $v_2(n)$ には白色雑音を用いる。

3 ノイズキャンセラ方式

上記の方法によりパラメータを推定し、それをカルマンフィルタに適用する。カルマンフィルタによる単一マイク形ノイズキャンセラの構成を図 3 に示す。

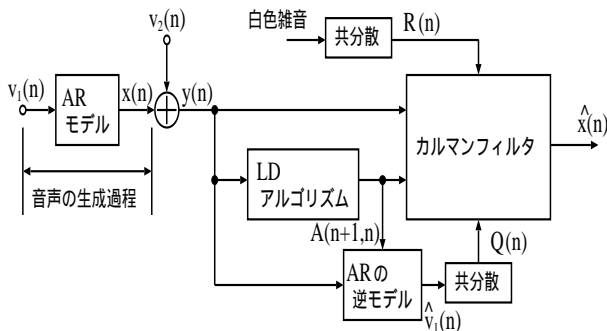


図 3: カルマンフィルタによる単一マイク形ノイズキャンセラ

本稿では図 3 の $x(n)$ には図 4 の音声信号を、 $v_2(n)$ には図 5 の白色雑音、または図 6 のパブル雑音を用いた。

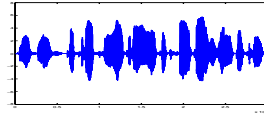


図 4: 音声信号

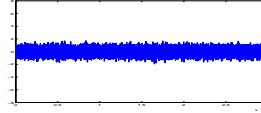


図 5: 白色雑音

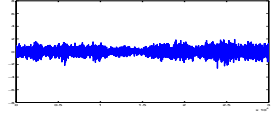


図 6: バブル雑音

これらより、3 つのパラメータそれぞれについての解析を行なう。

3.1 状態遷移行列 $A(n+1,n)$

状態遷移行列を構成する AR パラメータが、音声信号 $x(n)$ から Levinson-Durbin Algorithm により推定できたとする場合と、実環境に従い $y(n)$ を用いて推定した場合とで比較する。

3.2 共分散行列 $Q(n)$

音声信号から推定された $A(n+1,n)$ と $x(n)$ を用いて推定した共分散 $Q(n)$ を図 7 に、白色雑音が混在した信号から推定した $A(n+1,n)$ と $y(n)$ を用いて推定した共分散 $Q(n)$ を図 8 に、パブル雑音が混在した信号から推定した $A(n+1,n)$ と $y(n)$ を用いて推定した共分散 $Q(n)$ を図 9 に示す。

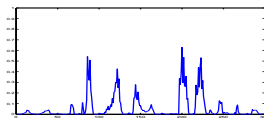


図 7: 理想的に推定した $Q(n)$

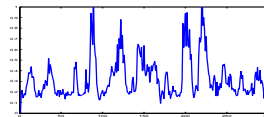


図 8: 実環境で推定した $Q(n)$ (白色雑音)

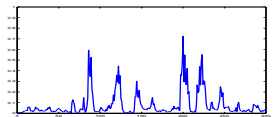


図 9: 実環境で推定した $Q(n)$ (パブル雑音)

式 (7)、(8) からわかるように $Q(n)$ から $E(n+1,n)$ まで伝達特性はループを構成している。従って、 $Q(n)$ における誤差の影響はループ利得だけ増幅される。

これに対しては、以下の 3 通りの方法により $Q(n)$ における誤差の影響を低減する方法を検討する。

1. 共分散行列 $Q(n)$ を定数倍することにより、共分散行列 $Q(n)$ の大きさを制御する (方法 1)

- 共分散行列 $Q(n)$ から共分散行列 $R(n)$ の定数倍を引き算することにより、共分散行列 $Q(n)$ における雑音の影響を低減する (方法 2)
- 式 (7)、(8) におけるループ利得を定数倍することにより低減し、 $Q(n)$ の誤差の影響が $E(n+1,n)$ で増幅されることを抑える (方法 3)

$$E(n+1, n) = \alpha A(n+1, n)E(n)A^T(n+1, n) + Q(n) \quad (15)$$

3.3 共分散行列 $R(n)$

白色雑音・パブル雑音それぞれの共分散 $R(n)$ を図 10、図 11 に、独立な白色雑音を用いて推定した $R(n)$ を図 12 に示す。

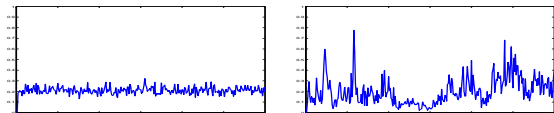


図 10: 白色雑音

図 11: パブル雑音

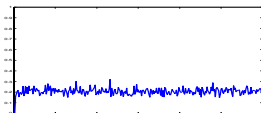


図 12: 独立な白色雑音を用いた $R(n)$

式 (3)、(4)、(5) より、 $R(n)$ が大きくなると $S(n)$ が大きくなり、 $G(n)$ が小さくなるため、式 (3) の修正項が小さくなり、結果的に他のパラメータの推定のずれの影響を抑えるような働きがあるのではないかと考えられる。

そこで図 12 の白色雑音の大きさを変化させ、解析する。

4 シミュレーション結果

本稿では、すべてのシミュレーションにおいて信号対雑音電力 (SNR) を約 15[dB] とした。また雑音の種類は、図 5、図 6 の白色雑音とパブル雑音の 2 種類に対してシミュレーションを行なった。

4.1 3つのパラメータを理想的に推定

3つのパラメータが全て理想的に推定できたとする場合の推定音声波形を図 13、図 14 に示す。

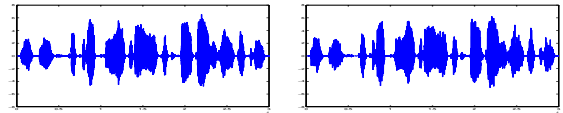


図 13: 理想的に推定 (白色雑音)SNR25.5[dB] 図 14: 理想的に推定 (パブル雑音)SNR22.1[dB]

どちらの場合も雑音が除去できていることがわかる。

4.2 3つのパラメータを実環境で推定

3つのパラメータを全て実環境で推定した場合の推定音声波形を図 15、図 16 に示す。

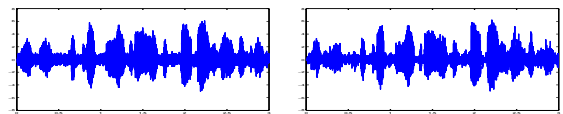


図 15: 実環境で推定 (白色雑音)SNR19.0[dB] 図 16: 実環境で推定 (パブル雑音)SNR17.2[dB]

どちらの場合も雑音が十分に除去できず、十分な推定ができたとは言えない。

そこで、3つのパラメータが推定にどれほど影響を与えているのかについて解析する。

4.3 $A(n+1,n)$ の推定誤差の影響

状態遷移行列 $A(n+1,n)$ の推定のずれの影響を調べるために、 $Q(n)$ と $R(n)$ の推定は理想的で、 $A(n+1,n)$ のみ実環境で推定した場合の推定音声波形の SNR は、白色雑音の場合 21.0[dB]、パブル雑音の場合 21.4[dB] であった。

4.4 $Q(n)$ の推定誤差の影響

$A(n+1,n)$ と $R(n)$ の推定は理想的で、 $Q(n)$ のみ実環境で推定した場合の推定音声波形の SNR は、白色雑音の場合 18.3[dB]、パブル雑音の場合 18.8[dB] であり、3つのパラメータのうち、最も推定精度が与える影響が大きいことがわかる。

4.4.1 方法 1 ($\alpha \times Q$) による改善

α を 0.1 から 1.0 まで変化させ、それぞれに対してシミュレーションを行なった。

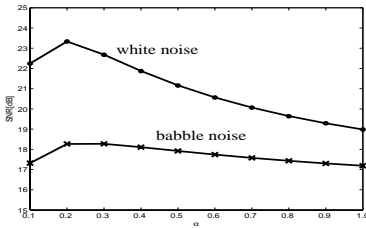


図 17: 方法 1 による α と SNR との関係

図 17 より、白色雑音、パブル雑音ともに $\alpha = 0.2$ の時が最も良い結果となり、その時の $Q(n)$ と推定音声波形を図 18、図 19、図 20、図 21 に示す。

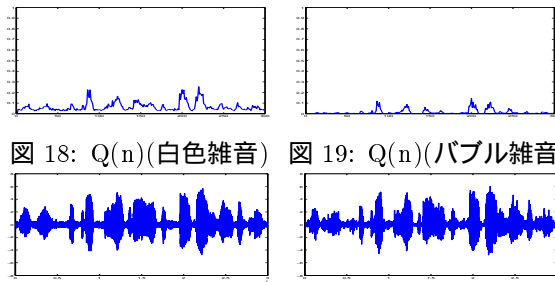


図 20: 推定音声波形 (白色雑音) SNR23.3[dB] 図 21: 推定音声波形 (パブル雑音) SNR18.3[dB]

4.4.2 方法 2 ($Q-\alpha \times R$) による改善

α を 0.0 から 1.0 まで変化させ、それぞれに対してシミュレーションを行なった。

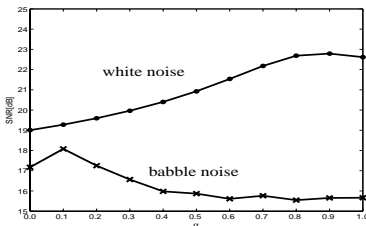


図 22: 方法 2 による α と SNR との関係

図 22 より、白色雑音においては $\alpha = 0.9$ の時、パブル雑音においては $\alpha = 0.1$ の時が最も良い結果となり、その時の $Q(n)$ と推定音声波形を図 23、図 24、図 25、図 26 に示す。

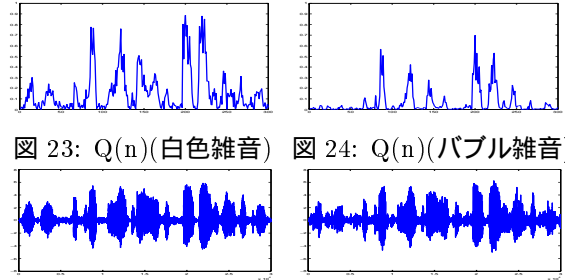


図 23: $Q(n)$ (白色雑音) 図 24: $Q(n)$ (パブル雑音)
図 25: 推定音声波形 (白色雑音) SNR22.8[dB] 図 26: 推定音声波形 (パブル雑音) SNR18.1[dB]

4.4.3 方法 3 ($\alpha \times$ ループ利得) による改善

α を 0.1 から 1.0 まで変化させ、それぞれに対してシミュレーションを行なった。

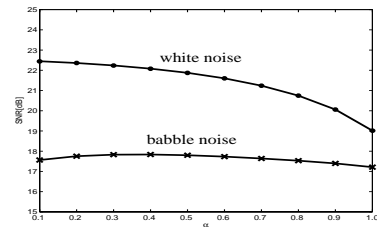


図 27: 方法 3 による α と SNR との関係

図 27 より、白色雑音においては $\alpha = 0.1$ の時、パブル雑音においては $\alpha = 0.4$ の時が最も良い結果となり、その時の推定音声波形を図 28、図 29 に示す。

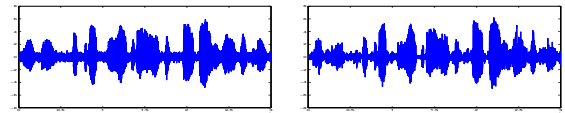


図 28: 推定音声波形 (白色雑音) SNR22.4[dB] 図 29: 推定音声波形 (パブル雑音) SNR17.8[dB]

4.5 $R(n)$ の推定誤差の影響

$A(n+1, n)$ と $Q(n)$ の推定は理想的で、 $R(n)$ のみ実環境で推定した場合の推定音声波形の SNR は、白色雑音の場合 25.5[dB]、パブル雑音の場合 21.7[dB] であり、他のパラメータの推定が正確であれば、 $R(n)$ の影響はほとんどないことがわかる。

また、 $R(n)$ を推定する白色雑音の大きさを変化させてシミュレーションを行なった。

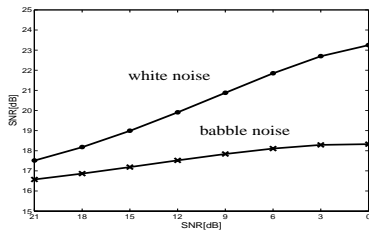


図 30: $R(n)$ の推定における α と SNR との関係

SNR を 0[dB] から 21[dB] まで 3[dB] ずつ変化させてシミュレーションを行なった結果、SNR を大きくするほど、つまり 0[dB] に近付くほど良い結果が得られた。その時の SNR は白色雑音においては 23.3[dB]、バブル雑音においては 18.3[dB] であった。

5 まとめ

単一マイク形ノイズキャンセラにカルマンフィルタを用いるために必要となる 3 つのパラメータについて解析を行なった。3 つのパラメータのうち、 $Q(n)$ の影響が最も大きく、次に $A(n+1, n)$ の影響が大きく、 $R(n)$ の影響はほとんどないことがわかった。

また $Q(n)$ 、 $R(n)$ の推定誤差の影響を低減する方法について考察し、その結果特に白色雑音については SNR を向上でき、推定誤差の影響を低減できた。しかし、バブル雑音などの有雑音に対しての効果は十分とは言えず、今後の課題として残されている。また $A(n+1, n)$ についての解析もまだまだ不十分である。

今後はさらに誤差の影響を低減する方法や、 α の値を適応的に制御する方法の検討が必要である。

また、3GPP が定める携帯電話用雑音抑圧技術の基準をクリアできるような性能を得るための新たな推定方法などが必要である。

参考文献

- [1] 杉山昭彦, 芹沢昌宏, 加藤正徳: “携帯電話用雑音抑圧技術の標準化”, 信学技報, SP2002-1, pp.1-6, Apr. 2002.
- [2] 古田訓, 高橋真哉: “3GPP 性能基準クリア AMR 音声コーデック用ノイズサプレッサとその評価結果”, 第 17 回 DSP シンポジウム, A5-2, 2002-11.
- [3] 加藤正徳, 杉山昭彦, 芹沢昌宏, 長谷川厚志, 水越明哉: “重み付き雑音推定に基づくノイズサブ

レッサの DSP 実現と 3GPP 評価”, 第 17 回 DSP シンポジウム, A5-3, 2002-11.

- [4] Masakiyo Fujimoto and Yasuo Ariki: “Noisy Speech Recognition Using Noise Reduction Method Based on Kalman Filter”, ICASSP2000, Vol.III, pp.1723-1726, 2000-06.
- [5] 藤本 雅清, 有木 康雄: “雑音環境下での音声認識を目的とした高速カルマンフィルタによる雑音除去法の提案”, 日本音響学会, 平成 12 年度春季研究発表会, 1-8-3, pp.5-6, 2000-03.
- [6] 藤本 雅清, 有木 康雄: “カルマンフィルタを用いた雑音環境下における音声認識の検討 - 雑音適応と雑音除去 -”, 日本音響学会, 平成 11 年度秋季研究発表会, 1-1-16, pp.31-32, 1999-09.
- [7] S.Haykin: “Adaptive Filter Theory.3rd Edition”, Prentice Hall Inc, 1996
- [8] S. ヘイキン: “適応フィルタ入門”, 現代工学社, 1987
- [9] K.K.Paliwal and A.Basu.: “A speech enhancement method based on Kalman filtering.”, in Proc.IEEE Int.Conf.Acoust.,Speech,Signal Processing(Dallas,TX).pp177-180.Apr.6-9,1987.
- [10] 竹内 康弘: “カルマンフィルタによる音声コーデックの雑音抑制”, 金沢大学 電気・情報工学科 卒業研究論文 1998-2.
- [11] 久保 慶太: “カルマンフィルタと音声強調による単一マイク・ノイズキャンセラ”, 金沢大学 修士論文 2000-2.