

学位申請論文

状態空間モデルに着目した 最小分散状態推定法とその応用

平成29年3月

名取 隆廣

目次

第1章	序論	3
1.1	本研究の背景	3
1.2	論文の構成	4
第2章	状態空間モデルを用いた状態量推定問題	5
2.1	緒言	5
2.2	ARモデルとカルマンフィルタを用いた状態量推定	5
2.2.1	問題設定	5
2.2.2	Step1:ARモデルを用いた状態量の定式化	6
2.2.3	Step2:カルマンフィルタを用いた状態量推定	6
2.2.4	ARモデルを用いたカルマンフィルタ手法の問題点	7
2.3	有色駆動源カルマンフィルタを用いた雑音抑圧法	8
2.4	雑音抑圧法と最小分散推定法	10
2.4.1	雑音抑圧法の解析	10
2.4.2	最小分散推定法 [2]	12
2.5	結言	12
第3章	MIMO MC-CDMA 通信におけるチャネルゲイン推定問題への適用	14
3.1	緒言	14
3.2	問題設定	16
3.2.1	送信信号の生成	17
3.2.2	受信信号の生成	19
3.3	ARモデルとカルマンフィルタを用いたチャネルゲイン推定法	21
3.3.1	Step 1: AR係数の算出	21
3.3.2	Step 2:カルマンフィルタを用いたチャネルゲイン推定	22
3.3.3	計算機シミュレーション	23
3.4	雑音抑圧法を用いたチャネルゲイン推定手法	26
3.4.1	提案手法の状態空間モデルとチャネルゲイン推定アルゴリズム	26
3.4.2	雑音抑圧法を適用したチャネルゲイン推定法の解析	29
3.5	演算量評価	33
3.6	計算機シミュレーション	35
3.6.1	チャネルゲイン推定精度評価	35
3.6.2	シンボルエラーレート評価	38
3.7	結言	39

第4章	自動カラオケ音源生成システムへの応用	41
4.1	緒言	41
4.2	問題設定	43
4.3	従来の伴奏信号抽出手法	43
4.3.1	Step1:伴奏信号の抽出	44
4.3.2	Step2:差信号とカラオケ信号の生成	45
4.4	雑音抑圧法を用いた伴奏信号抽出手法	46
4.4.1	Step1:ヴォーカル信号の抽出	47
4.4.2	Step2:雑音抑圧法 [3] を用いたカラオケ信号生成法	49
4.5	計算機シミュレーション	54
4.5.1	スペクトログラムによるステレオ感の評価	54
4.5.2	SNR を用いた誤差評価	56
4.5.3	MOS 評価	57
4.6	アプリケーション実装	60
4.6.1	演算量評価	60
4.6.2	アプリケーション実装	61
4.7	結言	63
第5章	結論	65
	謝辞	66
	参考文献	67

第1章 序論

1.1 本研究の背景

本論文で扱う状態量とは、(i) 観測値からは直接得ることのできないシステム内部の状態量（例：加速度や速度）または、(ii) 外乱などの影響により汚された所望信号（例：音声信号や元画像など）を指す。これら状態量を直接得ることは稀であり、観測値に含まれる状態量を抽出する手法は様々なアプリケーションへの応用においても非常に重要である。

状態量を何らかの数学的手法を用いて推定する問題を状態量推定問題と呼ぶ。カルマンフィルタ [1] はこの問題を解く手法の 1 つであり、状態量の時間変化を表した状態方程式と状態量が外乱により汚され観測される過程を表した観測方程式により定式化された状態空間モデルを構成することにより状態量を推定する。

状態量の時間変化が各時刻間において相関を有する場合、一般的に状態量の時間変化は自己回帰 (AR: Auto Regressive) モデルを用いて表現され、自己回帰係数 (AR 係数) を算出することにより状態方程式を構成する。ここで、AR モデルに用いる次数 (AR 次数) の決定問題が生じる。即ち、AR 次数が AR 係数の推定精度を決定づけるため、誤った次数を設定した場合、最終的に得られる状態量の推定精度が劣化する。また状態量の性質によっては AR 係数そのものの算出が困難な場合がある。

この問題点に対して田邊らは、雑音抑圧のための AR モデルを用いない有色駆動源を用いた状態量推定方法 [3]（以降、雑音抑圧法と呼ぶ）を提案した。これは文字通り、AR モデルを用いない代わりに状態方程式の駆動源に状態量と相関をもつ物理量を用いることで、状態量の時間変化を表している。AR モデルを用いないことにより、AR 次数の決定問題が解決され AR 係数の算出も必要がないため、計算コストが削減されているものの、推定される状態量についての議論がなされてこなかった。

そこで本研究では雑音抑圧法が最小分散推定法 [2] に帰着されることを示す。また、この雑音抑圧法を MIMO MC-CDMA 通信方式におけるチャネル推定法および自動カラオケ

音源生成システムに応用した手法について議論する。

1.2 論文の構成

本論文の構成は次のとおりである。2章では、状態量の時間変化をARモデルを用いて定式化した場合におけるカルマンフィルタによる状態量推定を述べ、その問題点を示す。そして問題点を解決すべく、田邊らが提案した雑音抑圧法について述べた後に、この手法の状態空間モデルの構造に着目することにより、雑音抑圧法が最小分散推定法に帰着されることを示す。3章では、MIMO MC-CDMA 通信方式に対するチャンネルゲイン推定法として雑音抑圧法を応用した手法を提案する。提案手法により推定されたチャンネルゲインの推定精度は、移動体の移動速度に依存せずまた、少ない演算量で推定が可能であることを示す。4章では、雑音抑圧法を用いた自動カラオケ音源生成手法を提案する。提案手法はステレオのカラオケ音源がリアルタイムで生成可能なことを示す。

第2章 状態空間モデルを用いた状態量推定問題

2.1 緒言

本章では、状態空間モデルを用いた状態量推定問題について述べる。

2.2 ARモデルとカルマンフィルタを用いた状態量推定

本章では、状態量をARモデルで定式化しカルマンフィルタを用いて状態量を推定する手法について述べる。

ARモデルとカルマンフィルタを用いた状態量推定は以下のStepによって実行される。Step1では、時間変動する状態量をARモデルで定式化した後にAR係数を算出する。Step2では、Step1で算出される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を適用することで逐次的に状態量を推定している。

2.2.1 問題設定

以降の議論のために、クリアな音声信号 $d(n)$ に雑音 $v(n)$ が加わった観測信号 $y(n)$ を以下

$$y(n) = d(n) + v(n) \quad (2.1)$$

のように定義し、観測信号からクリアな音声信号のみを推定する問題 [4] を考える。ここで、雑音は白色ガウス雑音を仮定し平均値と分散値はそれぞれ $E[v(n)] = 0$, $E[v^2(n)] = \sigma_v^2$ とする。

2.2.2 Step1:AR モデルを用いた状態量の定式化

音声信号 $d(n)$ を AR システムでモデル化すると

$$d(n) = \sum_{\ell=1}^{L_c} \alpha_{\ell} d(n-\ell) + e(n) \quad (2.2)$$

のように与えられる. ここで $\alpha_{\ell}(n)$ は, ℓ 次 AR 係数, L_c は AR 次数, 及び $e(n)$ は駆動源 (予測誤差) で白色信号とする.

次に AR 係数は, 式 (3.10) の予測誤差 $e(n)$ の平均二乗誤差を最小にするという基準を用いて算出すれば

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(L_c - 1) \\ R(1) & R(0) & \cdots & R(L_c - 2) \\ \vdots & \vdots & \ddots & \vdots \\ R(L_c - 1) & R(L_c - 2) & \cdots & R(0) \end{bmatrix} \begin{bmatrix} \alpha_1(n) \\ \alpha_2(n) \\ \vdots \\ \alpha_{L_c}(n) \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(L_c) \end{bmatrix} \quad (2.3)$$

と表すことができる. ここで $R(\ell-k)$ は音声信号の自己相関関数であり, $E[d(n-k)d(n-\ell)]$ と定義される.

2.2.3 Step2:カルマンフィルタを用いた状態量推定

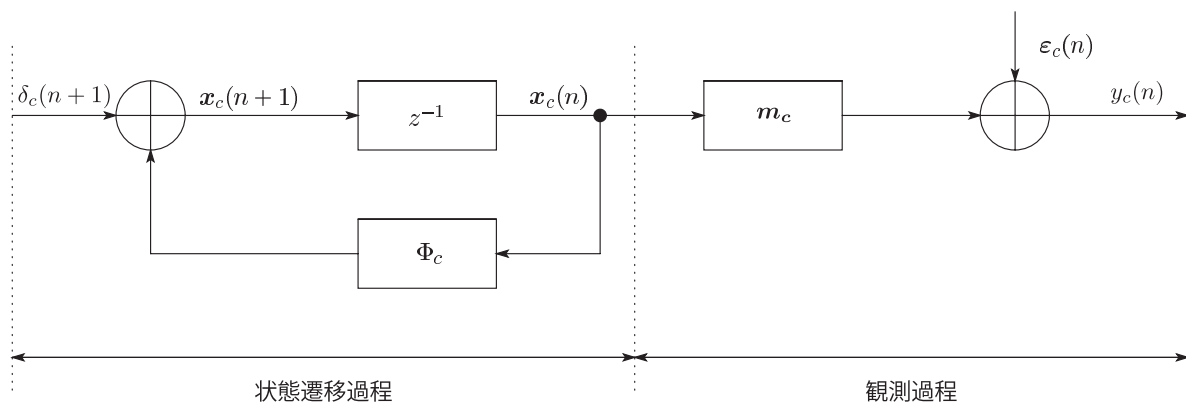


図 2.1: 従来手法の状態空間モデル

本節では Step1 で求めた AR 係数を用いて状態方程式と観測方程式からなる状態空間モデル (図 2.1) を構成し, こ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することで, 逐次的にクリアな音声信号を推定する手法について説明する.

L_c 次元状態ベクトル $\mathbf{x}_c(n)$ は, n 時刻から L サンプル過去の音声信号を並べたものとし

$$\mathbf{x}_c(n) = [d(n), d(n-1), \dots, d(n-L_c+1)]^T \quad (2.4)$$

のように定義される.

状態方程式は, 音声信号の時間変動を AR 係数を用いて表すので, 式 (3.10) を用いて

$$\text{[状態方程式]} \quad \mathbf{x}_c(n+1) = \Phi_c \mathbf{x}_c(n) + \mathbf{g}_c \delta_c(n+1) \quad (2.5)$$

のように与えられる. ここで, Step1 で求めた AR 係数を用いて $L_c \times L_c$ 行列の状態遷移行列 Φ_c は以下のように定義される.

$$\Phi_c = \begin{bmatrix} \alpha_1, \dots, \alpha_{L_c-1} & \alpha_{L_c} \\ I_{(L_c-1)} & \mathbf{0}_{(L_c-1) \times 1} \end{bmatrix} \quad (2.6)$$

また, L_c 次元駆動源ベクトル \mathbf{g}_c , 駆動源 δ_c は以下のように定義される.

$$\left. \begin{aligned} \mathbf{g}_c &= [1, 0, 0, \dots, 0]^T \\ \delta_c(n) &= e(n) \end{aligned} \right\} \quad (2.7)$$

次いで観測方程式は, クリアな音声信号 $d(n)$ が雑音 $v(n)$ の影響を受け劣化する過程を表したものであり, 式 (2.1) を用いることで, 次式のような観測方程式を得る.

$$\text{[観測方程式]} \quad y_c(n) = \mathbf{m}_c^T \mathbf{x}_c(n) + \varepsilon_c(n) \quad (2.8)$$

ここで L_c 次元観測ベクトル $\mathbf{m}_c(n)$, 観測雑音 $\varepsilon_c(n)$ を

$$\left. \begin{aligned} \mathbf{m}_c &= [1, 0, 0, \dots, 0]^T \\ \varepsilon_c(n) &= v(n) \end{aligned} \right\} \quad (2.9)$$

と定義している.

駆動源 $\delta_c(n)$ は白色信号でかつ, 状態ベクトル $\mathbf{x}_c(n)$ は駆動源と観測雑音 $\varepsilon_c(n)$ とは無相関という条件のもとで, 式 (2.5) と式 (2.8) をカルマンフィルタ理論に適用することにより, 逐次的にクリアな音声信号 $d(n)$ を推定している. 表 2.1 に推定アルゴリズムを示す.

2.2.4 AR モデルを用いたカルマンフィルタ手法の問題点

ここで, AR モデルを用いたカルマンフィルタ手法の問題点について述べる. この手法は, 音声信号の時間変動を AR システムで与えているため, AR 次数決定に起因する AR 係

表 2.1: AR モデルとカルマンフィルタを用いた雑音抑圧アルゴリズム

[Initialization]

$$\widehat{\boldsymbol{x}}_c(0|0) = \mathbf{0}, P_c(0|0) = I_{L_c}$$

$$R_{\delta_c}(n)[i, j] = \begin{cases} E \left[\left\{ y(n) - \sum_{\ell=1}^{L_c} \alpha_\ell d(n-\ell) \right\}^2 \right] - \sigma_v^2 & (i, j = 1) \\ 0 & (\text{other}) \end{cases}$$

$$r_{\epsilon_c}(n) = \sigma_v^2$$

[Iteration]

1. $P_c(n+1|n) = \Phi_c P_c(n|n) \Phi_c^T + R_{\delta_c}(n+1)$
2. $\boldsymbol{k}_c(n+1) = \left\{ P_c(n+1|n) \boldsymbol{m}_c^T \right\} \cdot \left\{ \boldsymbol{m}_c P_c(n+1|n) \boldsymbol{m}_c^T + r_{\epsilon_c}(n+1) \right\}^{-1}$
3. $\widehat{\boldsymbol{x}}_c(n+1|n) = \Phi_c \widehat{\boldsymbol{x}}_c(n|n)$
4. $\widehat{\boldsymbol{x}}_c(n+1|n+1) = \widehat{\boldsymbol{x}}_c(n+1|n) + \boldsymbol{k}_c(n+1) \cdot \left\{ y_c(n+1) - \boldsymbol{m}_c^T \widehat{\boldsymbol{x}}_c(n+1|n) \right\}$
5. $P_c(n+1|n+1) = \left\{ I - \boldsymbol{k}_c(n+1) \boldsymbol{m}_c^T \right\} P_c(n+1|n)$
6. $n = n + 1$ go back 1.

数精度劣化という問題が発生する。それゆえ、信頼性の低い AR 係数を式 (2.5) の状態方程式に用いた場合、音声信号の推定精度が劣化してしまう。また本章では音声信号を AR モデルで表したが、問題によっては AR 係数の算出が困難な場合が存在する。詳細は 3 章で述べることとする。

そこで次章では、AR システムのコンセプトを必要としない状態量推定法について述べる。

2.3 有色駆動源カルマンフィルタを用いた雑音抑圧法

本節では、状態量の変化を AR モデルを用いて定式化せず、有色駆動源と呼ばれる、駆動源と状態量とが相関をもつカルマンフィルタ [3] を用いた状態量推定法 (以降、雑音抑圧法と称する) について説明する。

雑音抑圧法における L_p 次元の状態ベクトルを

$$\boldsymbol{x}_p(n) = \left[d(n), d(n-1), \dots, d(n-L_p+1) \right]^T \quad (2.10)$$

と定義したとき、音声信号のみで状態方程式を表わすと

$$\text{[状態方程式]} \quad \boldsymbol{x}_p(n+1) = \Phi_p \boldsymbol{x}_p(n) + \boldsymbol{g}_p \delta_p(n+1) \quad (2.11)$$

となる．ここで， $L_p \times L_p$ 行列の状態遷移行列 Φ_p ， L_p 次元駆動源ベクトル \mathbf{g}_p および，駆動源 $\delta_p(n+1)$ は以下のように定義される．

$$\left. \begin{aligned} \Phi_p &= \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \\ \mathbf{g}_p &= [1, 0, 0, \dots, 0]^T \\ \delta_p(n+1) &= d(n+1) \end{aligned} \right\} \quad (2.12)$$

式 (2.12) の状態遷移行列は，状態量の時間変化を 0 と 1 のみで構成されるシフト行列である．この行列を状態ベクトル $\mathbf{x}_p(n)$ に作用させると，時刻 n から $(n-L_p+2)$ までの状態量が $(n+1)$ 時刻における状態ベクトルに含まれるようになる．しかし，時刻 $(n+1)$ の状態量は状態遷移行列の構造より，推定することができない．そこで，駆動源 $\delta_p(n+1)$ に時刻 $(n+1)$ における状態量 $d(n+1)$ を置くことで状態量の変化を表している．前節の駆動源は AR モデルの予測誤差としていたため，白色信号あることに対し，雑音抑圧法では音声信号とするため有色信号となる．

次いで観測方程式は前節の式 (2.8) と同様に

$$\text{[観測方程式]} \quad y_p(n) = \mathbf{m}_p^T \mathbf{x}_p(n) + \varepsilon_p(n) \quad (2.13)$$

となる．ここで L_p 次元観測ベクトル \mathbf{m}_p^T ，観測雑音 $\varepsilon_p(n+1)$ は次式となる．

$$\left. \begin{aligned} \mathbf{m}_p &= [1, 0, 0, \dots, 0]^T \\ \varepsilon_p(n) &= v(n) \end{aligned} \right\} \quad (2.14)$$

雑音抑圧法は式 (2.11) の状態方程式と式 (2.13) の観測方程式から成る状態空間モデルに対して，カルマンフィルタ理論 [1] 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する．雑音抑圧法のアルゴリズムを表 2.2 に示す．

雑音抑圧法 [3] は，音声信号の時間変化を AR モデルを用いて定式化することなく表している．それゆえ雑音抑圧法のメリットとして，(i) AR 次数決定問題に起因する状態量の推定精度劣化が発生せず，(ii) 式 (2.12) の状態遷移行列が”0”と”1”のみで構成されることによる演算量の軽減が期待できる．しかし，雑音抑圧法を用いて推定された状態量について議論がされてこなかった．次節では雑音抑圧法の理論的な裏付けをし，最終的に最小分散推定法に帰着されることを示す．

表 2.2: 有色駆動源カルマンフィルタを用いた雑音抑圧法のアルゴリズム

[Initialization]
$\widehat{\mathbf{x}}_p(0 0) = \mathbf{0}, P_p(0 0) = I_{L_p}$
$R_{\delta_p}(n)[i, j] = \begin{cases} E[y^2(n)] - \sigma_v^2 & (i, j = 1) \\ 0 & (\text{other}) \end{cases}$
$r_{\epsilon_p}(n) = \sigma_v^2$
[Iteration]
1. $P_p(n+1 n) = \Phi_p P_p(n n) \Phi_p^T + R_{\delta_p}(n+1)$
2. $\mathbf{k}_p(n+1) = \{P_p(n+1 n) \mathbf{m}_p^T\} \cdot \{ \mathbf{m}_p P_p(n+1 n) \mathbf{m}_p^T + r_{\epsilon_p}(n+1) \}^{-1}$
3. $\widehat{\mathbf{x}}_p(n+1 n) = \Phi_p \widehat{\mathbf{x}}_p(n n)$
4. $\widehat{\mathbf{x}}_p(n+1 n+1) = \widehat{\mathbf{x}}_p(n+1 n) + \mathbf{k}_p(n+1) \cdot \{y_p(n+1) - \mathbf{m}_p^T \widehat{\mathbf{x}}_p(n+1 n)\}$
5. $P_p(n+1 n+1) = \{I - \mathbf{k}_p(n+1) \mathbf{m}_p^T\} P_p(n+1 n)$
6. $n = n + 1$ go back 1.

2.4 雑音抑圧法と最小分散推定法

本節では雑音抑圧法を用いて推定された状態量が最小分散推定法に帰着されることを示す。

2.4.1 雑音抑圧法の解析

前節で述べた表 2.2 における雑音抑圧アルゴリズムについて、式 (2.12) の状態遷移行列の構造に注意しながら解析する。

まず、表 2.2 手順 3. の事前推定ベクトル $\widehat{\mathbf{x}}_p(n+1|n)$ の第 1 要素は、時刻 $(n+1)$ において出力される音声信号の事前推定値 $\widehat{d}(n+1|n)$ である。

ここで、式 (2.12) に示した状態遷移行列 Φ_p の構造に着目すれば、 $\widehat{d}(n+1|n)$ は時刻 n における最適推定値 $\widehat{\mathbf{x}}_p(n|n)$ の要素は一切使われずに更新され、結果 0 となる。また、 $\widehat{\mathbf{x}}_p(n+1|n)$ の第 1 要素以外は、1 つ前の時刻 n で実行された結果が順次反映される。すなわち、 $\widehat{\mathbf{x}}_p(n+1|n)$ の第 2 要素は $\widehat{\mathbf{x}}_p(n|n)$ の第 1 要素となり、これは時刻 $(n-1)$ での最適推定値であることから既知量となる。第 3 要素以降も同様である。

従って、時刻毎で更新の対象となるのは、手順 4. における $\widehat{\mathbf{x}}_p(n+1|n+1)$ の第 1 要素のみである。これより、 $\widehat{\mathbf{x}}_p(n+1|n+1)$ の更新では、 $\widehat{\mathbf{x}}_p(n+1|n+1)$ の第 1 要素のみに着目す

ればよい。これは、手順4.のカルマンゲインベクトル $\mathbf{k}_p(n+1)$ の第1要素のみを考慮すればよいことを意味する。

以上のことを踏まえると、表2.2の雑音抑圧アルゴリズムは次のように書き換えることができる。

まず、事後推定ベクトル $\widehat{\mathbf{x}}_p(n+1|n+1)$ における第1要素の算出に必要な要素のみを書き下すと

$$\begin{aligned}\widehat{\mathbf{x}}_p(n+1|n+1)[1] &= \widehat{\mathbf{x}}_p(n+1|n)[1] + \mathbf{k}_p(n+1)[1] \{y(n+1) - \mathbf{m}_p^T \widehat{\mathbf{x}}_p(n+1|n)\} \\ &= 0 + \mathbf{k}_p(n+1)[1] \{y(n+1) - 0\} \\ &= \mathbf{k}_p(n+1)[1] y(n+1)\end{aligned}\quad (2.15)$$

となる。ここで $\mathbf{a}[i]$ はベクトル \mathbf{a} の i 行目の要素を指定する記号である。

次いで表2.2手順2.のカルマンゲインベクトル $\mathbf{k}_p(n+1)$ は、式(2.15)より $\mathbf{k}_p(n+1)$ の第1要素のみに着目すると

$$\mathbf{k}_p(n+1)[1] = P_p(n+1|n)[1, 1] \cdot \{P_p(n+1|n)[1, 1] + r_{\epsilon_p}(n)\}^{-1}\quad (2.16)$$

のように表される。ここで $A[i, j]$ は行列 A の i 行 j 列目の要素を指定する記号である。

最後に表2.2手順1.の事前推定誤差共分散行列 $P_p(n+1|n)$ は、式(2.16)より $P_p(n+1|n)$ の1行1列目の要素のみに着目し、かつ $R_{\delta_p}(n+1)$ と $Q_p(n+1)$ の行列の構造を考慮すれば

$$\begin{aligned}P_L(n+1|n)[1, 1] &= P_{imp}[1, 1] + R_{\delta_p}(n+1)[1, 1] + Q_p(n+1)[1, 1] + Q_p^T(n+1)[1, 1] \\ &= 0 + \sigma_d^2(n+1) + 0 + 0 \\ &= \sigma_d^2(n+1)\end{aligned}\quad (2.17)$$

と書き表すことができる。ここで $P_{imp} = \Phi_p P_p(n|n) \Phi_p^T$ 、 $\sigma_d^2(n+1)$ は音声信号の分散値である。

新たに駆動源分散値を $p(n+1)$ 、カルマンゲインを $k(n+1)$ 、音声信号推定値を $\widehat{d}(n+1)$ とおけば、表2.2の雑音抑圧アルゴリズムは式(2.15)～(2.17)より

$$\left. \begin{aligned}1. \quad p(n+1) &= \sigma_d^2(n+1) \\ 2. \quad k(n+1) &= \frac{p(n+1)}{p(n+1) + r_{\epsilon_p}(n)} \\ 3. \quad \widehat{x}_p(n+1) &= k(n+1)y(n+1)\end{aligned} \right\} \quad (2.18)$$

のように表される。

表 2.3: 最小分散状態推定アルゴリズム

[Initialization]
$\sigma_d^2(n+1) = E[d^2(n+1)], r_{\epsilon_p}(n+1) = E[v^2(n+1)]$
[Iteration]
1. $p(n+1) = \sigma_d^2(n+1)$
2. $k(n+1) = \frac{p(n+1)}{p(n+1) + r_{\epsilon_p}(n+1)}$
3. $\widehat{x}_p(n+1) = k(n+1)y(n+1)$
4. $n = n+1$ go back 1.

2.4.2 最小分散推定法 [2]

m 次元の観測ベクトルを \mathbf{y} としたとき

$$\mathbf{y} = W\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (2.19)$$

という形の観測ベクトルが得られたとする。ここで W は既知の $m \times n$ 定数行列、 $\boldsymbol{\beta}$ は n 次元未知確率ベクトル、 $\boldsymbol{\epsilon}$ は m 次元確率誤差ベクトルとする。また、未知確率ベクトルと確率誤差ベクトルの共分散値をそれぞれ $R = E[\boldsymbol{\beta}\boldsymbol{\beta}^T]$, $Q = E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T]$ とする。このとき $E[\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|^2]$ を最小にする $\boldsymbol{\beta}$ の推定値 $\hat{\boldsymbol{\beta}}$ は

$$\hat{\boldsymbol{\beta}} = RW^T(WRW^T + Q)^{-1}\mathbf{y} \quad (2.20)$$

であり、最小分散推定法 [2] と呼ばれている。

さて、前節において議論した式 (2.18) の雑音抑圧法は、式 (2.20) で示した最小分散推定法において観測ベクトルがスカラーである場合に一致していることがわかる。それゆえ雑音抑圧法は最小分散推定法に帰着されることが示された。

最後に式 (2.18) に示した雑音抑圧法を最小分散状態推定法と呼ぶこととし、表 2.3 にアルゴリズムを示す。

2.5 結言

本章では、状態量を AR モデルで定式化した場合における問題点とその問題の解決策として提案された雑音抑圧法について述べた。雑音抑圧法の状態空間モデル、特に状態遷移

行列に着目したとき雑音抑圧法は最小分散推定法に帰着することを示した。

次章では雑音抑圧法を無線通信におけるチャネルゲイン推定法および、カラオケ音源生成手法に応用し、それぞれの手法に応用しても最終的に最小分散推定法に帰着されていることを示す。

第3章 MIMO MC-CDMA 通信における チャンネルゲイン推定問題への適用

本章では、雑音抑圧法 [3] を MIMO MC-CDMA 通信方式に対するチャンネルゲイン推定法に応用し、チャンネルゲインの推定精度が移動体の移動速度に依存しないことおよび、チャンネルゲイン推定に要する演算量が軽減されることを述べる。

3.1 緒言

スマートフォンの普及により、無線通信における通信トラフィックは以前と比較して大幅に増大している。これを受け、2020 年の実用化に向けて通信速度が 10Gbps を超えるような 5G(第 5 世代移動通信システム)の開発が進んでいる。ここで、高速で大容量な無線通信を可能にする技術をいくつか列挙する。

まず CDMA(Code Division Multiple Access)[5] は、第 3 世代移動通信システムで使われていた通信方式である。ユーザ毎に直交する拡散符号を送信信号に付加することで、複数のユーザが同一時に同一帯域を共有して通信が可能となる。また逆拡散処理の過程で干渉波の成分が広帯域に拡散されることになり、干渉成分の除去が容易となる [5]。

次いで OFDM(Orthogonal Frequency Division Multiplexing) [6] は、隣接する狭帯域のサブキャリア同士を直交させて周波数利用効率を高めることができ、また周波数選択性フェージングに強い通信方式であることが知られている。また、OFDM 通信方式はガードインターバル (GI: Guard Interval) を付加することにより通信容量を犠牲にしつつも、マルチパスフェージングへの耐性を持たせることが可能なことから、無線 LAN や地上デジタル放送など無線通信の基盤技術として利用されている。さらに、OFDM 通信方式に CDMA 通信方式を組み合わせた MC-CDMA(Multi Carrier-CDMA) [8] 通信方式が提案されている。

一方で、通信容量の観点からは、送受信器に複数本のアンテナを用いた MIMO(Multi-input Multi-output) システム [7] が知られており、送信アンテナ数を増加させることでシャ

ノン・ハートレーの通信容量の限界を超えた大容量通信を可能とする [7].

本論文では、これらの通信方式や通信システムの利点を併せ持つ MIMO MC-CDMA [9] 通信方式に着目した.

さて、一般的に無線通信は空間を伝搬媒体としているため、送信信号は伝送路変動 (チャンネルゲイン変動) の影響を受けシンボル間干渉 (ISI: Inter Symbol Interference) が生じる. これに対して OFDM 変調をベースとする MIMO MC-CDMA 通信方式では、伝送路の最大遅延時間よりも長い GI を付加することにより、OFDM シンボル間の干渉を除去可能である. しかしながら 1 つのサブキャリア単位で考えると、個々のサブキャリア信号は受信側でチャンネルゲインに起因する振幅と位相の歪みを受けるため [10, 11], 通信品質の劣化が生じてしまう. それゆえ、上記のような歪を補償するために、適応的にかつ適切にチャンネルゲインを推定する必要がある [6].

チャンネルゲイン変動に対応し、リアルタイムにチャンネルゲイン推定を行うためには、できるだけ直近の情報を活用し、かつ演算量を可能な限り軽減する必要がある. このような観点から、受信側で既知な情報 (パイロット信号) [12] を定期的を送信信号へ付加することで、伝送効率を犠牲にしてリアルタイムにチャンネルゲインを推定する手法は数多く研究されている [8, 18, 19, 20, 21, 22]. その中でも MIMO-OFDM 通信方式に適用されている期待値最大化アルゴリズムを用いた手法 [18] は、高い精度でチャンネルゲイン推定を可能としている一方で、精度を得るために多くのデータを用いた繰返し演算が必要となる. この手法に対して、MIMO-OFDM 通信のための最小二乗法を用いた手法 [19, 20] は、処理がシンプルな反面、事前にチャンネルの統計的情報を知る必要がある.

これらの問題を解決する一手法として、演算効率と推定精度のバランスに優れた自己回帰 (AR: Auto Regressive) システムとカルマンフィルタ理論 [1] を用いたチャンネル推定法が挙げられる [21, 22, 8]. これらの手法は、推定対象となる各送受信アンテナ間ごとのチャンネルゲイン変動を AR モデルで定式化した上で、チャンネルゲインを推定するものである. また、文献 [21, 22, 8] の提案に先立ち、文献 [23] では移動体の速度とチャンネルゲインのモデル化次数 (AR 次数) の関係が報告されている. ただし文献 [23] ではフェージング環境下にある一般的な通信システムを対象としており、その適用を MIMO MC-CDMA システムに限定しているものではない.

さて、上記チャンネルゲイン推定法は 2 つの問題点

1. チャネルゲイン変動を AR モデルで定式化した上でチャネルゲイン推定を行う場合には、文献 [23] で示されているように、移動体の速度がその推定精度を左右する。しかしながら文献 [21, 22, 8] ではその欠点が解決されていない。
2. MIMO MC-CDMA 通信は、送受信アンテナ数の増加に伴いモデル化すべきチャネル数も増加するため、チャネルゲイン推定に要する演算量が増加する。

を有する。

そこで本章では、雑音抑圧法を用いた MIMO MC-CDMA チャネル推定法について検討を行い、上記の問題点を改善することを試みる。特に (1) チャネルゲイン変動の定式化方法、(2) 移動速度の変化に対するチャネルゲイン推定精度評価、(3) および送受信アンテナ数の増加に対する演算量、について議論を行う。

本手法は、任意の送受信アンテナ間に対する MIMO MC-CDMA チャネルゲイン変動を AR モデルを用いずに定式化した状態方程式、および上記チャネルゲインの影響を受け、かつ回線雑音が重畳した受信信号を表す観測方程式からなる状態空間モデルを構成することで、チャネルゲインを推定している。

ここで本論文で使用する記号について説明する。大文字のアルファベット A は周波数領域信号、大文字のボールド体 \mathbf{A} は周波数領域信号のベクトル表記、大文字のカリグラフィック体 \mathcal{A} は周波数領域信号の行列表記、 I_a は $a \times a$ の単位行列、 $O_{a \times b}$ は $a \times b$ の零行列、 $E[\cdot]$ は期待値演算、 a^* は a の共役複素数、 A^T は行列 A の転置、 A^H は行列 A のエルミート転置、 A^{-1} は行列 A の逆行列、 $\|\mathbf{a}\|$ はベクトル \mathbf{a} のノルム、 $A \otimes B$ は行列 A と B のクロネッカー積、および A_c と A_p のように表わされる下付き文字は、それぞれ従来手法または提案手法に用いられる記号であることを示す。

3.2 問題設定

本節では、MIMO MC-CDMA 通信方式の送信信号と受信信号の生成過程について説明する。

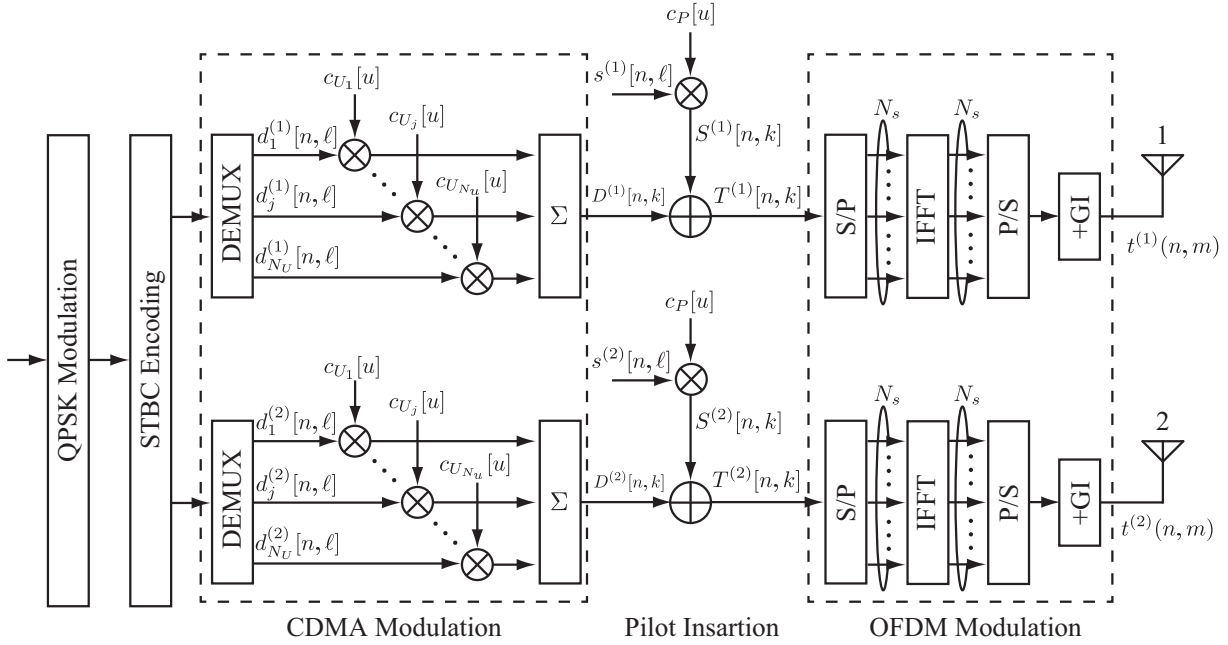


図 3.1: MIMO MC-CDMA 通信方式 送信機ブロック図

3.2.1 送信信号の生成

図 3.1 に MIMO MC-CDMA 通信方式 [9] の送信機を示す. なお簡単化のために, 送信アンテナ数 $P = 2$, 受信アンテナ数 $Q = 2$ の 2×2 MIMO システムを想定して説明する. まず各ユーザのデータシンボルはあらかじめ QPSK (Quadrature Phase Shift Keying) などの 1 次変調が施される. MIMO システムでは複数のアンテナを用いて通信容量を増加させる利用方法の他に, 受信信号の受信品質向上のために最大比合成ダイバーシチ [7] と呼ばれる利用方法もある. 本論文では, 受信信号とチャネルゲインの複素演算のみを用いて容易に最大比合成ダイバーシチ効果が得られる, 時空間ブロック符号化 (STBC: Space-Time Block Coding) [7] をプリコーディングとして施す.

次いで CDMA 変調および OFDM 変調を施すために, N_U 人のユーザが送信する複数シンボル (シンボル番号 $\ell = 1, 2, \dots, N_\ell$) のデータを同時に n ブロック目のデータとして扱う. n ブロック目のデータシンボルはユーザごとに割り当てられた符号長 N_c の拡散符号 $c_{U_j}[u] (u = 0, \dots, N_c - 1)$ を用いて拡散し, サブキャリア方向に配置する. p 番目のアンテナから送信される j ユーザ, n ブロックにおける ℓ シンボル目のデータシンボルを $d_j^{(p)}[n, \ell]$ とすると, N_U 人のデータシンボルが多重化された k サブキャリア目のデータシンボル $D^{(p)}[n, k]$

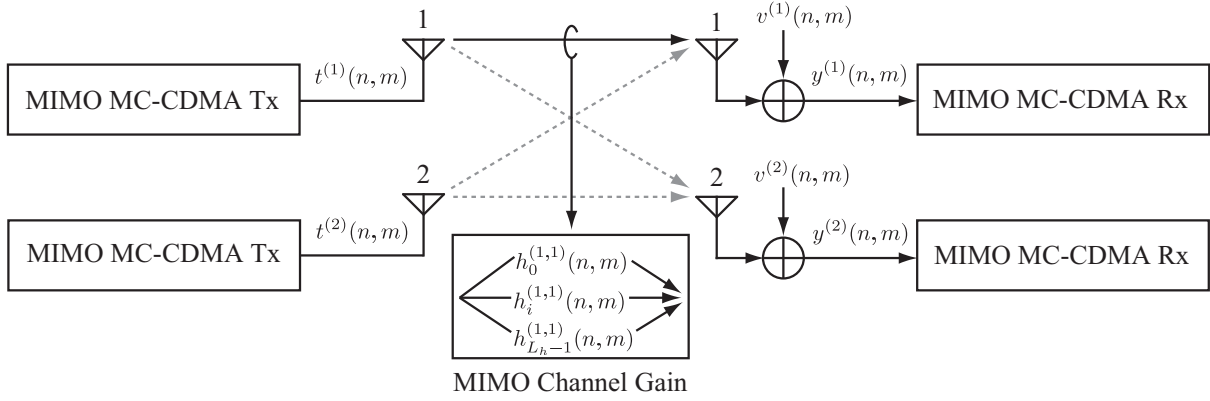


図 3.2: MIMO MC-CDMA 通信方式 チャネルモデル

は

$$D^{(p)}[n, k] = \sum_{j=1}^{N_U} d_j^{(p)}[n, \text{floor}(k/N_c)] c_{U_j}[\text{mod}(k/N_c)] \quad k = 0, 1, \dots, N_s - 1 \quad (3.1)$$

と表すことができる。ここで $N_s (= N_\ell \times N_c)$ はサブキャリア数、 $\text{floor}(\cdot)$ は床関数、 $\text{mod}(\cdot)$ は剰余算である。

次に、チャネルゲインを推定するためにデータシンボルに付加するパイロットシンボル $S^{(p)}[n, \ell]$ は、パイロット信号用の符号長 N_c の拡散符号 $c_P[u] (u = 0, \dots, N_c - 1)$ を用いて拡散することで得られ次式

$$S^{(p)}[n, k] = s^{(p)}[n, \text{floor}(k/N_c)] c_P[\text{mod}(k/N_c)] \quad k = 0, 1, \dots, N_s - 1 \quad (3.2)$$

のように表される。ここで $s^{(p)}[n, \ell]$ は p 番目送信アンテナに付加される、 n ブロック ℓ シンボル目のパイロットシンボルである。

p 本目のアンテナから送信される n ブロック k サブキャリア目の送信シンボル $T^{(p)}[n, k]$ は

$$T^{(p)}[n, k] = D^{(p)}[n, k] + S^{(p)}[n, k] \quad (3.3)$$

となる。各ブロックは N_s 本のサブキャリアから構成され、 N_s 点の逆高速フーリエ変換 (IFFT: Inverse Fast Fourier Transform) 処理により OFDM 変調されたのち、GI を付加し p 本目のアンテナより変調信号 $t^{(p)}(n, m)$ が送信される。なお添え字 m は n ブロック内におけるサンプル番号である。

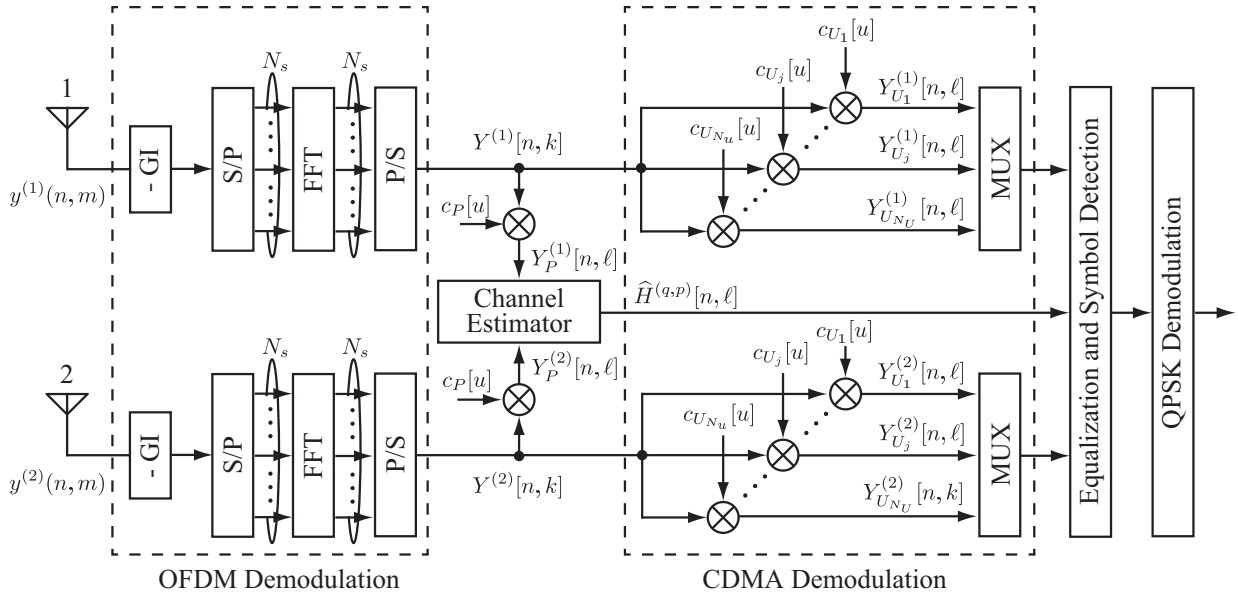


図 3.3: MIMO MC-CDMA システム 受信機ブロック図

3.2.2 受信信号の生成

図 3.2 は MIMO MC-CDMA 通信方式におけるチャネルのモデルを表したものである。送信信号は n 時刻の $(p - q)$ アンテナ間における、 i 遅延波のチャネルゲイン $\{h_i^{(q,p)}(n, m)\}$ の影響を受け、さらに加法的白色ガウス雑音 (AWGN: Additive White Gaussian Noise) $v^{(q)}(n, m)$ が付加されて受信される。このとき図 3.3 をもとに、 q 本目のアンテナに受信される受信信号 $y^{(q)}(n, m)$ は [9]

$$y^{(q)}(n, m) = \sum_{p=1}^P \sum_{i=0}^{L_h-1} h_i^{(q,p)}(n, m) t^{(p)}(n, m - i) + v^{(q)}(n, m) \quad (3.4)$$

$q = 1, 2, \dots, Q$

となる。ここで AWGN の $v^{(q)}(n, m)$ は $E[v^{(q)}(n, m)] = 0$, $E[\{v^{(q)}(n, m)\}^2] = \sigma_v^2$ を満たすものとする。また各アンテナ間のチャネルゲイン $\{h_i^{(q,p)}(n, m)\}$ は、広義定常無相関散乱チャネル (WSSUS: Wide-Sense Stationary Uncorrelated Scattering) [15, 23] に従うレイリーフェージングモデルとする。

図 3.3 は MIMO MC-CDMA システムにおける受信機を表したものである。GI を除去した受信信号 $y^{(q)}(n, m)$ は、 N 点の高速フーリエ変換 (FFT: Fast Fourier Transform) 処理され、受信シンボル $Y^{(q)}[n, k]$ へ変換される。ここで $Y^{(q)}[n, k]$ には、パイロットシンボルを含む受信シンボル $Y_P^{(q)}[n, \ell]$ とユーザのデータシンボルを含む受信シンボル $Y_{U_i}^{(q)}[n, \ell]$ が多重されてい

る。それゆえそれぞれのシンボルに分離するために、データシンボル用の拡散符号 $c_{U_j}[u]$ 、パイロットシンボル用の拡散符号 $c_P[u]$ を用いて逆拡散処理をする。逆拡散処理後のパイロットシンボルを含む受信シンボル $Y_P^{(q)}[n, \ell]$ は

$$\begin{aligned} Y_P^{(q)}[n, \ell] &= \sum_{u=0}^{N_c-1} Y^{(q)}[n, \ell \cdot N_c + u] c_P[u] \\ &= \sum_{p=1}^P H_{ave}^{(q,p)}[n, \ell] s^{(p)}[n, \ell] + V_P^{(q)}[n, \ell] \end{aligned} \quad (3.5)$$

$q = 1, 2, \dots, Q$

ただし、 $H_{ave}^{(q,p)}[n, \ell]$ は $(q-p)$ アンテナ間の n ブロックにおける ℓ シンボルに係る周波数領域のチャネルゲイン、 $V_P^{(q)}[n, \ell]$ は q 本目のアンテナに受信された n ブロックにおける ℓ シンボルに係る周波数領域の雑音でありそれぞれ

$$\left. \begin{aligned} H_{ave}^{(q,p)}[n, \ell] &= \frac{1}{N_c} \sum_{u=0}^{N_c-1} H^{(q,p)}[n, \ell \cdot N_c + u] \\ V_P^{(q)}[n, \ell] &= \sum_{u=0}^{N_c-1} V^{(q)}[n, \ell \cdot N_c + u] c_P[u] \end{aligned} \right\} \quad (3.6)$$

のようにして得られる [9].

チャネルゲイン推定には、式 (3.5) で与えられたパイロットシンボルを含む受信シンボルとパイロットシンボルを用い、推定されたチャネルゲインを用いて等化処理を経ることで送信信号が復元される。

以後の議論のために Q 次元受信信号ベクトル $\mathbf{Y}_P[n, \ell]$ を

$$\mathbf{Y}_P[n, \ell] = [Y_P^{(1)}[n, \ell], Y_P^{(2)}[n, \ell], \dots, Y_P^{(Q)}[n, \ell]]^T \quad (3.7)$$

と定義したとき、式 (3.5) は

$$\mathbf{Y}_P[n, \ell] = \mathbf{S}[n, \ell] \mathbf{H}_{ave}[n, \ell] + \mathbf{V}_P[n, \ell] \quad (3.8)$$

と表現できる。ここで $Q \times QP$ パイロットシンボル行列 $\mathbf{S}[n, \ell]$ 、 QP 次元チャネルゲインベクトル $\mathbf{H}_{ave}[n, \ell]$ 、 Q 次元 AWGN ベクトル $\mathbf{V}_P[n, \ell]$ は、それぞれ次式となる。

$$\left. \begin{aligned} \mathbf{S}[n, \ell] &= [s^{(1)}[n, \ell], s^{(2)}[n, \ell], \dots, s^{(P)}[n, \ell]] \otimes I_Q \\ \mathbf{H}_{ave}[n, \ell] &= [H_{ave}^{(1,1)}[n, \ell], \dots, H_{ave}^{(Q,1)}[n, \ell], \dots, H_{ave}^{(1,P)}[n, \ell], \dots, H_{ave}^{(Q,P)}[n, \ell]]^T \\ \mathbf{V}_P[n, \ell] &= [V_P^{(1)}[n, \ell], V_P^{(2)}[n, \ell], \dots, V_P^{(Q)}[n, \ell]]^T \end{aligned} \right\} \quad (3.9)$$

本章の目的は、 Q 本のアンテナによって受信されたパイロットシンボルを含む受信シンボル $\mathbf{Y}_p[n, \ell]$ と送信シンボルに挿入したパイロットシンボルを用い、高速移動環境下での全送受信アンテナ間のチャネルゲイン $\{H_{ave}^{(q,p)}[n, \ell]\}$ を少ない演算量でかつ正確に推定することである。

3.3 ARモデルとカルマンフィルタを用いたチャネルゲイン推定法

本章では、ARモデルとカルマンフィルタを用いたチャネルゲイン推定法について、MC-CDMA通信方式におけるチャネルゲイン推定法 [8](以降、従来手法とする) を MIMO システムに拡張して議論する。

3.3.1 Step 1: AR 係数の算出

アンテナ $(p-q)$ 間の n ブロックにおける ℓ シンボルのチャネルゲイン $\{H_{ave}^{(q,p)}[n, \ell]\}$ を AR モデルを用いて定式化すると

$$H_{ave}^{(q,p)}[n, \ell] = \sum_{i=1}^{L_c} \alpha_i H_{ave}^{(q,p)}[n-i, \ell] + W^{(q,p)}[n, \ell] \quad (3.10)$$

のように与えられる。ここで $\{\alpha_i\}$ は i 次 AR 係数、 L_c は AR 次数、および $W^{(q,p)}[n, \ell]$ はアンテナ $(p-q)$ 間の n ブロックにおける ℓ シンボルのチャネルゲインモデル化誤差 (駆動源) でかつ白色性を有している。

本節では AR 係数を次式で与えられるチャネルゲインの自己相関関数

$$\begin{aligned} R_{HH}(\tau) &= E \left[H_{ave}^{(q,p)}[n, \ell] \left\{ H_{ave}^{(q,p)}[n-\tau, \ell] \right\}^* \right] \\ &= J_0(2\pi f_D T_F \tau) \end{aligned} \quad (3.11)$$

を用いて、Yule-Walker 方程式 [8] を解くことにより AR 係数を算出する。ここで $J_0(\cdot)$ は第 1 種 0 次ベッセル関数、 $f_D (= v/\lambda)$ は移動速度 v と搬送波の波長 λ より算出される最大ドップラー周波数、及び T_F は OFDM シンボル周期である。

3.3.2 Step 2:カルマンフィルタを用いたチャネルゲイン推定

QPL_c 次元状態ベクトル $\mathbf{X}_c[n, \ell] = [\mathbf{H}_{ave}^T[n, \ell], \mathbf{H}_{ave}^T[n-1, \ell], \dots, \mathbf{H}_{ave}^T[n-L_c+1, \ell]]^T$ は、式(3.9)のチャネルゲインベクトル $\mathbf{H}_{ave}[n, \ell]$ を n ブロックから L_c ブロック過去のチャネルゲインベクトルを並べたものとして定義する。式(3.10)を用いた全送受信アンテナ間のチャネルゲイン変動を表わす状態方程式は

$$\mathbf{X}_c[n+1, \ell] = \mathcal{F}_c \mathbf{X}_c[n, \ell] + \mathcal{G}_c \Delta_c[n, \ell] \quad (3.12)$$

のように与えられる。ここで Step 1 で求めた AR 係数が対角に並ぶ $QP \times QP$ 行列 $\mathcal{A}_i = \alpha_i I_{QP}$ を用いて、 $QPL_c \times QPL_c$ 状態遷移行列 \mathcal{F}_c は以下のように定義する。

$$\mathcal{F}_c = \left[\begin{array}{c|c} \mathcal{A}_1, \dots, \mathcal{A}_i, \dots, \mathcal{A}_{L_c-1} & \mathcal{A}_{L_c} \\ \hline I_{\eta_c} & \mathbf{O}_{\eta_c \times QP} \end{array} \right] \quad (3.13)$$

ただし $\eta_c = QP(L_c - 1)$ とする。また $QPL_c \times QP$ 行列 \mathcal{G}_c 、 QP 次元駆動源ベクトル $\Delta_c[n, \ell]$ は以下のように定義される。

$$\left. \begin{aligned} \mathcal{G}_c &= [I_{QP}, \mathbf{O}_{QP \times \eta_c}]^T \\ \Delta_c[n, \ell] &= [W^{(1,1)}[n, \ell], \dots, W^{(Q,1)}[n, \ell], \dots, W^{(1,P)}[n, \ell], \dots, W^{(Q,P)}[n, \ell]]^T \end{aligned} \right\} \quad (3.14)$$

次いで、送信信号がチャネルゲインの影響を受けさらに雑音加わり Q 本のアンテナに受信される過程を表した観測方程式は、式(3.8)を用いることで

$$\mathbf{Y}_c[n, \ell] = \mathcal{M}_c[n, \ell] \mathbf{X}_c[n, \ell] + \mathbf{E}_c[n, \ell] \quad (3.15)$$

となる。ここで $Q \times QPL_c$ 観測行列 $\mathcal{M}_c[n, \ell]$ 、 Q 次元観測雑音ベクトル $\mathbf{E}_c[n, \ell]$ は式(3.9)の $\mathbf{S}[n, \ell]$ と $\mathbf{V}_P[n, \ell]$ を用いて以下のように定義している。

$$\left. \begin{aligned} \mathcal{M}_c[n, \ell] &= [\mathbf{S}[n, \ell], \mathbf{O}_{Q \times \eta_c}] \\ \mathbf{E}_c[n, \ell] &= \mathbf{V}_P[n, \ell] \end{aligned} \right\} \quad (3.16)$$

駆動源ベクトル $\Delta_c[n, \ell]$ は白色信号でかつ、状態ベクトル $\mathbf{X}_c[n, \ell]$ は駆動源ベクトルおよび観測雑音ベクトル $\mathbf{E}_c[n, \ell]$ と無相関という条件のもとで、式(3.12)と式(3.15)をカルマンフィルタ理論 [1] に適用することにより、逐次的にチャネルゲインベクトル $\mathbf{H}_{ave}[n, \ell]$ を推定している。従来手法のアルゴリズムは表 3.1 で与えられる。

ここで従来手法の問題点について整理する。従来手法は、

表 3.1: 従来手法のカルマンフィルタアルゴリズム

[Initialization]

$$\widehat{\mathbf{X}}_c[0|0, \ell] = \mathbf{0}, \mathcal{P}_c[0|0, \ell] = I_{QPL_c}, f_D: \text{Known}$$

$$C_{E_c}[n, \ell] = E \left[\mathbf{E}_c[n, \ell] \mathbf{E}_c^H[n, \ell] \right] = \sigma_v^2 I_Q$$

$$R_{HH}(\tau) = J_0(2\pi f_D T_F \tau)$$

$$C_{\Delta_c}[j, k][n, \ell] = \begin{cases} E \left[\left\{ H_{ave}^{(q,p)}[n, \ell] - \sum_{i=1}^{L_c} \alpha_i H_{ave}^{(q,p)}[n-i, \ell] \right\}^2 \right] & (j = k) \\ 0 & (\text{other}) \end{cases}$$

[Iteration]

• Step1: Estimation of the AR coefficient using Eq. (3.10) and (3.11).

• Step2: Execution of the Kalman Algorithm.

$$1. \mathcal{P}_c[n|n-1, \ell] = \mathcal{F}_c \mathcal{P}_c[n-1|n-1, \ell] \mathcal{F}_c^H + \mathcal{G}_c C_{\Delta_c}[n, \ell] \mathcal{G}_c^H$$

$$2. \mathcal{K}_c[n, \ell] = \left\{ \mathcal{P}_c[n|n-1, \ell] \mathcal{M}_c^H[n, \ell] \right\} \cdot \left\{ \mathcal{M}_c[n, \ell] \mathcal{P}_c[n|n-1, \ell] \mathcal{M}_c^H[n, \ell] + C_{E_c}[n, \ell] \right\}^{-1}$$

$$3. \widehat{\mathbf{X}}_c[n|n-1, \ell] = \mathcal{F}_c \widehat{\mathbf{X}}_c[n-1|n-1, \ell]$$

$$4. \widehat{\mathbf{X}}_c[n|n, \ell] = \widehat{\mathbf{X}}_c[n|n-1, \ell] + \mathcal{K}_c[n, \ell] \cdot \left\{ \mathbf{Y}_p[n, \ell] - \mathcal{M}_c[n, \ell] \widehat{\mathbf{X}}_c[n|n-1, \ell] \right\}$$

$$5. \widehat{\mathbf{H}}_c[n, \ell] = \mathcal{G}_c^H \widehat{\mathbf{X}}_c[n|n, \ell]$$

$$6. \mathcal{P}_c[n|n, \ell] = \left\{ I_{QPL_c} - \mathcal{K}_c[n, \ell] \mathcal{M}_c[n, \ell] \right\} \mathcal{P}_c[n|n-1, \ell]$$

7. $n = n + 1$ go back 1.

1. 式 (3.10) で用いる AR 次数と式 (3.11) の AR 係数の算出に必要な最大ドップラー周波数が移動体の速度に依存するため、移動速度を正確に得ることができなければチャンネルゲイン推定精度の劣化が生じる。

2. MIMO 特有のアンテナ数に依存して演算量が増加する。

などの問題点がある。次節では、具体例によりこれら従来手法の問題点について検証する。

3.3.3 計算機シミュレーション

本節では表 3.2 の諸元をもとに、前節にて説明した従来手法のチャンネルゲイン推定精度を評価した。シミュレーション緒元を表 3.2 に示す。

表 3.2: シミュレーション諸元 [9]

Subcarrier modulation scheme	QPSK
Carrier Frequency	$f_c = 4.0[\text{GHz}]$
MIMO System	
Number of Tx antennas	$P = 2$
Number of Rx antennas	$Q = 2$
STBC Encoding	Alamouti Code
CDMA	
Number of Users	2
Spreading code	Walsh code
Spreading gain	16
OFDM	
Effective Symbol Length	$T_e = 82.4[\mu\text{s}]$
Guard Interval Length	$T_g = 20[\mu\text{s}]$
OFDM Symbol Duration	$T_F = 102.4[\mu\text{s}]$
Sampling Duration	$T_s = 50[\text{ns}]$
Number of Subcarrier	$K = 2048$
Channel Environment	
Fading Model	Rayleigh Fading[13]
Delay Profile	COST207[13]
Number of Path	$L_h = 6$
Other parameter	
AR Oder	$L_c = 2$
Number of Trials	100
Maximum Doppler Frequency	$f_D = 9.3, 138.9, 694.4[\text{Hz}]$
SNR	$0, 2, \dots, 20[\text{dB}]$

伝送路環境は、市街地を想定した指数関数近似モデルの COST207[13] を用い、遅延波数は表 3.2 のガードインターバル長 T_g に収まるよう最大で $L_h = 6$ とした。また本シミュレーションは、最大遅延時間をガードインターバル長 T_g とすれば、コヒーレンス帯域幅 $B_c = 1/T_g$ は OFDM シンボルの帯域幅 $B_{OFDM} = 1/T_F$ よりも大きいことより、フラットフェージング環境を想定している。

従来手法のチャネルゲイン推定精度の評価を、次式に示す正規化誤差ノルム (NMSE: Normalized Mean Square Error)[14]

$$\text{NMSE} = \frac{E \left[\|\mathbf{H}_{ave}[n, \ell] - \widehat{\mathbf{H}}_{ave}[n, \ell]\|^2 \right]}{E \left[\|\mathbf{H}_{ave}[n, \ell]\|^2 \right]} \quad (3.17)$$

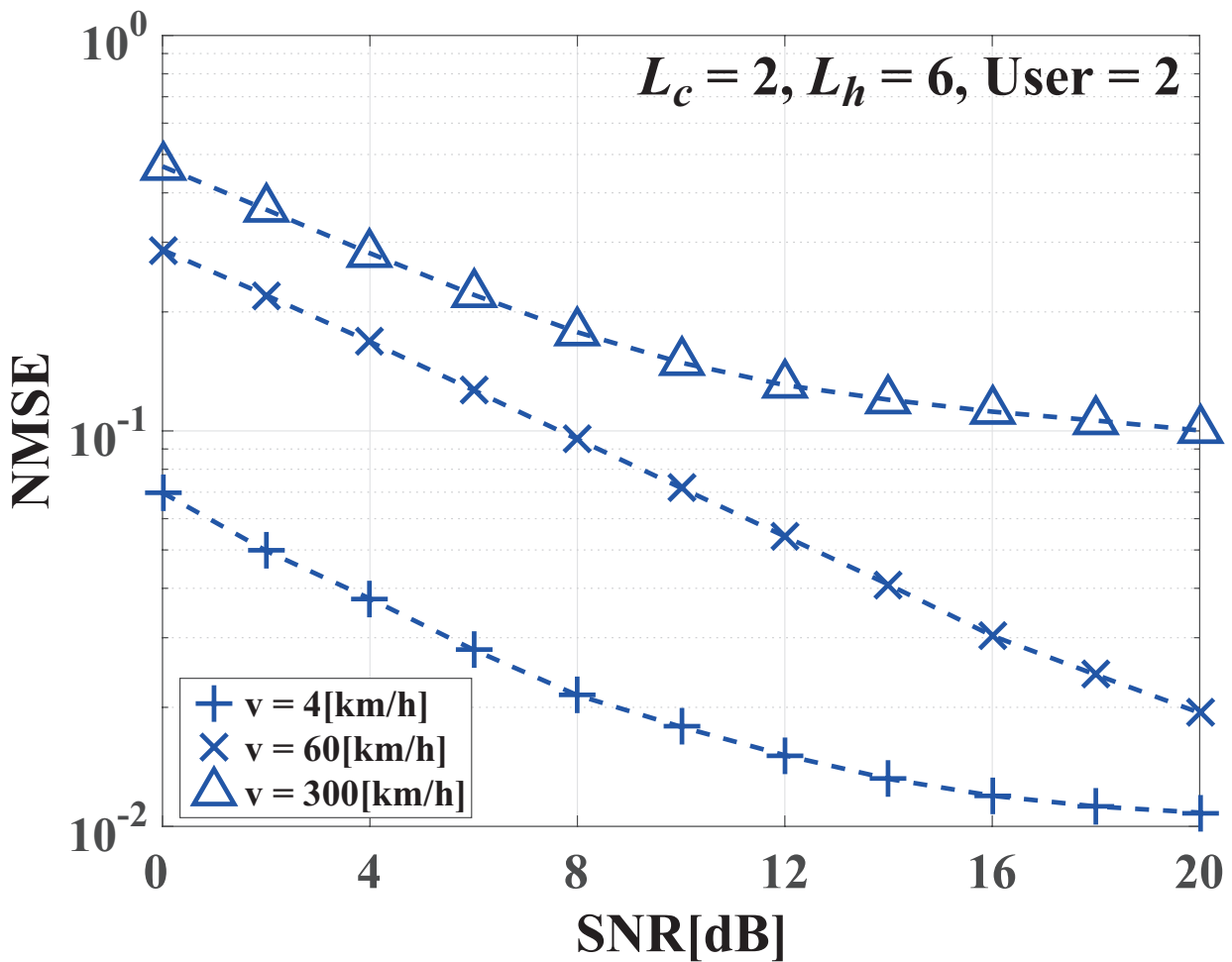


図 3.4: 従来手法の NMSE 特性

を用い、信号対雑音電力比 (SNR: Signal to Noise Ratio) を

$$\text{SNR} = 10 \log_{10} \frac{E \left[\|S[n, \ell] \mathbf{H}_{ave}[n, \ell]\|^2 \right]}{E \left[\|V[n, \ell]\|^2 \right]} [\text{dB}] \quad (3.18)$$

となるように与えた。

図 3.4 は、表 3.2 をもとに送受信アンテナ数をそれぞれ $P = 2, Q = 2$ 、遅延波数 $L_h = 6$ の MIMO MC-CDMA 通信システムにおける、移動体の速度を変化させた場合のチャネルゲイン推定精度を NMSE を用いて評価したものである。移動速度 v は $v = 4, 60, 300$ [km/h] のように変化させることにより、人の歩行、道路を走行する車、高速鉄道等の様々な移動環境を想定した。ただし、AR 次数については文献 [15, 14, 21, 17] より最適値であるとされる $L_c = 2$ を用いた。

図 3.4 より従来手法は、移動速度 v の変化に対して AR 次数を常に $L_c = 2$ としているこ

とより、正確な AR 係数が算出できずチャンネルゲインの推定精度が劣化していることが確認できる。

また図 3.4 の結果は、移動速度 v を既知として与えているが、実環境下では未知な情報である。それゆえ移動速度 v を何らかの手法を用いて推定しない限り、AR システムを用いる従来手法の実用化は困難といえる。

さらに式 (3.12)~(3.16) にて与えられる状態空間モデルを構成するベクトルや行列のサイズは、送受信アンテナ数 P, Q をパラメータに持つため、MIMO MC-CDMA チャンネルゲインの推定に必要な演算量は送受信アンテナ数に依存して増加すると考えられる。一例を挙げれば表 3.1 に示す従来手法の手順 1. $\mathcal{P}_c[n|n-1, \ell]$ の更新には、 $2Q^2P^2L_c^2$ 回の演算量 (乗算回数) が必要となり、演算量の軽減が望まれる。

これらの問題を解決するために移動体の移動速度に依存することなく、かつ低演算量で MIMO MC-CDMA チャンネルゲイン推定可能な手法について次章で述べる。

3.4 雑音抑圧法を用いたチャンネルゲイン推定手法

本章では MIMO MC-CDMA 通信方式におけるチャンネルゲイン推定に雑音抑圧法を適用した手法について議論する。

3.4.1 提案手法の状態空間モデルとチャンネルゲイン推定アルゴリズム

提案手法の QPL_p 次元状態ベクトル $\mathbf{X}_p[n, \ell]$ は式 (3.9) のチャンネルゲインベクトル $\mathbf{H}_{ave}[n, \ell]$ を n ブロックから L_p ブロック過去のチャンネルゲインベクトルを並べたものとし、以下のように定義する。

$$\mathbf{X}_p[n, \ell] = \left[\mathbf{H}_{ave}^T[n, \ell], \mathbf{H}_{ave}^T[n-1, \ell], \dots, \mathbf{H}_{ave}^T[n-L_p+1, \ell] \right]^T \quad (3.19)$$

このとき、全送受信アンテナ間のチャンネルゲイン変動を、雑音抑圧法 [3] を基にして次のように考える。

まず $(n+1)$ ブロックのチャンネルゲインを、 n ブロックにおけるチャンネルゲインに状態遷移行列を作用させ、 $(n+1)$ ブロックにおいて新たなチャンネルゲインを加えることでチャンネルゲイン変動を表現する。

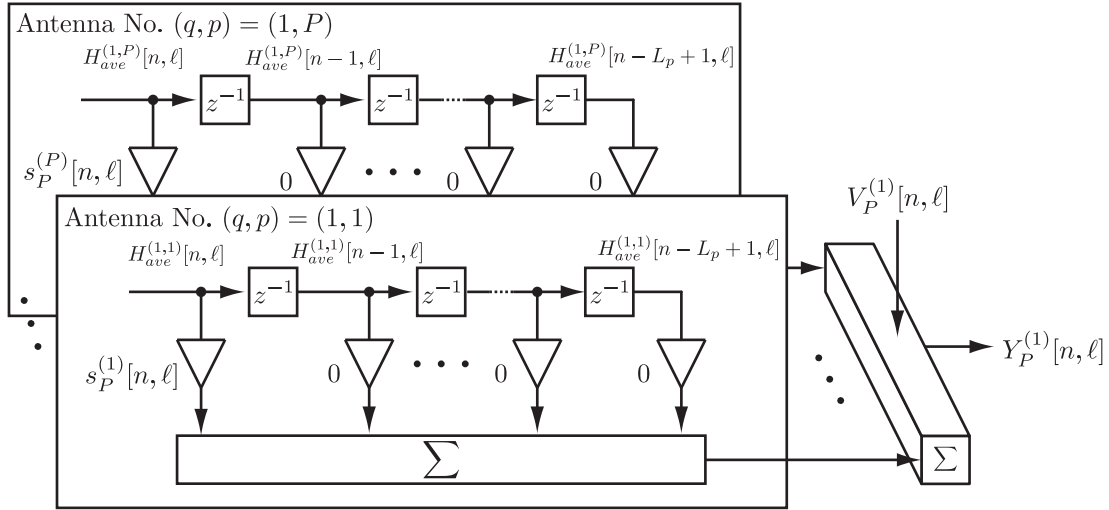


図 3.5: n ブロック目の状態遷移モデル

図 3.5 は、 n ブロックにおける送信信号が 1 本目から P 本目のすべての送信アンテナから送信され、1 本目の受信アンテナに受信されるまでに受けるチャンネルゲインと雑音の影響を FIR 型のフィルタとして表したものである。また図 3.6 は、図 3.5 から 1 ブロック遷移した $(n+1)$ ブロックにおける上記と同様の様子を表した図である。

図 3.5 の n ブロックと図 3.6 の $(n+1)$ ブロックのチャンネルゲイン変動 $\{H_{ave}^{(q,p)}[n, \ell]\}$ に着目すると、図 3.6 の $(n+1)$ ブロックにおける状態ベクトル $\mathbf{X}_p[n+1, \ell]$ は、図 3.5 の n ブロックから $(n-L_p+1)$ ブロックまでのチャンネルゲインのうち、最も過去にあたるチャンネルゲイン $\{H_{ave}^{(q,p)}[n-L_p+1, \ell]\}$ を除いて、新たに $(n+1)$ ブロックのチャンネルゲイン $\{H_{ave}^{(q,p)}[n+1, \ell]\}$ を加えたものとして考えることができる。

上記の関係を状態方程式を用いて表すと

$$\mathbf{X}_p[n+1, \ell] = \mathcal{F}_p \mathbf{X}_p[n, \ell] + \mathcal{G}_p \Delta_p[n+1, \ell] \quad (3.20)$$

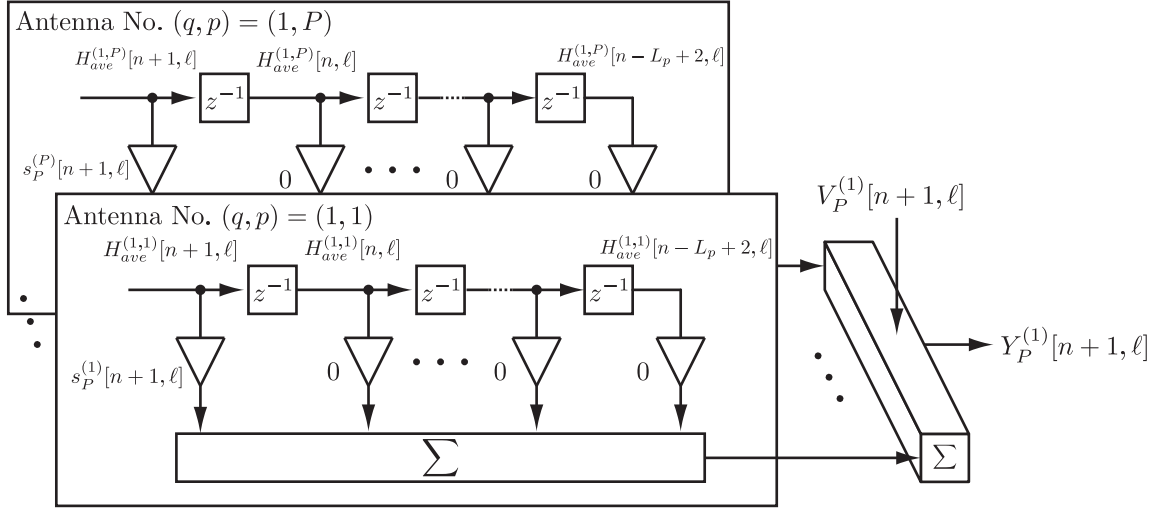


図 3.6: $(n+1)$ ブロック目の状態遷移モデル

と与えられる．ここで $QPL_p \times QPL_p$ 状態遷移行列 \mathcal{F}_p を

$$\mathcal{F}_p = \begin{matrix} & \overbrace{\hspace{10em}}^{QPL_p} \\ \left. \begin{matrix} QP \\ \vdots \\ \eta_p \end{matrix} \right\} & \left[\begin{array}{cccccc} 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 1 & & \mathbf{O} & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots & & \vdots \\ \mathbf{O} & & 1 & 0 & \cdots & 0 \end{array} \right] & (3.21) \\ & \underbrace{\hspace{4em}}_{\eta_p} & \underbrace{\hspace{4em}}_{QP} \end{matrix}$$

のように定義した．ただし $\eta_p = QP(L_p - 1)$ とする．また， $QPL_p \times QP$ 行列 \mathcal{G}_p ， QP 次元駆動源ベクトル $\Delta_p[n+1, \ell]$ をそれぞれ以下のように定義する．

$$\left. \begin{aligned} \mathcal{G}_p &= \left[I_{QP}, \mathbf{O}_{QP \times \eta_p} \right]^T \\ \Delta_p[n+1, \ell] &= \mathbf{H}_{ave}[n+1, \ell] \end{aligned} \right\} \quad (3.22)$$

次に式 (3.8) より観測方程式は

$$\mathbf{Y}_p[n, \ell] = \mathcal{M}_p[n, \ell] \mathbf{X}_p[n, \ell] + \mathbf{E}_p[n, \ell] \quad (3.23)$$

と与えられる．ここで $Q \times QPL_p$ 観測行列 $\mathcal{M}_p[n, \ell]$ ， Q 次元観測雑音ベクトル $\mathbf{E}_p[n, \ell]$ をそれぞれ式 (3.9) の $\mathcal{S}[n, \ell]$ と $\mathbf{V}_p[n, \ell]$ を用いて以下のように定義している．

$$\left. \begin{aligned} \mathcal{M}_p[n, \ell] &= \left[\mathcal{S}[n, \ell], \mathbf{O}_{Q \times \eta_p} \right] \\ \mathbf{E}_p[n, \ell] &= \mathbf{V}_p[n, \ell] \end{aligned} \right\} \quad (3.24)$$

表 3.3: 提案手法のチャネルゲイン推定アルゴリズム

[Initialization]

$$\widehat{\mathbf{X}}_p[0|0, \ell] = \mathbf{0}, \mathcal{P}_p[0|0, \ell] = I_{QPL_p}, f_D: \text{Unknown}$$

$$\mathbf{C}_{E_p}[n, \ell] = E \left[\mathbf{E}_p[n, \ell] \mathbf{E}_p^H[n, \ell] \right] = \sigma_v^2 I_Q$$

$$\mathbf{R}_{HH}(0) = E \left[\mathbf{H}_{ave}^{(q,p)}[n, \ell] \left\{ \mathbf{H}_{ave}^{(q,p)}[n, \ell] \right\}^* \right] = J_0(0)$$

$$\mathbf{C}_{\Delta_p}[j, k][n, \ell] = \begin{cases} J_0(0) & (j = k) \\ 0 & (\text{other}) \end{cases}$$

[Iteration]

$$1. \mathcal{P}_p[n|n-1, \ell] = \mathcal{F}_p \mathcal{P}_p[n-1|n-1, \ell] \mathcal{F}_p^H + \mathcal{G}_p \mathbf{C}_{\Delta_p}[n, \ell] \mathcal{G}_p^H$$

$$2. \mathcal{K}_p[n, \ell] = \left\{ \mathcal{P}_p[n|n-1, \ell] \mathbf{M}_p^H[n, \ell] \right\} \cdot \left\{ \mathbf{M}_p[n, \ell] \mathcal{P}_p[n|n-1, \ell] \mathbf{M}_p^H[n, \ell] + \mathbf{C}_{E_p}[n, \ell] \right\}^{-1}$$

$$3. \widehat{\mathbf{X}}_p[n|n-1, \ell] = \mathcal{F}_p \widehat{\mathbf{X}}_p[n-1|n-1, \ell]$$

$$4. \widehat{\mathbf{X}}_p[n|n, \ell] = \widehat{\mathbf{X}}_p[n|n-1, \ell] + \mathcal{K}_p[n, \ell] \left\{ \mathbf{Y}_p[n, \ell] - \mathbf{M}_p[n, \ell] \widehat{\mathbf{X}}_p[n|n-1, \ell] \right\}$$

$$5. \widehat{\mathbf{H}}_p[n, \ell] = \mathcal{G}_p^H \widehat{\mathbf{X}}_p[n|n, \ell]$$

$$6. \mathcal{P}_p[n|n, \ell] = \left\{ I_{QPL_p} - \mathcal{K}_p[n, \ell] \mathbf{M}_p[n, \ell] \right\} \mathcal{P}_p[n|n-1, \ell]$$

$$7. n = n + 1 \text{ go back 1.}$$

最後に、式 (3.20) と (3.23) に示した状態空間モデルを用いて逐次的にチャネルゲインベクトル $\mathbf{H}_{ave}[n, \ell]$ を推定する。本手法のアルゴリズムは表 3.3 で与えられる。

さて、2 章において雑音抑圧法は最小分散推定法に帰着されることを述べた。次節では MIMO MC-CDMA 通信方式のチャネルゲイン推定法として雑音抑圧法を適用し得られた表 3.3 のアルゴリズムが、最小分散推定法に帰着されるかについて議論する。

3.4.2 雑音抑圧法を適用したチャネルゲイン推定法の解析

本節では、2 章 2.4.1 節と同様に状態遷移行列、観測行列の構造及び、 n ブロックにおけるチャネルゲインの推定に必要な要素に着目し表 3.3 のアルゴリズムを整理する。

まず、表 3.3 手順 3. のチャネルゲイン事前推定ベクトル $\widehat{\mathbf{X}}_p[n|n-1, \ell]$ の 1 行目から QP 行目は、 n ブロックにおいて出力されるチャネルゲインの事前推定値 $\widehat{\mathbf{H}}_p[n, \ell]$ である。

ここで、式 (3.21) に示した状態遷移行列 \mathcal{F}_p の構造に着目すれば、 $\widehat{\mathbf{H}}_p[n, \ell]$ は $(n-1)$ ブ

ロックにおける最適推定値 $\widehat{\mathbf{X}}_p[n-1|n-1, \ell]$ の要素は一切使われずに更新され、結果 $\mathbf{0}$ となる。また、 $\widehat{\mathbf{X}}_p[n|n-1, \ell]$ の 1 行目から QP 行目以外は、1 つ前の $(n-1)$ ブロックで実行された結果が順次反映される。すなわち、 $\widehat{\mathbf{X}}_p[n|n-1, \ell]$ の $(QP+1)$ 行目から $2QP$ 行目は $\widehat{\mathbf{X}}_p[n-1|n-1, \ell]$ の 1 行目から QP 行目の要素となり、これは $(n-2)$ ブロックでの最適推定値であることから既知量となる。 $(2QP+1)$ 行目から $3QP$ 行目以降の要素も同様である。

従って、ブロック毎で更新の対象となるのは、手順 4. における $\widehat{\mathbf{X}}_p[n|n, \ell]$ の 1 行目から QP 行目のみである。これより、 $\widehat{\mathbf{X}}_p[n|n, \ell]$ の更新では、 $\widehat{\mathbf{X}}_p[n|n, \ell]$ の 1 行目から QP 行目のみに着目すればよい。これは、手順 4. のカルマンゲイン行列 $\mathcal{K}_p[n, \ell]$ の 1 行目から QP 行目のみを考慮すればよいことを意味する。

以上のことを踏まえると、表 3.3 のチャンネルゲイン推定アルゴリズムは次のように書き換えることができる。

まず、事後推定ベクトル $\widehat{\mathbf{X}}_p[n|n, \ell]$ における 1 行目から QP 行目の算出に必要な要素のみを書き下すと

$$\widehat{\mathbf{X}}_p[n|n, \ell] = \widehat{\mathbf{X}}_p[n|n-1, \ell] + \mathcal{K}_p[n, \ell] \left\{ \mathbf{Y}_p[n, \ell] - \mathcal{M}_p[n, \ell] \widehat{\mathbf{X}}_p[n|n-1, \ell] \right\}$$

のように、網掛けした要素が必要となる。ここで $\eta_p = QP(L_p - 1)$ である。改めて $\mathbf{a}[i:j]$ はベクトル \mathbf{a} の i 行目から j 行目の要素、 $A[i:j, k:l]$ は行列 A の i 行目から j 行目かつ、 k 列目から l 列目の要素を指定する記号としたとき、事後推定ベクトル $\widehat{\mathbf{X}}_p[n|n, \ell]$ の 1 行目から QP 行目は

$$\begin{aligned} \widehat{\mathbf{X}}_p[n|n, \ell][1:QP] &= \widehat{\mathbf{X}}_p[n|n-1, \ell][1:QP] + \mathcal{K}_p[n, \ell][1:QP, 1:Q] \\ &\quad \cdot \left\{ \mathbf{Y}_p[n, \ell] - \mathcal{M}_p[n, \ell] \widehat{\mathbf{X}}_p[n|n-1, \ell] \right\} \\ &= \mathcal{K}_p[n, \ell][1:QP, 1:Q] \mathbf{Y}_p[n, \ell] \end{aligned} \quad (3.25)$$

となる。

次いで表 3.3 手順 2. のカルマンゲイン行列 $\mathcal{K}_p[n, \ell]$ は, 式 (3.25) より $\mathcal{K}_p[n, \ell]$ の 1 行目から QP 行目かつ 1 列目から Q 列目の要素のみに着目すると

$$\mathcal{K}_p[n, \ell] = \left\{ \begin{array}{l} \mathcal{P}_p[n|n-1, \ell] \mathcal{M}_p^H[n, \ell] \\ \cdot \left\{ \mathcal{M}_p[n, \ell] \mathcal{P}_p[n|n-1, \ell] \mathcal{M}_p^H[n, \ell] + \mathbf{C}_{E_p}[n, \ell] \right\}^{-1} \end{array} \right\}$$

$$\left[\begin{array}{c} \overbrace{\phantom{\mathcal{K}_p[n, \ell]}}^{Q} \\ \underbrace{\phantom{\mathcal{K}_p[n, \ell]}}_{\eta_p} \end{array} \right] = \left[\begin{array}{c} \overbrace{\phantom{\mathcal{K}_p[n, \ell]}}^{Q} \\ \underbrace{\phantom{\mathcal{K}_p[n, \ell]}}_{\eta_p} \end{array} \right] \cdot \left\{ \mathbf{S}[n, k] \mathbf{C}_\Delta[n, k] \mathbf{S}^H[n, k] + \mathbf{C}_{E_p}[n, \ell] \right\}^{-1}$$

のように, 網掛けした要素が必要となる. 以上よりカルマンゲイン行列 $\mathcal{K}_p[n, \ell]$ の 1 行目から QP 行目は

$$\begin{aligned} \mathcal{K}_p[n, \ell][1:QP, 1:Q] &= \left\{ \mathcal{P}_p[n|n-1, \ell][1:QP, 1:QP] \mathcal{M}_p^H[n, \ell][1:QP, 1:Q] \right\} \\ &\cdot \left\{ \mathcal{M}_p[n, \ell][1:QP, 1:Q] \mathcal{P}_p[n|n-1, \ell][1:QP, 1:QP] \mathcal{M}_p^H[n, \ell][1:QP, 1:Q] + \mathbf{C}_{E_p}[n, \ell] \right\}^{-1} \\ &= \left\{ \mathbf{C}_\Delta[n, k] \mathbf{S}^H[n, k] \right\} \left\{ \mathbf{S}[n, k] \mathbf{C}_\Delta[n, k] \mathbf{S}^H[n, k] + \mathbf{C}_{E_p}[n, \ell] \right\}^{-1} \end{aligned} \quad (3.26)$$

となる.

最後に表 3.3 手順 1. の事前推定誤差共分散行列 $\mathcal{P}_p[n|n-1, \ell]$ は, 式 (3.26) より $\mathcal{P}_p[n|n-1, \ell]$ の 1 行目から QP 行目かつ 1 列目から QP 列目の要素のみに着目し, $R_{\delta_p}(n+1)$ の行列の構造を考慮すれば

$$\mathcal{P}_p[n|n-1, \ell] = \mathcal{F}_p \mathcal{P}_p[n-1|n-1, \ell] \mathcal{F}_p^H + \mathcal{G}_p \mathbf{C}_{\Delta_p}[n, \ell] \mathcal{G}_p^H$$

$$\left[\begin{array}{c|c} \overbrace{\phantom{\mathcal{P}_p[n|n-1, \ell]}}^{Q} & \mathbf{O}_{QP \times \eta_p} \\ \hline \mathbf{O}_{\eta_p \times QP} & \phantom{\mathcal{P}_p[n|n-1, \ell]} \end{array} \right] = \left[\begin{array}{c|c} \overbrace{\phantom{\mathcal{P}_p[n|n-1, \ell]}}^{Q} & \mathbf{O}_{QP \times \eta_p} \\ \hline \mathbf{O}_{\eta_p \times QP} & \phantom{\mathcal{P}_p[n|n-1, \ell]} \end{array} \right] + \left[\begin{array}{c|c} \overbrace{\phantom{\mathcal{P}_p[n|n-1, \ell]}}^{Q} & \mathbf{O}_{QP \times \eta_p} \\ \hline \mathbf{O}_{\eta_p \times QP} & \mathbf{O}_{\eta_p} \end{array} \right]$$

のように, 網掛けした要素が必要となる. これらより事前誤差共分散行列 $\mathcal{P}_p[n|n-1, \ell]$ の

表 3.4: 雑音抑圧法を用いたチャネルゲイン推定アルゴリズム

<p>[Initialization]</p> $C_{\Delta_p}[n, \ell] = E \left[H_{ave}^{(q,p)}[n, \ell] \left\{ H_{ave}^{(q,p)}[n, \ell] \right\}^* \right] = J_0(0),$ $C_{E_p}[n, \ell] = E \left[\mathbf{E}_p[n, \ell] \mathbf{E}_p^H[n, \ell] \right] = \sigma_v^2 I_Q$ <p>[Iteration]</p> <ol style="list-style-type: none"> 1. $\mathcal{P}_{p1}[n, \ell] = C_{\Delta_p}[n, \ell]$ 2. $\mathcal{K}_{p1}[n, \ell] = \left\{ \mathcal{P}_{p1}[n, \ell] \mathbf{S}^H[n, k] \right\} \left\{ \mathbf{S}[n, k] \mathcal{P}_{p1}[n, \ell] \mathbf{S}^H[n, k] + C_{E_p}[n, \ell] \right\}^{-1}$ 3. $\widehat{\mathbf{H}}_{p1}[n, \ell] = \mathcal{K}_{p1}[n, \ell] \mathbf{Y}_p[n, \ell]$ 4. $n = n + 1$ go back 1.

1 行目から QP 行目は

$$\begin{aligned}
 \mathcal{P}_p[n|n-1, \ell][1:QP, 1:QP] &= \\
 \mathcal{F}_p[1:QP, 1:QP] \mathcal{P}_p[n-1|n-1, \ell][1:QP, 1:QP] \mathcal{F}_p^H[1:QP, 1:QP] & \\
 + \mathcal{G}_p C_{\Delta_p}[n, \ell] \mathcal{G}_p^H & \quad (3.27) \\
 &= O + C_{\Delta_p}[n, \ell] \\
 &= C_{\Delta_p}[n, \ell]
 \end{aligned}$$

と書き表すことができる。

新たに事前誤差共分散行列を $\mathcal{P}_{p1}[n, \ell]$, カルマンゲイン行列を $\mathcal{K}_{p1}[n, \ell]$, チャネルゲイン推定値を $\widehat{\mathbf{H}}_{p1}[n, \ell]$ とおけば, 表 3.3 のチャネルゲイン推定アルゴリズムは式 (3.25) ~ (3.27) より

$$\left. \begin{aligned}
 1. \quad \mathcal{P}_{p1}[n, \ell] &= C_{\Delta_p}[n, \ell] \\
 2. \quad \mathcal{K}_{p1}[n, \ell] &= \left\{ \mathcal{P}_{p1}[n, \ell] \mathbf{S}^H[n, k] \right\} \left\{ \mathbf{S}[n, k] \mathcal{P}_{p1}[n, \ell] \mathbf{S}^H[n, k] + C_v[n, k] \right\}^{-1} \\
 3. \quad \widehat{\mathbf{H}}_{p1}[n, \ell] &= \mathcal{K}_{p1}[n, \ell] \mathbf{Y}_p[n, \ell]
 \end{aligned} \right\} \quad (3.28)$$

のように表される。最後に雑音抑圧法を用いたチャネルゲイン推定アルゴリズムを表 3.4 に示す。表 3.4 のアルゴリズムは, 観測シンボルベクトル $\mathbf{Y}_p[n, \ell]$ に対してチャネルゲインの共分散行列 $\mathcal{P}_{p1}[n, \ell]$ と雑音の共分散行列 $C_v[n, k]$ から構成される重み行列 $\mathcal{K}_{p1}[n, \ell]$ が掛けられ, チャネルゲインを推定している。これは 2.4.2 節において説明した最小分散推定法に一致しており, 雑音抑圧法を MIMO MC-CDMA 通信方式におけるチャネルゲイン推定に応用した場合, 推定されたチャネルゲインは, 真のチャネルゲインとの誤差分散を最小にするように推定されていることが示された。

ここで, 雑音抑圧法を用いたチャネルゲイン推定法について考察する。まず提案手法の状態方程式は, 0 と 1 のみによって構成される状態遷移行列 \mathcal{F}_p を用いてチャネルゲインの

表 3.5: 従来手法と提案手法の演算量比較

	Conv.[8]	Prop.
1. $\mathcal{P}[n n-1, \ell]$	$2Q^2P^2L_c^2$	0
2. $\mathcal{K}[n, \ell]$	$Q^3(PL_c + 1) + Q^2(2P^2L_c + P)$	$Q^3(P + 1) + Q^2P + Q^2P^2$
3. $\widehat{\mathcal{X}}[n n-1, \ell]$	QPL_c	0
4. $\widehat{\mathcal{X}}[n n, \ell]$	$Q^2PL_c + QP(L_c + 1)$	Q^2P
5. $\widehat{\mathcal{H}}[n, \ell]$	0	0
6. $\mathcal{P}[n n, \ell]$	$Q^3P^3L_c^2 + Q^2P^2L_c$	0
Total	$Q^3(P^3L_c^2 + PL_c + 1) + Q^2(2P^2L_c^2 + 3P^2L_c + PL_c + P) + Q(PL_c + P)$	$Q^3(P + 1) + Q^2(P^2 + 2P)$

変動を表現している．つまり従来手法の状態方程式のように移動体の速度に依存する AR 係数を含まないため，提案手法のチャンネルゲイン推定精度は移動速度の影響を受けないと考えられる．

また従来手法は，AR 係数と表 3.1 の手順 1 に用いられる駆動源ベクトル $\Delta_c[n, \ell]$ の自己相関行列 $C_{\Delta_c}[n, \ell]$ を算出するために，最大ドップラー周波数 f_D を必要としていた．しかしながら本手法は，式 (3.22) の駆動源ベクトル $\Delta_p[n+1, \ell]$ をチャンネルゲインベクトル $\mathbf{H}_{ave}[n+1, \ell]$ と置くことにより，最大ドップラー周波数 f_D を算出する必要がない．

具体的には，表 3.4 の手順 1 において駆動源ベクトル $\Delta_p[n+1, \ell] = \mathbf{H}_{ave}[n+1, \ell]$ の自己相関行列 $C_{\Delta_p}[n, \ell] = E \left[\mathbf{H}_{ave}^{(q,p)}[n, \ell] \left\{ \mathbf{H}_{ave}^{(q,p)}[n, \ell] \right\}^* \right]$ より，行列 $C_{\Delta_p}[n, \ell]$ の対角成分は同一ブロックにおけるチャンネルゲインの相関 $E \left[\mathbf{H}_{ave}^{(q,p)}[n, \ell] \left\{ \mathbf{H}_{ave}^{(q,p)}[n, \ell] \right\}^* \right] = J_0(0)$ となる．それゆえ本手法は最大ドップラー周波数算出のために移動速度を事前情報として必要としない．

加えて本手法は AR 係数の算出を必要とせず，また多くの 0 を要素に多く含む状態遷移行列 \mathcal{F}_p を表 3.3 の手順 1, 3 に用いられていることより，表 3.4 における整理されたチャンネルゲイン推定アルゴリズムは大幅に演算量が軽減されている．

3.5 演算量評価

表 3.5 は，表 3.1 の従来手法と表 3.4 の提案手法で必要な演算量を示したものである．また図 3.7 は，表 3.5 の AR 次数を $L_c = 2$ のように設定したとき，送受信アンテナ数を $P = Q = 2$,

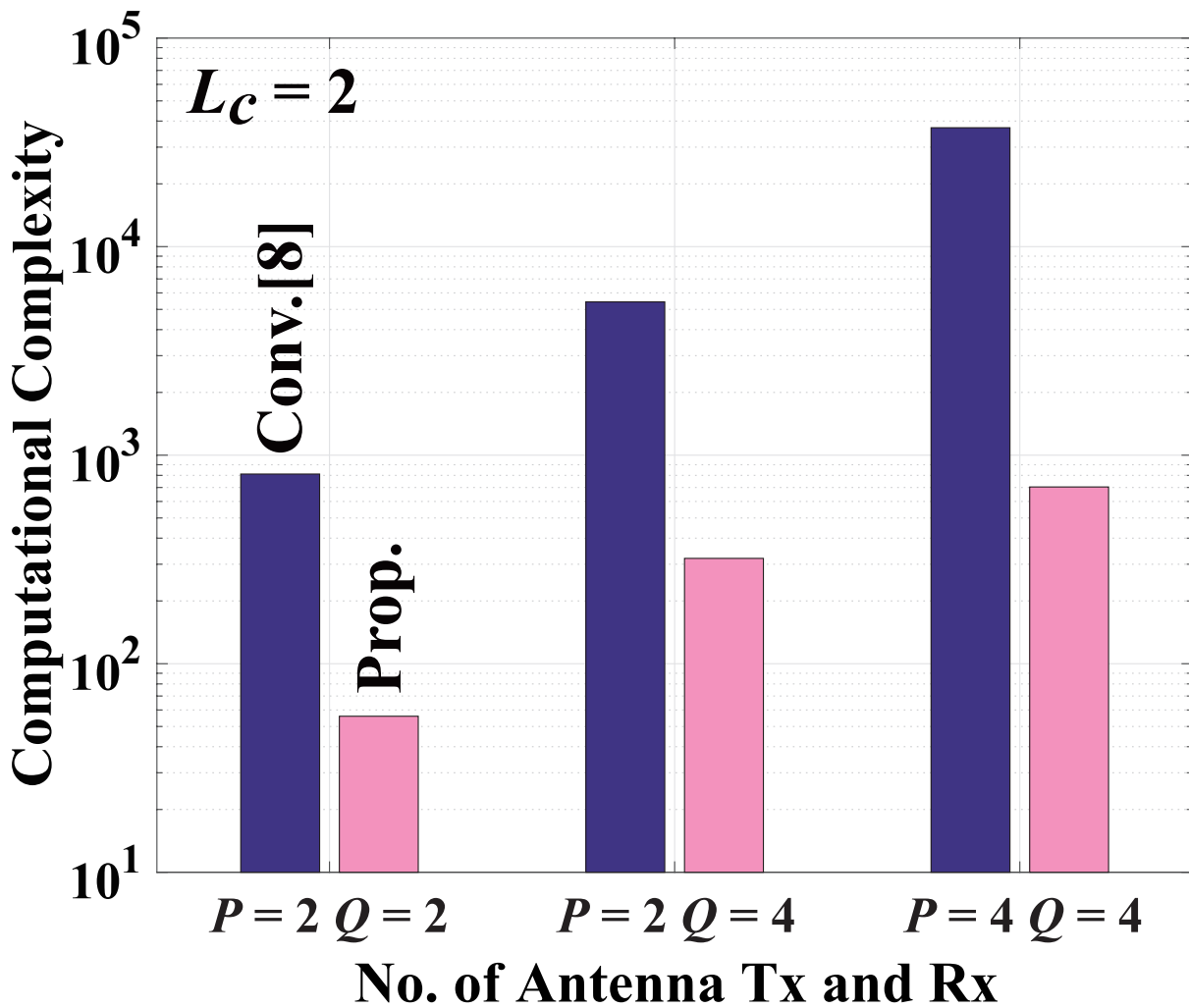


図 3.7: 送受信アンテナ数の変化に対する演算量評価

$P = 2, Q = 4$, $P = Q = 4$ のように変化させた場合の従来手法と本手法の演算量を示している。ただし、従来手法における AR 係数算出は様々な手法が存在するため、本論文では AR 係数算出に必要な演算量は考慮しないこととする。従って、表 3.1 に示される従来手法の演算量は、実際には更に増加することに注意しなければならない。

表 3.5 と図 3.7 より、従来手法は演算量が AR 次数 L_c に依存していることに対して、提案手法は状態ベクトルのサイズ L_p に依存することなくチャンネルゲイン推定を可能としている。これは、提案手法における式 (3.22) で与えられる $QPL_p \times QP$ 駆動源行列 \mathcal{G}_p と、式 (3.24) で与えられる $Q \times QPL_p$ 観測遷移行列 $M_p[n, \ell]$ それぞれの零以外の要素が、送受信アンテナ P, Q のサイズのみ依存する行列であることに起因する。それゆえ表 3.3 の手順 2 ~ 5 において、演算量は L_p に依存せず、結果として表 3.4 のチャンネルゲイン推定に要す

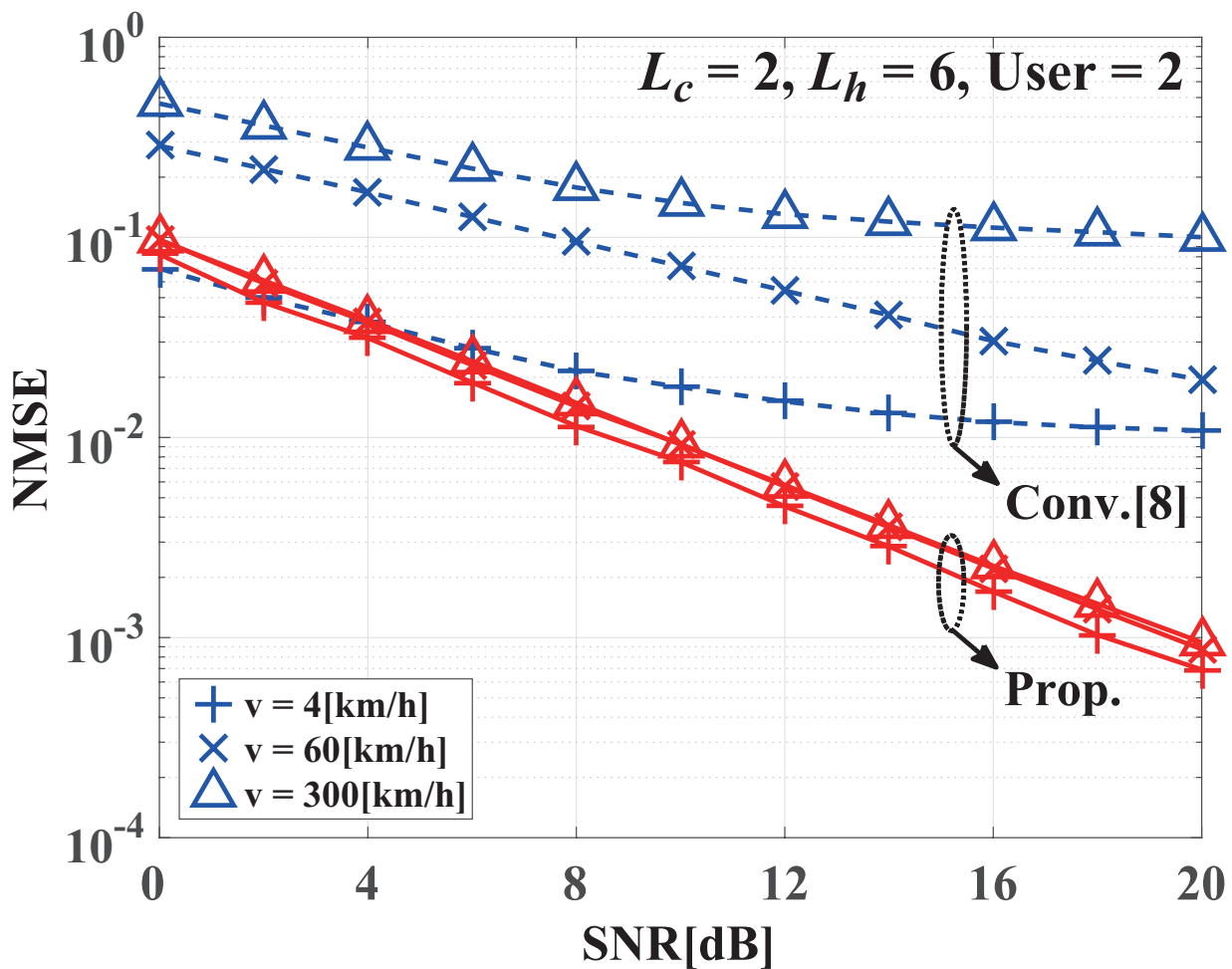


図 3.8: 移動速度 v の変化に対する従来手法と提案手法のチャンネルゲイン推定精度比較

る演算量は $L_p (> 0)$ のサイズに無関係である。

3.6 計算機シミュレーション

本章では、提案手法の有効性を確認するために計算機シミュレーションを行い、その結果について考察する。なおシミュレーション諸元やチャンネルゲインの推定精度の評価量、フェージングと雑音の条件は、3.3.3 節に用いたものと同様である。

3.6.1 チャンネルゲイン推定精度評価

図 3.8 は送信アンテナ数 $P = 2$ 、受信アンテナ数 $Q = 2$ 、遅延波数 $L_h = 6$ の MIMO MC-CDMA 通信システムにおける、移動体の速度を変化させた場合のチャンネルゲイン推定精度

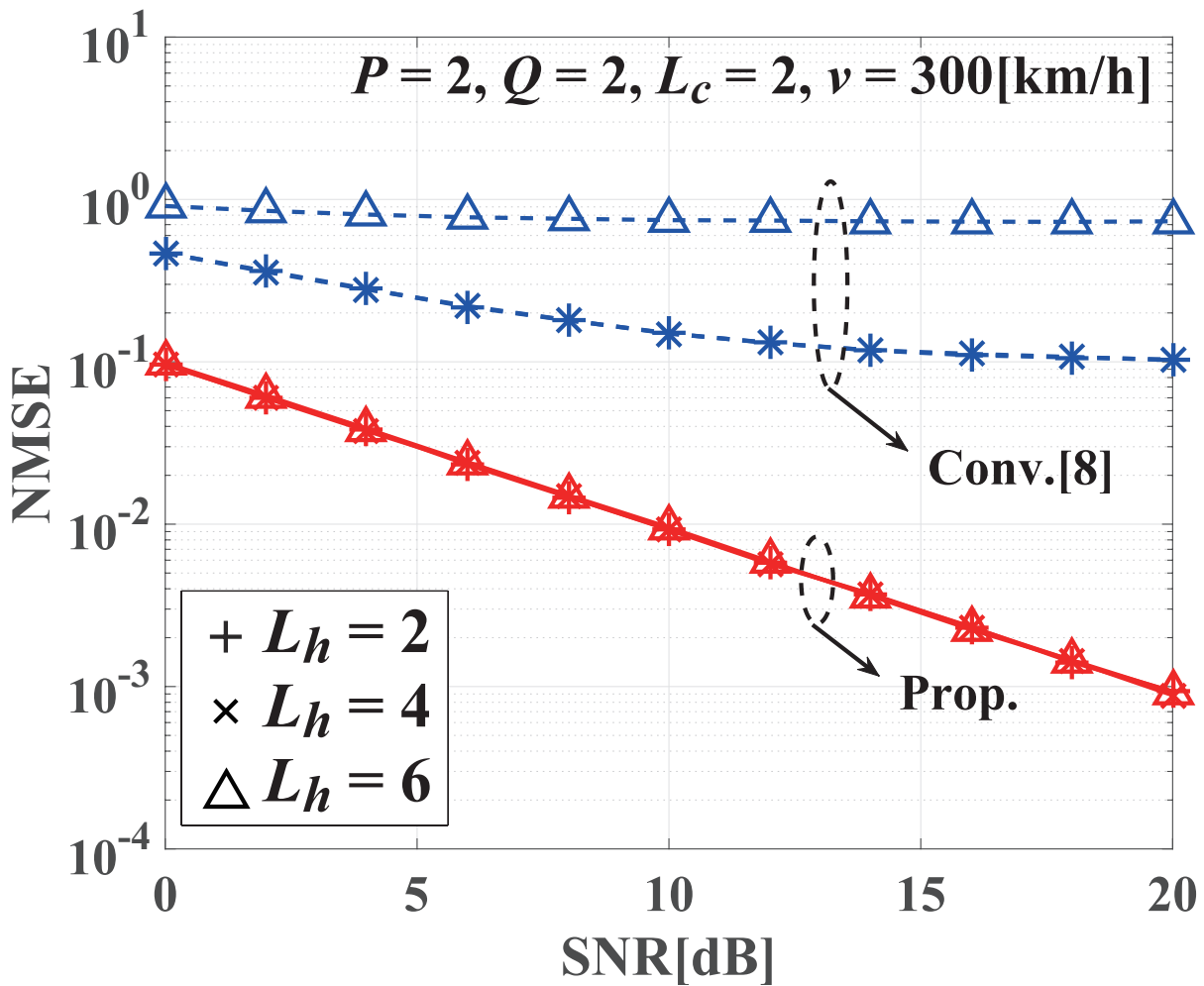


図 3.9: 遅延波数 L_h に対する従来手法と提案手法のチャネルゲイン推定精度比較

を NMSE を用いて評価したものである。移動速度は $v = 4, 60, 300[\text{km/h}]$ のように変化させた。また従来手法の AR 次数は 3.3.3 節と同様に AR 次数として最適とされている $L_c = 2$ とした。

図 3.8 より、従来手法は移動体の速度の変化に従いチャネルゲインの推定精度が劣化している。3.3.3 節でも述べたが、Step:1 で移動速度に対して適切でない AR 次数を用いて算出された AR 係数を Step:2 の状態遷移行列 \mathcal{F}_c に用いたことより、チャネルゲインの推定精度が劣化したと考えられる。

一方提案手法は、移動体の速度に関係なく従来手法よりも推定精度が良好であり、常に一定のチャネルゲイン推定精度を保っている。これは移動体の速度に依存する変数を式 (3.20) の状態方程式に用いていないからである。

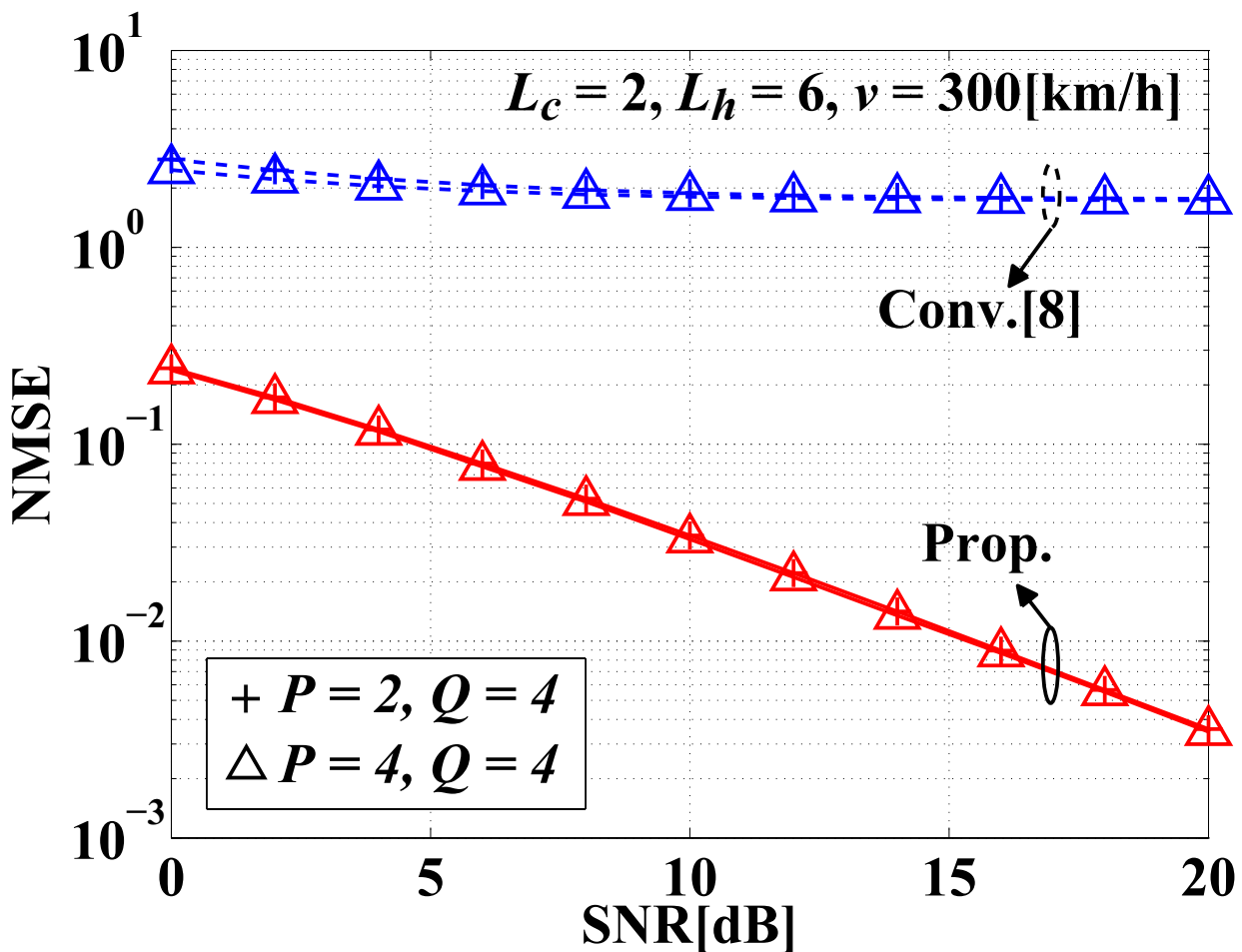


図 3.10: 送信アンテナ数 P の変化に対する従来手法と提案手法のチャネルゲイン推定精度比較

図 3.9 は送受信アンテナ数をそれぞれ $P = 2, Q = 2$ となる MIMO MC-CDMA 通信システムにおいて、時速 $v = 300[\text{km/h}]$ の高速移動環境を想定した場合、遅延波数を $L_h = 2, 4, 6$ のように変化させたときのチャネルゲインの推定精度評価をした。従来手法の AR 次数は $L_c = 2$ のように設定している。また図 3.10 は時速 $v = 300[\text{km/h}]$ の高速移動環境でかつ、遅延波数が $L_h = 6$ のような劣悪な環境において、送信アンテナ数の変化に対するチャネルゲインの推定精度を評価したものである。送信アンテナ数を $P = 2, 4$ 、受信アンテナ数が $Q = 4$ の MIMO MC-CDMA 通信システムとし、従来手法の AR 次数は $L_c = 2$ のように設定した。

図 3.9, 3.10 より提案手法のチャネルゲイン推定精度は遅延波数、および送信アンテナ数に殆ど依存しないことが確認される。しかしながら、これに関する理論的根拠を現段階で

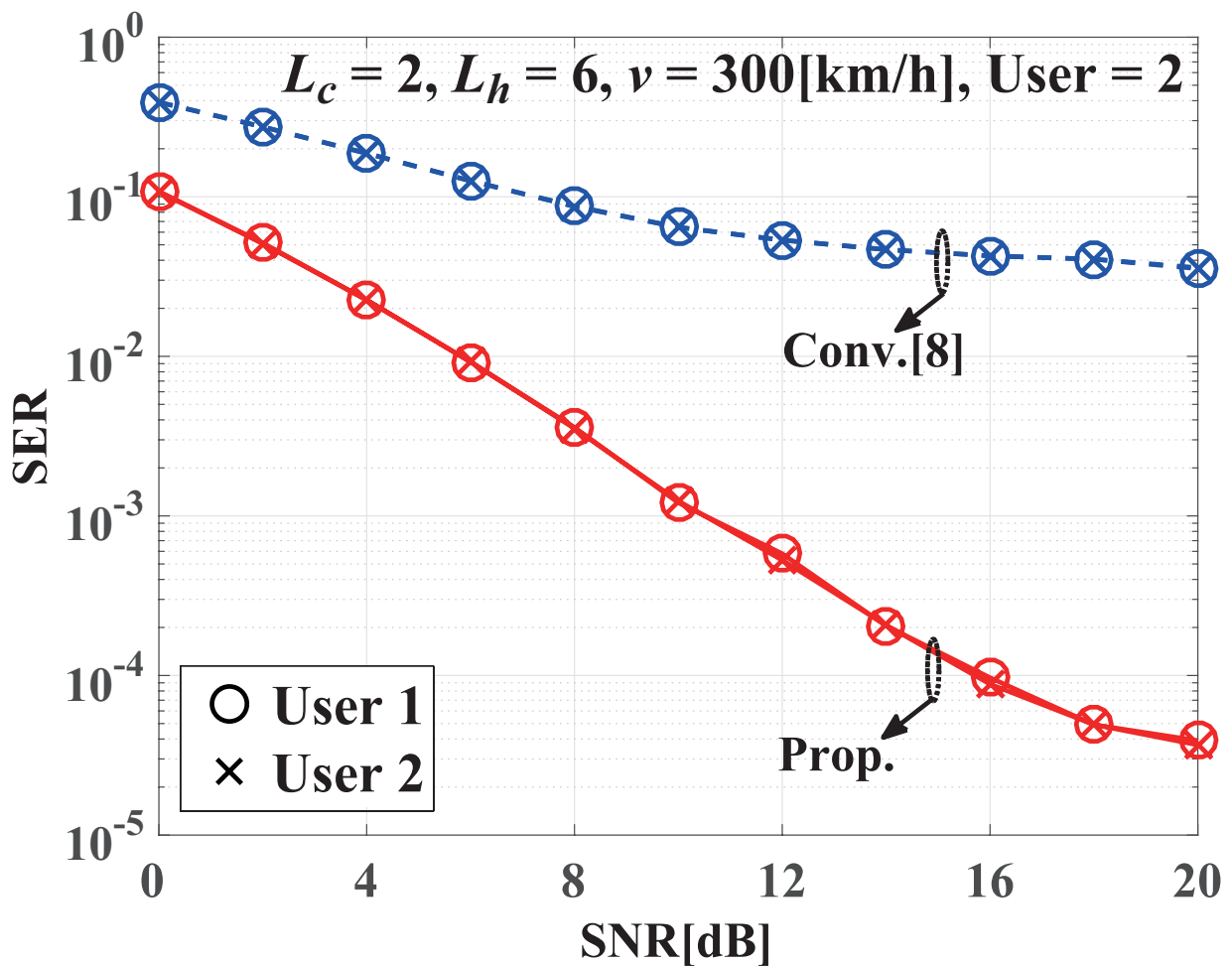


図 3.11: 従来手法と提案手法のシンボルエラーレート評価

は定量的に示すのは困難であり、今後考察を行う予定である。

また図 3.8~3.10 より、従来手法は移動速度を既知情報として与えているにもかかわらず、チャンネルゲイン推定精度が劣化している。それに対して移動速度が未知である本手法は従来手法に比べ、低 SNR においても良好な NMSE を得ていることがわかる。

さらに提案手法は、雑音抑圧法を用いても、従来手法よりチャンネルゲイン推定精度が高いことがわかる。

3.6.2 シンボルエラーレート評価

従来手法と提案手法より得られたチャンネルゲインを用いて、シンボルエラーレート (SER: Symbol Error Ratio) 評価をする。

ここで逐次的にチャンネルゲインを推定するために、パイロットシンボルを挿入する必要がある。一般的に文献 [24] の図 3.2 に示されているように、各ブロック内のサブキャリアごとにパイロットシンボルを配置し、また受信側では、パイロットシンボルを従来手法と提案手法のそれぞれのチャンネルゲイン推定法に用いることで、チャンネルゲインを推定している。しかし本論文ではチャンネルゲイン推定に主眼を置いているため、パイロットシンボルの挿入方法については議論せず、すべてのサブキャリアにパイロット信号を挿入することとする。

以上のように推定されたチャンネルゲインの結果から従来手法、提案手法ともに最小二乗平均誤差 (MMSE : Minimum Mean Square Error) 基準 [6] の等化器を設計し、情報データを復元可能である。本論文では同様の手順を各受信ブロックごとに繰り返すことにより、逐次的に情報データを復元している。

図 3.11 は送受信アンテナ数がそれぞれ $P = 2, Q = 2$ の MIMO MC-CDMA 通信システムにおいて、送信フレームを数を 128 とし、時速 $v = 300[\text{km/h}]$ の高速移動環境でかつ遅延波数が $L_h = 6$ のような劣悪な通信環境を想定した。また、従来手法の AR 次数は $L_c = 2$ のように設定した。

図 3.11 より、提案手法は従来手法に比べ、チャンネルゲインの推定精度が向上したことにより良好な SER 特性を示している。従って、提案手法は高速移動体通信環境下においても対応可能な手法である。

3.7 結言

本章では、雑音抑圧法を用いた高速移動体通信における MIMO MC-CDMA 通信方式のためのチャンネルゲイン推定法について検討を行い、具体例によりその有効性を検証した。以下、本章で得られた検討結果を列記する。

1. 移動体の速度がチャンネルゲインの推定精度に依存する問題に対して、AR システムを用いない新たな状態空間モデルを構成することで解決できた。
2. 送受信アンテナ数増加に対して、0 と 1 を多く含む状態遷移行列を構成することにより低演算量であることから、シンプルで実用的な手法といえる。

また、最終的に得られたチャネルゲイン推定アルゴリズムは、最小分散推定法に帰着されることを明らかにした。

以上のことより、提案手法は移動体の速度に影響されることなく、また従来手法と比較して低演算量で良好なチャネルゲイン推定が可能な手法と言える。

今後の課題として、様々なフェージング環境下でのチャネルゲイン推定精度評価、さらなる演算量の軽減などが挙げられる。

第4章 自動カラオケ音源生成システムへの応用

本章では、雑音抑圧法を自動カラオケ生成に応用することにより、ステレオのカラオケ音源をリアルタイムに生成可能な手法を提案する。

4.1 緒言

近年、スマートフォンに代表されるように小型で高性能な電子機器が急速に普及している。同時に通信速度の向上も相まって、音楽や映像データなどのデジタルコンテンツをいつでもどこでも楽しむことのできる時代となった。

ところで広い世代に渡って親しまれている音楽コンテンツとして「カラオケ」が挙げられる [25]。カラオケを楽しむためには音響設備の整った店 (カラオケ BOX 等) を利用することが一般的である。しかしながら生活の多様化に伴い、時間的、金銭的な余裕が無くカラオケ BOX に行くことのできないユーザやカラオケに行く前に家で練習をしたいというユーザ、さらに花見や忘年会などの各種イベントの際にカラオケをしたいなど、ユーザはカラオケをする場所を限定しない新たな楽しみ方を求めていると考えられる。

そこで本論文では、市販されている CD (Compact Disc) や音楽配信サイトなどから得ることのできる、ヴォーカル信号に実際の楽器による伴奏信号が加わった楽曲信号より、手軽にカラオケ信号を生成可能とする手法について議論を行う。

はじめに以降の議論のために、本論文が対象とする楽曲信号の構成方法と用語の定義について図 4.1 を交えて説明する。

楽曲信号はヴォーカルやギター、ドラムなど個別に録音された音素材に対してエフェクトや音像定位処理を施し、さらにステレオトラックとして出力することによりステレオ感を有するステレオ楽曲信号が構成される [26]。このとき一般的にヴォーカルは中央に定位するように配置されるため、ヴォーカル信号 $d(n)$ は図 4.1 のように左右のチャンネルに等し

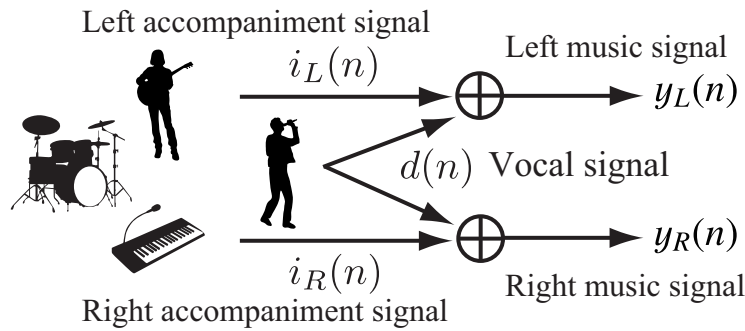


図 4.1: ステレオ楽曲信号の構成

く含まれるものとする。また各楽器は任意の位置に定位しているものとし、左チャンネルの伴奏信号を $i_L(n)$ 、右チャンネルの伴奏信号を $i_R(n)$ 、左チャンネルの伴奏信号とヴォーカル信号が加わった左楽曲信号を $y_L(n)$ 、右チャンネルの伴奏信号とヴォーカル信号が加わった右楽曲信号を $y_R(n)$ として議論を進める。

図 4.1 のように複数の信号が混合した音源より所望の信号を取り出す手法として、独立成分分析に基づくブラインド音源分離 [27] が知られている。しかし複数音源が統計的に互いに独立であるという仮定を必要とするため、本論文が対象する楽曲信号への適用は困難である。それゆえ本論文では、ヴォーカル信号が中央定位するという特徴に着目したステレオ楽曲信号に対するカラオケ信号を生成する方法について議論を行う。

上記の特徴を用いたカラオケ信号生成手法の 1 つとして、**Out of Phase Stereo(OOPS)** 手法 [28] が挙げられる。これは左楽曲信号から右楽曲信号を減算(その逆でもよい)することで左右の楽曲信号に同程度含まれるヴォーカル信号を除去している。これは減算のみの処理なので非常にシンプルでかつ、リアルタイムにカラオケ信号を生成できるものの、生成されたカラオケ信号はモノラル信号となる問題がある。また中央に定位している楽器の信号も除去される問題も存在する。

一方、ステレオ構造を維持したカラオケ信号を生成可能な手法として、差信号合成型カラオケ信号生成法 [29] が挙げられる。これは、ステレオ楽曲信号を周波数領域に変換し左右チャンネルのスペクトルの差を算出した後に、差の小さいスペクトルをヴォーカルとして抑圧し、差の大きいスペクトルを暫定カラオケスペクトルとして抽出する手法である。この手法の特徴は、暫定カラオケスペクトルにミュージカルノイズが発生するものの、暫定カラオケ信号に左右楽曲信号の差信号を加えることにより、上記ノイズを抑圧できる点に

ある。しかし、この手法の差信号を加える操作は左右のチャンネルに対して同一成分の信号を加えることを意味する。その結果、左右のチャンネルに別々の成分が入るステレオ楽曲信号の構造を考慮していない補正方法のため、左チャンネルもしくは右チャンネルに本来は存在しない音が出力され、ステレオ感が減少する問題が生じる。

上記に鑑み、本章ではヴォーカルが中央定位していることを前提とし、差信号合成型カラオケ信号生成法 [29] を従来手法と位置付け、著者らが既に提案した雑音抑圧技術 [3] を応用したカラオケ信号生成手法を提案する。

4.2 問題設定

図 4.1 を用いてステレオ楽曲信号を定量的に表現する。

ステレオ楽曲信号は、左伴奏信号 $i_L(n)$ と右伴奏信号 $i_R(n)$ に中央定位するようにヴォーカル信号 $d(n)$ が加わるため、左楽曲信号 $y_L(n)$ と右楽曲信号 $y_R(n)$ はそれぞれ

$$\left. \begin{aligned} y_L(n) &= d(n) + i_L(n) \\ y_R(n) &= d(n) + i_R(n) \end{aligned} \right\} \quad (4.1)$$

と表わす。ここで n はサンプル時刻である。また一般的に伴奏はヴォーカルに合わせて演奏されるため、両者の信号間の独立性は仮定できないと考えられる。従って左右伴奏信号とヴォーカル信号の平均値および、左右伴奏信号とヴォーカル信号との相関をそれぞれ

$$\left. \begin{aligned} E[i_L(n)] &= E[i_R(n)] = E[d(n)] = 0, \\ E[i_L(n)d(n)] &= 0, E[i_R(n)d(n)] = 0 \end{aligned} \right\} \quad (4.2)$$

のように仮定して以降の議論を行う。ただし $E[\cdot]$ は期待値演算を表わす。

提案法の目的は、ステレオ楽曲信号 $y_L(n)$, $y_R(n)$ のみから、カラオケ信号 $i_L(n)$, $i_R(n)$ を生成することである。

4.3 従来 of 伴奏信号抽出手法

本章では、ステレオ楽曲信号の左右チャンネル間の差信号を用いた差信号補正型カラオケ信号生成法 [29] について簡単に説明し、その問題点を指摘する。

4.3.1 Step1:伴奏信号の抽出

各時刻 n におけるステレオ楽曲信号 $y_L(n)$, $y_R(n)$ に対して, フレーム長 L_F , シフト長 L_S の短時間フーリエ変換 (STFT : Short-Time Fourier Transform)[30] を施すことにより, 得られたステレオ楽曲スペクトル $Y_L(\ell, k)$, $Y_R(\ell, k)$ が

$$\left. \begin{aligned} Y_L(\ell, k) &= D(\ell, k) + I_L(\ell, k) \\ Y_R(\ell, k) &= D(\ell, k) + I_R(\ell, k) \end{aligned} \right\} \quad (4.3)$$

のように与えられる. ここで ℓ はフレーム番号, k は周波数ビン番号, $D(\ell, k)$ はヴォーカルスペクトル, $I_L(\ell, k)$ と $I_R(\ell, k)$ はそれぞれ左チャンネルと右チャンネルの伴奏スペクトルである. また人間の聴覚は位相情報に対して鈍感であることより, 振幅スペクトルのみを扱うこととし, さらに位相スペクトル $\phi_L(\ell, k)$, $\phi_R(\ell, k)$ は元の楽曲スペクトルを用いて次式

$$\left. \begin{aligned} \phi_L(\ell, k) &= \tan^{-1} \frac{\Im[Y_L(\ell, k)]}{\Re[Y_L(\ell, k)]} \\ \phi_R(\ell, k) &= \tan^{-1} \frac{\Im[Y_R(\ell, k)]}{\Re[Y_R(\ell, k)]} \end{aligned} \right\} \quad (4.4)$$

により, あらかじめ算出しておく. この位相情報は, IFFT の際に振幅スペクトルに付与し音声信号を生成するために活用されることが文献 [29] で述べられている. なお $\Re[\cdot]$, $\Im[\cdot]$ はそれぞれ複素数の実数部と虚数部を取り出す記号を意味する.

従来手法 [29] は, ヴォーカルスペクトルと伴奏スペクトルを判別するために, 次式

$$A_e(\ell, k) = \frac{\{|Y_L(\ell, k) - Y_R(\ell, k)|\}^2}{\{|Y_L(\ell, k) + Y_R(\ell, k)|\}^2} \quad (k_L \leq k \leq k_H) \quad (4.5)$$

で得られる $A_e(\ell, k)$ を利用している. ここで k_L , k_H はそれぞれ, ヴォーカルスペクトルが存在する周波数帯域 (以降, ヴォーカル帯域と称する) の下限周波数ビン番号と上限周波数ビン番号である.

$A_e(\ell, k)$ は, 左右の楽曲スペクトルが等しいほど 0 に近づき異なるほど 1 に近くなる. そこで $A_e(\ell, k)$ に対して適当な閾値 α を設定し, α 未満となった周波数ビンに対する左右の楽曲スペクトル $|Y_L(\ell, k)|$, $|Y_R(\ell, k)|$ をヴォーカルスペクトルとして抑圧する. 逆に α 以上となった周波数ビンに対する左右の楽曲スペクトルは暫定伴奏スペクトルとして残す. これらの処理を 1 フレーム内のすべての周波数ビンに対して実行することにより, 暫定カラオ

ケスペクトル $\check{I}_L(\ell, k)$, $\check{I}_R(\ell, k)$ を次式

$$\left. \begin{aligned} |\check{I}_L(\ell, k)| &= \begin{cases} k_0 \cdot |Y_L(\ell, k)| & (0 \leq A_e(\ell, k) < \alpha) \\ |Y_L(\ell, k)| & (\alpha \leq A_e(\ell, k) \leq 1) \\ |Y_L(\ell, k)| & (k < k_L, k_H < k) \end{cases} \\ k &= 0, 1, \dots, [(L_F/2)]_G - 1 \\ |\check{I}_R(\ell, k)| &= \begin{cases} k_0 \cdot |Y_R(\ell, k)| & (0 \leq A_e(\ell, k) < \alpha) \\ |Y_R(\ell, k)| & (\alpha \leq A_e(\ell, k) \leq 1) \\ |Y_R(\ell, k)| & (k < k_L, k_H < k) \end{cases} \\ k &= 0, 1, \dots, [(L_F/2)]_G - 1 \end{aligned} \right\} \quad (4.6)$$

のように得ている。ここで、 $[\cdot]_G$ はガウス記号、 k_0 はヴォーカルスpekトルの抑圧係数とし、 α と k_0 は共に 0 以上 1 未満の範囲で設定されている。

4.3.2 Step2:差信号とカラオケ信号の生成

Step1 により算出された暫定カラオケスペクトル $\check{I}_L(\ell, k)$, $\check{I}_R(\ell, k)$ に対して IFFT をし、得られた暫定カラオケ信号を $\check{i}_L(n)$, $\check{i}_R(n)$ とする。

さて、文献 [31] によれば、周波数領域においてspeクトルの抑圧操作をすると周波数speクトルの孤立によるトーン性のノイズ (ミュージカルノイズ) が発生すると報告されている。それゆえヴォーカルスpekトルの抑圧操作により暫定カラオケspeクトル $\check{I}_L(\ell, k)$, $\check{I}_R(\ell, k)$ にもミュージカルノイズが重畳されると考えられる。そこで、左右チャンネル間の差信号 $\Delta y_L(n)$, $\Delta y_R(n)$ を次式

$$\left. \begin{aligned} \Delta y_L(n) &= y_L(n) - y_R(n) \\ &= i_L(n) - i_R(n) \\ \Delta y_R(n) &= y_R(n) - y_L(n) \\ &= i_R(n) - i_L(n) \end{aligned} \right\} \quad (4.7)$$

のように求め、暫定カラオケ信号 $\check{i}_L(n)$, $\check{i}_R(n)$ に加えることで、ミュージカルノイズを補償した確定カラオケ信号 $\widehat{i}_L(n)$, $\widehat{i}_R(n)$

$$\left. \begin{aligned} \widehat{i}_L(n) &= k_1 \cdot \check{i}_L(n) + k_2 \cdot \Delta y_L(n) \\ \widehat{i}_R(n) &= k_1 \cdot \check{i}_R(n) + k_2 \cdot \Delta y_R(n) \end{aligned} \right\} \quad (4.8)$$

を得ている。なお k_1 , k_2 は重み係数でありそれぞれ 0 以上 1 未満の範囲で設定する。

ここで Step2 について考察する。式 (4.7) により算出される左右チャンネル間の差信号 $\Delta y_L(n)$, $\Delta y_R(n)$ は、右チャンネルの楽曲信号から左チャンネルの楽曲信号を、または左チャンネルの楽曲

表 4.1: 従来手法と提案手法のカラオケ信号生成手順

従来手法	提案手法
ステレオ楽曲信号を周波数領域へ変換 (FFT:2 回)	
Step1:伴奏信号の抽出 <ul style="list-style-type: none"> ● 左右チャンネルのスペクトルより $A_e(\ell, k)$ を算出 ● ステレオの“伴奏スペクトル”を抽出 ● 左右チャンネルを時間領域へ変換 (IFFT:2 回) 	Step1:ヴォーカル信号の抽出 <ul style="list-style-type: none"> ● 左右チャンネルのスペクトルより $A_e(\ell, k)$ を算出 ● モノラルの“ヴォーカルスペクトル”を抽出 ● ヴォーカルスペクトルを時間領域へ変換 (IFFT:1 回)
Step2:カラオケ信号の生成 <ul style="list-style-type: none"> ● 暫定伴奏信号と差信号を合成 	Step2:カラオケ信号の生成 <ul style="list-style-type: none"> ● 左右チャンネル別々に雑音抑圧法 [3] を適用

信号から右チャンネルの楽曲信号を減算しており、これらは符号が異なる絶対値が等しい信号である。それゆえ式 (4.8) のように暫定カラオケ信号へ差信号を加えると、左右のチャンネルに同一成分の信号が存在することとなり、ステレオ感が減少すると考えられる。

そこで次章において、雑音抑圧法 [3] を応用した、ステレオ型カラオケ信号生成法を提案する。

4.4 雑音抑圧法を用いた伴奏信号抽出手法

表 4.1 は、従来手法のカラオケ信号生成法とこれから提案するカラオケ信号生成法の手順を簡単にまとめたものである。

提案手法は 2 章で説明した雑音抑圧法 [3] を自動カラオケ生成手法に応用する。これは、非抑圧対象信号の時間変化を表す状態方程式と、観測信号が得られる過程を記述した観測方程式から成る状態空間モデルを構成し、カルマンフィルタ理論 [1] を適用することで雑音抑圧を実現する。ここではあらかじめヴォーカル信号の統計量を算出する必要があるため、提案手法の Step1 では従来手法の Step1 を応用してヴォーカル信号を抽出する。次いで Step2 では、左右のチャンネルごとに独立した状態空間モデルを構成しステレオのカラオケ信号を生成するためのアルゴリズムを導出する。以下、提案手法の手順について詳しく説明する。

4.4.1 Step1:ヴォーカル信号の抽出

本節では、4.3.1節の伴奏信号判別法を用いたヴォーカル信号の抽出方法について、図4.2に従って説明する。

図4.2(a), (b)のように左右の楽曲信号に対してSTFTを施し、得られた左右の楽曲スペクトルから式(4.5)によって $A_e(\ell, k)$ を得た様子を図4.2(c)に示す。

左右のチャンネルに同程度の強度でヴォーカルスペクトルが存在する場合、その周波数ビンにおける $A_e(\ell, k)$ は極端に小さくなると考えられる。そこで図4.2(c)のように適当な閾値 α を設け、 $A_e(\ell, k)$ と α の大小関係を1フレーム内(ここでは ℓ 番目のフレーム)のすべての周波数ビン $k = 0, 1, \dots, [(L_F/2)]_G - 1$ ごとに比較し、 $A_e(\ell, k)$ が α 未満となったビンに対応する楽曲スペクトル $|Y_L(\ell, k)|$ を暫定ヴォーカルスペクトル $|\check{D}(\ell, k)|$ として残す(図4.2(d)の周波数ビン k_v に対応するスペクトル)。

また後述するようにStep2においては正確なヴォーカル信号の統計量を算出するため、 α 以上となった楽曲スペクトルは伴奏スペクトルとしてすべて抑圧する処理をする(図4.2(d)の周波数ビン k_a に対応するスペクトル)。

即ち、図4.2(e)に示す暫定ヴォーカルスペクトル $|\check{D}(\ell, k)|$ を以下

$$|\check{D}(\ell, k)| = \begin{cases} |Y_L(\ell, k)| & (0 \leq A_e(\ell, k) < \alpha) \\ 0 & (\alpha \leq A_e(\ell, k) \leq 1) \\ 0 & (k < k_L, k_H < k) \end{cases} \quad (4.9)$$

$$k = 0, 1, \dots, [(L_F/2)]_G - 1$$

のように設定する。式(4.9)からわかるように、 $|\check{D}(\ell, k)|$ は周波数ビン k の値とそれに対応した $A_e(\ell, k)$ の値に応じて $|Y_L(\ell, k)|$ 、あるいはゼロとなる。また式(4.9)は左楽曲スペクトル $|Y_L(\ell, k)|$ を用いて暫定ヴォーカルスペクトルを抽出しているが、ヴォーカルの中央定位の仮定より右楽曲スペクトル $|Y_R(\ell, k)|$ を用いてもよい。

次に閾値 α について考察を行う。図4.2(c)より、 α を0に近づけすぎると暫定ヴォーカルスペクトルは得られず、あるいは α を大きくしすぎると暫定ヴォーカルスペクトルは伴奏スペクトルを含むようになるため好ましくない。また $A_e(\ell, k)$ は、対象とする楽曲信号によって取りうる値が異なることが予想され、その都度 α を決定する必要がある。これらのことを踏まえると、閾値 α を一意に決定することは困難である。なお、4.5章の音質評価に用いる α は実験的に求めたものを利用している。

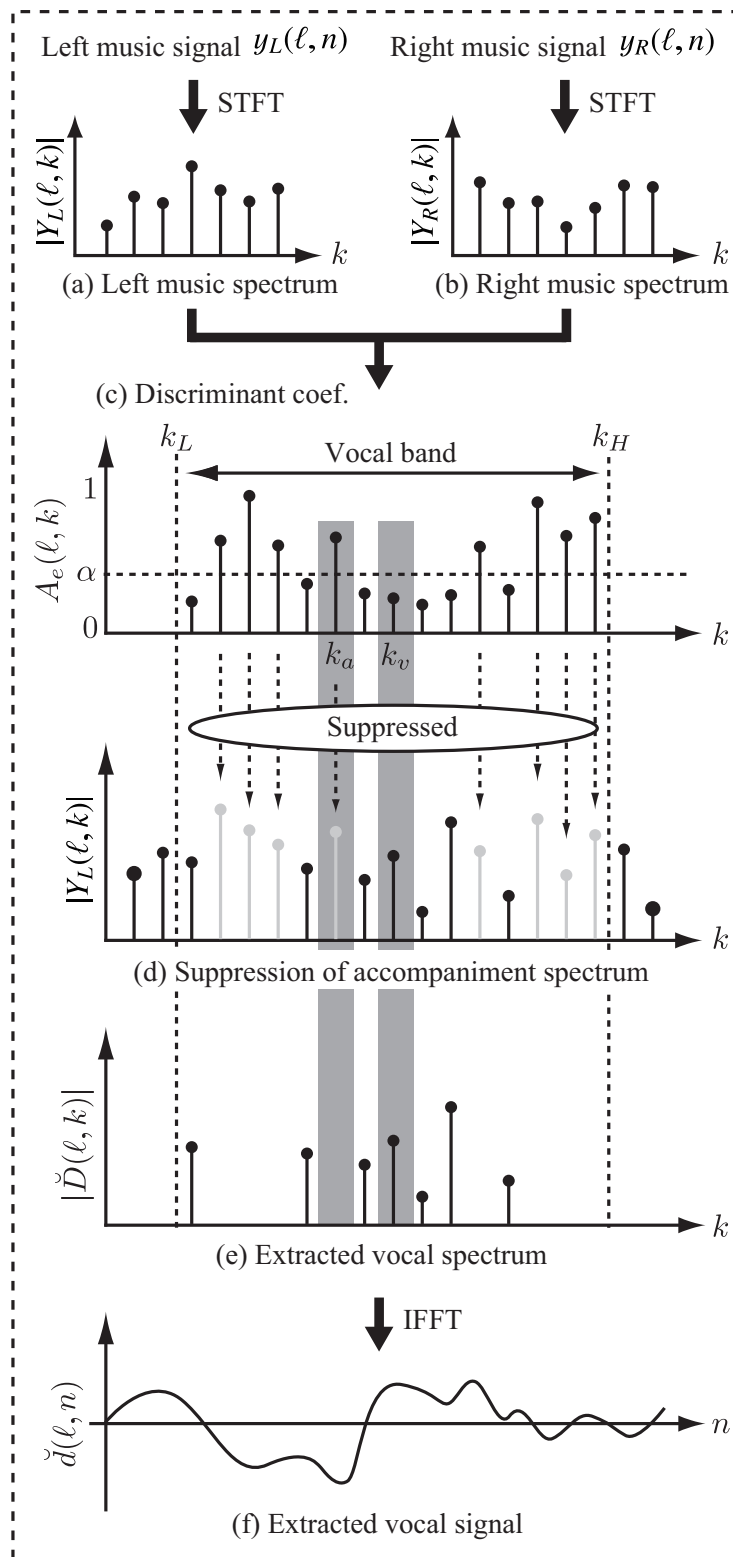


図 4.2: 提案手法 Step1 の流れ

抽出された暫定ヴォーカルスペクトル $|\check{D}(\ell, k)|$ は、これを IFFT をすることで図 4.2 (f) のような暫定ヴォーカル信号 $\check{d}(\ell, n)$ に変換される。STFT によりフレームごとに処理をしているため、1 フレーム前の暫定ヴォーカル信号とオーバーラップさせることにより、最終的に推定ヴォーカル信号 $\widehat{d}(\ell, n)$ を

$$\widehat{d}(\ell, n) = \check{d}(\ell - 1, n - L_s) + \check{d}(\ell, n) \quad (4.10)$$

のように得る。なおこれ以降オーバーラップ処理を終えた推定ヴォーカル信号を扱うため、フレーム番号 ℓ を省略し $\widehat{d}(n)$ と表記する。

4.4.2 Step2:雑音抑圧法 [3] を用いたカラオケ信号生成法

本節では、著者らが提案した雑音抑圧法 [3] を用いたカラオケ信号生成法について説明する。まず左チャンネルの伴奏信号 $i_L(n)$ の時間変化を表す状態方程式を

$$\mathbf{x}_L(n+1) = \Phi \mathbf{x}_L(n) + \boldsymbol{\delta}_L(n+1) \quad (4.11)$$

のように定義する。ここで L_p 次元状態ベクトル $\mathbf{x}_L(n+1)$ は、時刻 n から L_p サンプル過去の伴奏信号をそれぞれ並べたベクトルであり

$$\mathbf{x}_L(n+1) = [i_L(n+1), i_L(n), \dots, i_L(n-L_p+2)]^T \quad (4.12)$$

と定義する。

また、文献 [3] の雑音抑圧法において、時刻 $(n+1)$ における状態量は過去の情報を用いずに状態量の変化を表している。それゆえ、 $L_p \times L_p$ 状態遷移行列 Φ と L_p 次元駆動源ベクトル $\boldsymbol{\delta}_L(n+1)$ を

$$\Phi = \begin{bmatrix} 0 & 0 & \dots & \dots & 0 \\ 1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix} \quad (4.13)$$

$$\boldsymbol{\delta}_L(n+1) = [i_L(n+1), 0, \dots, 0]^T \quad (4.14)$$

のように定義する。ここで、式 (4.14) の駆動源ベクトルは時刻 $(n+1)$ における左伴奏信号 $i_L(n+1)$ により構成され、状態ベクトル $\mathbf{x}_L(n+1)$ と相関をもつことに注意されたい [3]。

次に左チャンネルの伴奏信号 $i_L(n)$ に真のヴォーカル信号 $d(n)$ が加わり，左チャンネルの楽曲信号 $y_L(n)$ が生成される過程を表す観測方程式は式 (4.1) より

$$y_L(n) = \mathbf{m}^T \mathbf{x}_L(n) + \epsilon(n) \quad (4.15)$$

となる．ここで L_p 次元観測遷移ベクトル \mathbf{m} と観測雑音 $\epsilon(n)$ を以下

$$\left. \begin{aligned} \mathbf{m} &= \left[\begin{array}{cccc} 1 & 0 & \cdots & 0 \end{array} \right]^T \\ \epsilon(n) &= d(n) \end{aligned} \right\} \quad (4.16)$$

のように定義した．

さて，式 (4.11) と式 (4.15) にて定義した左チャンネルの状態空間モデルに対してカルマンフィルタ理論 [1] を適用することにより，以下のカラオケ生成アルゴリズム (以降，アルゴリズム 1 と称する) が導出される．

$$\left. \begin{aligned} 1. P_L(n+1|n) &= \Phi P_L(n|n) \Phi^T + R_{\delta_L}(n+1) \\ &\quad + Q_L(n+1) + Q_L^T(n+1) \\ 2. \mathbf{k}_L(n+1) &= \{P_L(n+1|n) \mathbf{m}\} \\ &\quad \cdot \{\mathbf{m}^T P_L(n+1|n) \mathbf{m} + \sigma_d^2(n+1)\}^{-1} \\ 3. \widehat{\mathbf{x}}_L(n+1|n) &= \Phi \widehat{\mathbf{x}}_L(n|n) \\ 4. \widehat{\mathbf{x}}_L(n+1|n+1) &= \widehat{\mathbf{x}}_L(n+1|n) \\ &\quad + \mathbf{k}_L(n+1) \cdot \{y_L(n+1) - \mathbf{m}^T \widehat{\mathbf{x}}_L(n+1|n)\} \\ 5. P_L(n+1|n+1) &= \\ &\quad \{I_{L_p} - \mathbf{k}_L(n+1) \mathbf{m}^T\} P_L(n+1|n) \end{aligned} \right\} \quad (4.17)$$

ここで， $P_L(n+1|n)$ は $L_p \times L_p$ の事前推定誤差共分散行列， $P_L(n+1|n+1)$ は $L_p \times L_p$ の事後推定誤差共分散行列， $R_{\delta_L}(n+1)$ は $R_{\delta_L}(n+1) = E[\delta_L(n+1) \delta_L^T(n+1)]$ のように定義される $L_p \times L_p$ の駆動源共分散行列， $Q_L(n+1)$ は $L_p \times L_p$ の相関行列， $\mathbf{k}_L(n+1)$ は L_p 次元カルマンゲインベクトル， $\sigma_d^2(n+1)$ はヴォーカル信号の分散値， $\widehat{\mathbf{x}}_L(n+1|n)$ および $\widehat{\mathbf{x}}_L(n+1|n+1)$ はそれぞれ事前推定ベクトルと事後推定ベクトルである．また I_M は， $M \times M$ の単位行列を表す．

アルゴリズム 1 手順 1. の $Q_L(n+1)$ は，状態ベクトル $\mathbf{x}_L(n+1)$ と駆動源ベクトル $\delta_L(n+1)$ 双方に左チャンネル伴奏信号が含まれがゆえに相関項として残る行列である．本論文では詳細な検討は行わないが， $Q_L(n+1)$ の 1 行 1 列目の要素は 0 となることに注意する．なお，行列 $Q_L(n+1)$ の扱いについては文献 [3] と同様である．

ところで，アルゴリズム 1 手順 3. の事前推定ベクトル $\widehat{\mathbf{x}}_L(n+1|n)$ の第 1 要素は，時刻 $(n+1)$ において出力される左チャンネル伴奏信号の事前推定値 $\widehat{i}_L(n+1|n)$ である．ここで，式

(4.13)に示した状態遷移行列 Φ の構造に着目すれば、 $\hat{i}_L(n+1|n)$ は時刻 n における最適推定値 $\hat{\mathbf{x}}_L(n|n)$ の要素は一切使われずに更新され、結果 0 となる。また、 $\hat{\mathbf{x}}_L(n+1|n)$ の第 1 要素以外は、1 つ前の時刻 n で実行された結果が順次反映される。すなわち、 $\hat{\mathbf{x}}_L(n+1|n)$ の第 2 要素は $\hat{\mathbf{x}}_L(n|n)$ の第 1 要素となり、これは時刻 $(n-1)$ での最適推定値であることから既知量となる。第 3 要素以降も同様である。

従って、時刻毎で更新の対象となるのは、手順 4. における $\hat{\mathbf{x}}_L(n+1|n+1)$ の第 1 要素のみである。これより、 $\hat{\mathbf{x}}_L(n+1|n+1)$ の更新では、 $\hat{\mathbf{x}}_L(n+1|n+1)$ の第 1 要素のみに着目すればよい。これは、手順 4. のカルマンゲインベクトル $\mathbf{k}_L(n+1)$ の第 1 要素のみを考慮すればよいことを意味する。

以上のことを踏まえると、アルゴリズム 1 は次のように書き換えることができる。

まず、事後推定ベクトル $\hat{\mathbf{x}}_L(n+1|n+1)$ における第 1 要素の算出に必要な要素のみを書き下すと

$$\begin{aligned}\hat{\mathbf{x}}_L(n+1|n+1)[1] &= \hat{\mathbf{x}}_L(n+1|n)[1] \\ &\quad + \mathbf{k}_L(n+1)[1] \{y_L(n+1) - \mathbf{m}^T \hat{\mathbf{x}}_L(n+1|n)\} \\ &= 0 + \mathbf{k}_L(n+1)[1] \{y_L(n+1) - 0\} \\ &= \mathbf{k}_L(n+1)[1] y_L(n+1)\end{aligned}\tag{4.18}$$

となる。ここで $\mathbf{a}[i]$ はベクトル \mathbf{a} の i 行目の要素を指定する記号である。

次いでアルゴリズム 1 手順 2. のカルマンゲインベクトル $\mathbf{k}_L(n+1)$ は、式 (4.18) より $\mathbf{k}_L(n+1)$ の第 1 要素のみに着目すると

$$\begin{aligned}\mathbf{k}_L(n+1)[1] &= P_L(n+1|n)[1, 1] \\ &\quad \cdot \{P_L(n+1|n)[1, 1] + \sigma_d^2(n+1)\}^{-1}\end{aligned}\tag{4.19}$$

のように表される。ここで $A[i, j]$ は行列 A の i 行 j 列目の要素を指定する記号である。

最後にアルゴリズム 1 手順 1. の事前推定誤差共分散行列 $P(n+1|n)$ は、式 (4.19) より $P(n+1|n)$ の 1 行 1 列目の要素のみに着目し、かつ $R_{\delta_L}(n+1)$ と $Q_L(n+1)$ の行列の構造を考慮すれば

$$\begin{aligned}P_L(n+1|n)[1, 1] &= P_{mp}[1, 1] + R_{\delta_L}(n+1)[1, 1] \\ &\quad + Q_L(n+1)[1, 1] + Q_L^T(n+1)[1, 1] \\ &= 0 + \sigma_{i_L}^2(n+1) + 0 + 0 \\ &= \sigma_{i_L}^2(n+1)\end{aligned}\tag{4.20}$$

と書き表すことができる。ここで $P_{mp} = \Phi P_L(n|n) \Phi^T$ 、 $\sigma_{i_L}^2(n+1)$ は左チャネル伴奏信号の分散値である。

新たに左チャンネル駆動源分散値を $p_L(n+1)$ ，左チャンネルカルマンゲインを $k_L(n+1)$ ，左チャンネル伴奏信号推定値を $\widehat{x}_L(n+1)$ とおけば，アルゴリズム 1 は式 (4.18) ~ (4.20) より

$$\left. \begin{aligned} 1. p_L(n+1) &= \sigma_{i_L}^2(n+1) \\ 2. k_L(n+1) &= \frac{p_L(n+1)}{p_L(n+1) + \sigma_d^2(n+1)} \\ 3. \widehat{x}_L(n+1) &= k_L(n+1)y_L(n+1) \end{aligned} \right\} \quad (4.21)$$

のように帰着され，これをアルゴリズム 2 と呼ぶこととする．ここでアルゴリズム 2 は，最小分散推定法 [2] に対応している．つまり提案手法は，左チャンネルの伴奏信号 $i_L(n)$ とヴォーカル信号 $d(n)$ の平均値と分散値が既知であれば， $E\left[\{i_L(n) - \widehat{i}_L(n)\}^2\right]$ の誤差分散を最小とする推定値を得ることが可能な手法と言える．なお，ヴォーカル信号および左チャンネル伴奏信号の分散値は，左チャンネル観測信号 $y_L(n)$ と Step1 によって推定されたヴォーカル信号 $\widehat{d}(n)$ を用いて

$$\left. \begin{aligned} \sigma_d^2(n) &= \frac{1}{N} \sum_{i=0}^{N-1} \widehat{d}^2(n-i) \\ \sigma_{i_L}^2(n) &= \frac{1}{N} \sum_{i=0}^{N-1} y_L^2(n-i) - \sigma_d^2(n) \end{aligned} \right\} \quad (4.22)$$

のように計算される．ここで N は分散値を計算するためのサンプル数である．

これまでの議論では左チャンネルの状態空間モデルを対象としていたが，右チャンネルの状態空間モデルにおいても同様な議論ができ，式 (4.11) ~ (4.22) 中の添え字 L を $L \rightarrow R$ とすることにより，右チャンネルカラオケ信号生成アルゴリズムが得られる．

最後に提案手法のカラオケ生成アルゴリズムを表 4.2 に示す．

表 4.2: 提案手法のカラオケ生成アルゴリズム

[Iteration]

for $\ell = 1, 2, \dots, [(f_s \times \text{total length of music})/L_s]_G$

Step1 Extraction of the vocal signal

1. STFT for the stereo music signal
2. Calculation of the discriminant coef.

$$A_e(\ell, k) = \frac{\{|Y_L(\ell, k) - Y_R(\ell, k)|\}^2}{\{|Y_L(\ell, k) + Y_R(\ell, k)|\}^2}$$

$(k_L \leq k \leq k_H)$

3. Extraction of the temporary vocal spectrum

$$|\check{D}(\ell, k)| = \begin{cases} |Y_L(\ell, k)| & (0 \leq A_e(\ell, k) < \alpha) \\ 0 & (\alpha \leq A_e(\ell, k) \leq 1) \\ 0 & (k < k_L, k_H < k) \end{cases}$$

$k = 0, 1, \dots, [(L_F/2)]_G - 1$

4. The vocal signal is calculated by IFFT and overlap add.

$$\widehat{d}(\ell, n) = \check{d}(\ell - 1, n - L_s) + \check{d}(\ell, n)$$

for $n = (\ell - 1)L_s, (\ell - 1)L_s + 1, \dots, \ell L_s$

Step2 Generate of the Karaoke signal

- Calculation of the vocal variance

$$\sigma_d^2(n) = \frac{1}{N} \sum_{i=0}^{N-1} \widehat{d}^2(n - i)$$

- Covariance matrix of the driving source

$$\sigma_{i_L}^2(n) = \frac{1}{N} \sum_{i=0}^{N-1} y_L^2(n - i) - \sigma_d^2(n)$$

$$\sigma_{i_R}^2(n) = \frac{1}{N} \sum_{i=0}^{N-1} y_R^2(n - i) - \sigma_d^2(n)$$

- Extraction algorithm for the Karaoke signal

5. $p_L(n + 1) = \sigma_{i_L}^2(n)$

5. $p_R(n + 1) = \sigma_{i_R}^2(n)$

6. $k_L(n + 1) = \frac{p_L(n + 1)}{p_L(n + 1) + \sigma_d^2(n)}$

6. $k_R(n + 1) = \frac{p_R(n + 1)}{p_R(n + 1) + \sigma_d^2(n)}$

7. $\widehat{x}_L(n + 1) = k_L(n + 1)y_L(n + 1)$

7. $\widehat{x}_R(n + 1) = k_R(n + 1)y_R(n + 1)$

for end

for end

表 4.3: シミュレーションに用いた楽曲信号

Music Title	Artist	Category	Tag
青い春	back number	J-POP	A
Feel the love	浜崎あゆみ	J-POP	B
忘れもの	秋元順子	演歌	C
川千鳥	氷川きよし	演歌	D
Never Say Never	Justin Bieber	洋楽	E
Applause	Ledy GaGa	洋楽	F

4.5 計算機シミュレーション

本章では客観評価と主観評価により，提案手法の音質を評価する．表 4.3 は音質評価に用いたステレオ楽曲信号の一覧である．ここで表 4.3 の各ステレオ楽曲信号の CD に収録されていたインストゥルメンタルトラック (伴奏のみの楽曲) を真のカラオケ信号として以降の各評価に用いる．

ここで表 4.3 におけるサンプリング周波数 f_s を $f_s = 44.1[\text{kHz}]$ ，周波数領域変換のための STFT は窓関数にハニング窓，フレーム長は $L_F = 4096$ ，シフト長は $L_S = 2048$ として実行した．またヴォーカル帯域は $k_L = 200[\text{Hz}]$ ， $k_H = 15,000[\text{Hz}]$ ，従来手法 [29] における暫定カラオケ信号と差信号に対する重み係数をそれぞれ $k_1 = k_2 = 0.5$ とした．

ヴォーカルスペクトル抽出のために用いる閾値 α は，4.4 章でも述べたとおり一意に決定することは困難であるため，実験的に求めた数値として $\alpha = 0.01$ とした．また，ヴォーカル信号の分散値を算出するためのサンプル数を，楽曲信号やヴォーカル信号が非定常信号であるということを考慮して $N = 3$ のように設定した．

4.5.1 スペクトログラムによるステレオ感の評価

本節では従来手法 [29] と提案手法により生成されたカラオケ信号について，スペクトログラムを用いてステレオ感の評価をした．

表 4.3 は Tag A の楽曲信号に対する 22 ~ 27[sec] 間の，(a) に左チャンネルの真のカラオケスペクトル，(b) に右チャンネルの真のカラオケスペクトルを示している．また (c)，(d) はそれぞれ，従来手法を用いて生成した左カラオケスペクトルと右カラオケスペクトル，(e)，

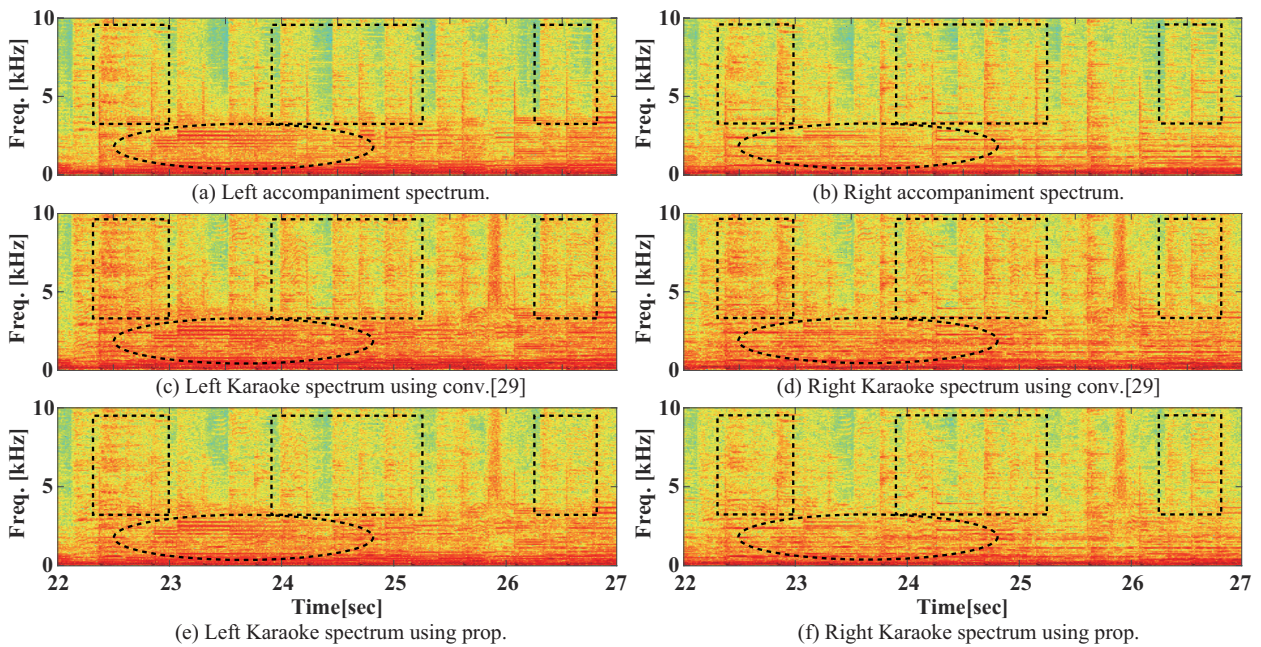


図 4.3: Tag A に対するカラオケ信号のスペクトログラム比較

(f) はそれぞれ提案手法を用いて生成した左カラオケスペクトルと右カラオケスペクトルである。また図の横軸は時刻，縦軸は周波数を表しスペクトルのパワーが強いほど図中の濃淡は濃くなる。

まず，真のカラオケスペクトル (a), (b) と従来手法によって生成されたカラオケスペクトル (c), (d) を比較する。従来手法により生成されたカラオケスペクトルは真のカラオケスペクトルよりも全体的にスペクトルのパワーが強いことが観察できる。また，図中の破線で囲まれた領域を比較するとスペクトルの濃淡分布が異なっている。さらに，従来手法の左右チャンネルのスペクトル (c), (d) を比較すると，破線で囲まれた領域のスペクトルは似た濃淡分布を示している。これは従来手法の式 (4.8) からわかるように，左右チャンネル間の差信号を暫定カラオケ信号に加えたことが原因であると考えられる。

次に真のカラオケスペクトル (a), (b) と提案手法によって生成されたカラオケスペクトル (e), (f) を比較する。破線に囲まれた左右のカラオケスペクトルを比較すると，真のカラオケスペクトルに非常に近い濃淡分布を示していることが確認できる。また左右のカラオケスペクトルは独立した濃淡分布となっている。つまり提案手法によって生成されたカラオケ信号は，真のカラオケ信号に近く左右のスペクトルは異なることから，ステレオ感を有しているといえる。

表 4.4: SNR による従来手法と提案手法のカラオケ信号推定精度評価

Tag	SNR of Conv[29].		SNR of Prop.	
	Left[dB]	Right[dB]	Left[dB]	Right[dB]
A	5.1	4.2	8.2	7.7
B	4.8	4.4	8.0	7.7
C	2.7	3.6	5.4	5.9
D	3.5	3.3	5.9	5.7
E	2.6	1.8	4.8	4.6
F	5.5	5.2	8.0	8.0

4.5.2 SNR を用いた誤差評価

本節では，信号対雑音比 (SNR:Signal to Noise Ratio) を用いて真のカラオケ信号と従来手法 [29]，提案手法により生成されたカラオケ信号との誤差をそれぞれ評価した．評価式は以下のとおりである．

$$\text{SNR} = 10\log_{10} \left\{ \frac{E \left[i_L^2(n) \right]}{E \left[\left\{ i_L(n) - \widehat{i}_L(n) \right\}^2 \right]} \right\} [\text{dB}] \quad (4.23)$$

また右のカラオケ信号 $\widehat{i}_R(n)$ に対しても同様に SNR を算出した．

表 4.4 は表 4.3 の楽曲信号に対して従来，提案手法を用いて生成したカラオケ信号と真のカラオケ信号との誤差を SNR を用いて評価した結果である．SNR 値が大きいほど各手法によって生成されたカラオケ信号が真のカラオケ信号に近いことを意味している．

表 4.4 より，A~F すべての楽曲信号において提案手法の方が従来手法よりも SNR が大きいことが確認できる．しかし提案手法の SNR のみに着目すれば，生成されたカラオケ信号は真のカラオケ信号に必ずしも近いとは言い難い．これは，実験に使用した楽曲信号のヴォーカル信号に対してエフェクトが掛けられており，エフェクト処理は中央定位しないヴォーカル信号成分として残留した結果，SNR が低下したと考えられる．ここで楽曲 C, D, E は他の楽曲よりも SNR が小さいが，これは次のような理由が考えられる．まず楽曲 C, D はヴォーカル信号にエコー処理が掛けられているため，従来・提案手法ともにエコーの部分は抑圧しきれず SNR が小さくなったと考えられる．また楽曲 E はコーラスが左右

に広がって定位しており，このコーラス部分が抑圧しきれず SNR が小さくなったと考えられる。

4.5.3 MOS 評価

本節では従来手法 [29] と提案手法により生成されたカラオケ信号に対して MOS による聞き取り調査を行った結果を示す。調査環境は次のとおりである。

1. 音は PC で再生し，被験者はヘッドホンを装着
2. 最初に真のカラオケ信号を提示
3. 従来手法もしくは提案手法により生成したカラオケ信号を提示
4. もう一度真のカラオケ信号を提示
5. (iii) において提示しなかった手法により生成したカラオケ信号を提示
6. ステレオ楽曲信号 A~F について (ii)~(v) を繰り返す

なお，音量は被験者自身に調節をしてもらった。また，ステレオ楽曲信号 A~F はランダムな順番で提示した。

図 4.4 と図 4.5 は，男性 3 人女性 3 人の計 6 人に対してステレオ楽曲信号 A~F より生成されたカラオケ信号を，図 4.4 は“音質の劣化の有無”，図 4.5 は“ステレオ感”という観点で聞いてもらい，それぞれ表 4.5 と表 4.6 に基づき評価した結果の平均値である。

図 4.4 の音質に関する MOS 評価より，提案手法は従来手法よりもほとんどの楽曲について 4 以上のスコアが得られた。このことから提案手法によって得られたカラオケ信号は，SNR 評価の観点からそれほど精度が良いとは言えないものの，聴感上十分な品質を維持していると言える。また図 4.5 のステレオ感に関する MOS 評価についても，提案手法の方が高いスコアを得ることができた。このことから提案手法は高音質を維持しながら，ステレオ感のあるカラオケ信号生成法であるといえる。

表 4.7 は従来手法と提案手法の MOS 評価値に対して有意差があるのか，t 検定により検定した結果である。有意水準は 5% と定め，有意差が認められた場合は“有”，認められない場合は“無”とそれぞれ表記した。

表 4.5: 音質におけるスコアとカテゴリ

Score	Category
5	劣化が全く認められない
4	劣化が認められるが気にならない
3	劣化がわずかに気になる
2	劣化が気になる
1	劣化が非常に気になる

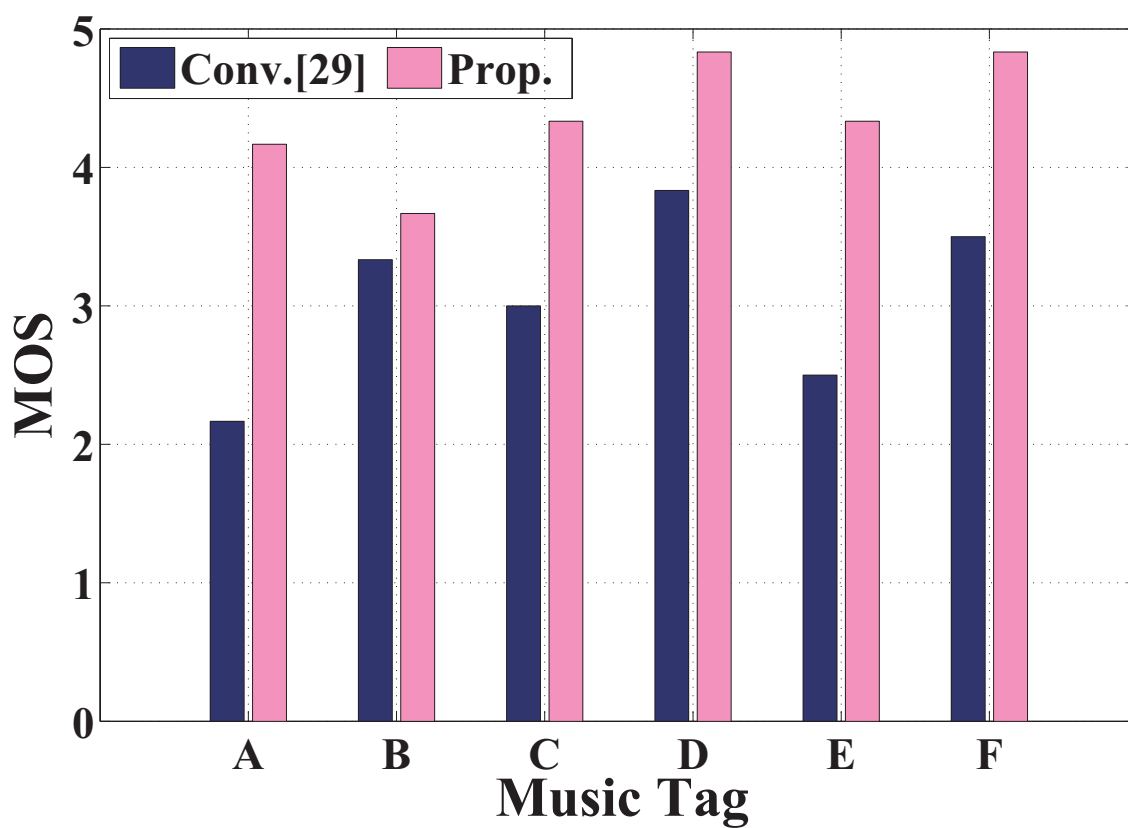


図 4.4: 従来手法と提案手法による音質に関する評価結果

表 4.6: ステレオ感におけるスコアとカテゴリ

Score	Category
5	ステレオ感は完全にある
4	ステレオ感はおおむねある
3	ステレオ感はややある
2	ステレオ感はあまりない
1	ステレオ感はほぼない

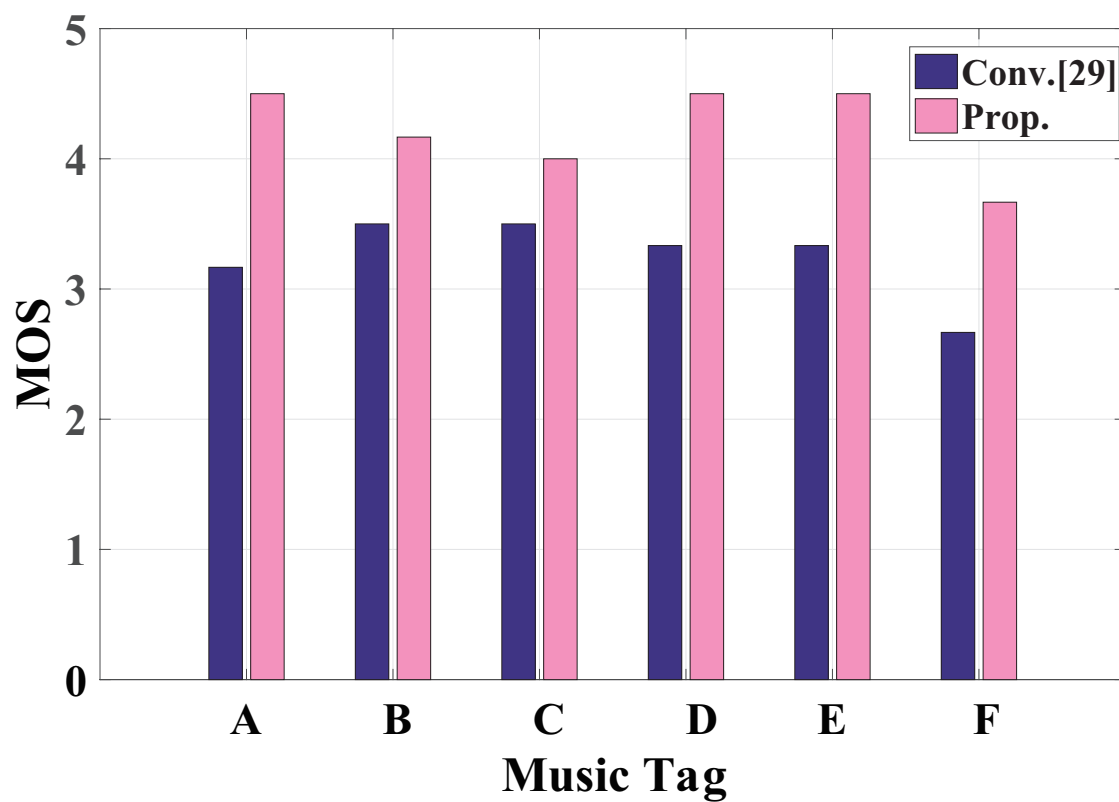


図 4.5: 従来手法と提案手法によるステレオ感に関する評価結果

表 4.7: OS 評価に対する t 検定

評価基準	Music Tag					
	A	B	C	D	E	F
音質劣化の有無に対する有意差	有	無	有	有	有	有
ステレオ感の有意差	有	無	無	有	有	有

表 4.7 より，“音質の劣化の有無”に関しては Tag B 以外，“ステレオ感”に関しては Tag B, C 以外に有意差が認められ，Tag B, C のような楽曲を除き，従来手法よりも提案手法の方が有効であることが確認された。

ここで，Tag B のような pop ミュージックには，楽器の定位を急激に左右に振るエフェクトが掛けられている楽曲が存在する．このような楽曲に対して従来手法のように左右チャンネルに同一成分の信号を加えても，ステレオ感が損なわれたようには知覚されず，提案手法との間に差異が認められなかったと考えられる．また Tag C のような演歌は，残響などのエフェクトが付加されていることが多く，従来もしくは提案手法ではこれらのエフェクトが残ってしまうため，“ステレオ感”に関して有意差が認められなかったと考えられる．

従って，提案手法はエフェクト効果が比較的に少ない楽曲に対して特に有効であると考えられる．

4.6 アプリケーション実装

本章では提案手法の実用性を示すために，演算量評価と携帯電話端末への実装について述べる．

4.6.1 演算量評価

本節では従来手法 [29] と提案手法の演算量評価を示す．表 4.8 は，ステレオ楽曲信号から 1 サンプルのカラオケ信号を生成するために必要な乗算回数を各ステップごとにまとめた結果である．

ここで表中の N_b は，ボーカル帯域 $k_L \sim k_H$ に存在する周波数ビンの数である．また表中の手順 3. は，従来手法では式 (4.6) におけるヴォーカルスペクトルの抑圧操作，提案手法では式 (4.9) における伴奏スペクトルの抑圧操作に必要な乗算回数の最大数である．

表 4.8: 従来手法と提案手法の演算量比較

Procedure	Multiplications	
	Conv.[29]	Prop.
1.	$2L_F + \frac{2L_F \cdot \log_2 L_F}{2}$	$2L_F + \frac{2L_F \cdot \log_2 L_F}{2}$
2.	$5N_b$	$5N_b$
3.	$2N_b$ (max)	N_b (max)
4.	$2L_F + \frac{2L_F \cdot \log_2 L_F}{2}$	$L_F + \frac{L_F \cdot \log_2 L_F}{2}$
Step2	4	8
Total	$7N_b + \frac{4L_F \cdot \log_2 L_F}{2} + 4$	$6N_b + \frac{3L_F \cdot \log_2 L_F}{2} + 8$

表 4.8 より Step1 の手順 3., 4. は, 従来手法よりも提案手法の方が乗算回数が少ない. 提案手法はヴォーカルスペクトルのみを抽出するため, 手順 3. の抑圧操作は左右どちらかのチャンネルに施せばよくまた, 手順 4. の IFFT 操作も 1 チャンネル分のみでよいことが乗算回数減少の要因である.

図 4.6 は表 4.8 を基に, 1 フレーム中のサンプル数 L_F を 1024, 2048, 4096 のように増加させたときの従来手法と提案手法の演算量比較である. なおヴォーカル帯域に存在する周波数ビンの数 N_b は, サンプリング周波数 $f_s = 44.1$ [kHz], ヴォーカル帯域を $k_L = 200$ [Hz], $k_H = 15,000$ [Hz] として算出した.

図 4.6 より, サンプル数 L_F が増加するにつれ従来手法と提案手法の演算量の差が広がっていることが分かる. つまり提案手法は, 周波数分解能を上げてヴォーカル抽出処理をする場合においても少ない演算量でカラオケ信号が生成可能であるといえる.

4.6.2 アプリケーション実装

本節では携帯電話端末用 OS である iOS や Android に対して提案手法を実装したステレオ型自動カラオケアプリケーション“カシレボ! JOYSOUND”について, iOS に実装した場合の操作画面を図 4.7 に示す.

開発した“カシレボ! JOYSOUND”は, 提案手法のヴォーカル抑圧だけでなく, 共同開

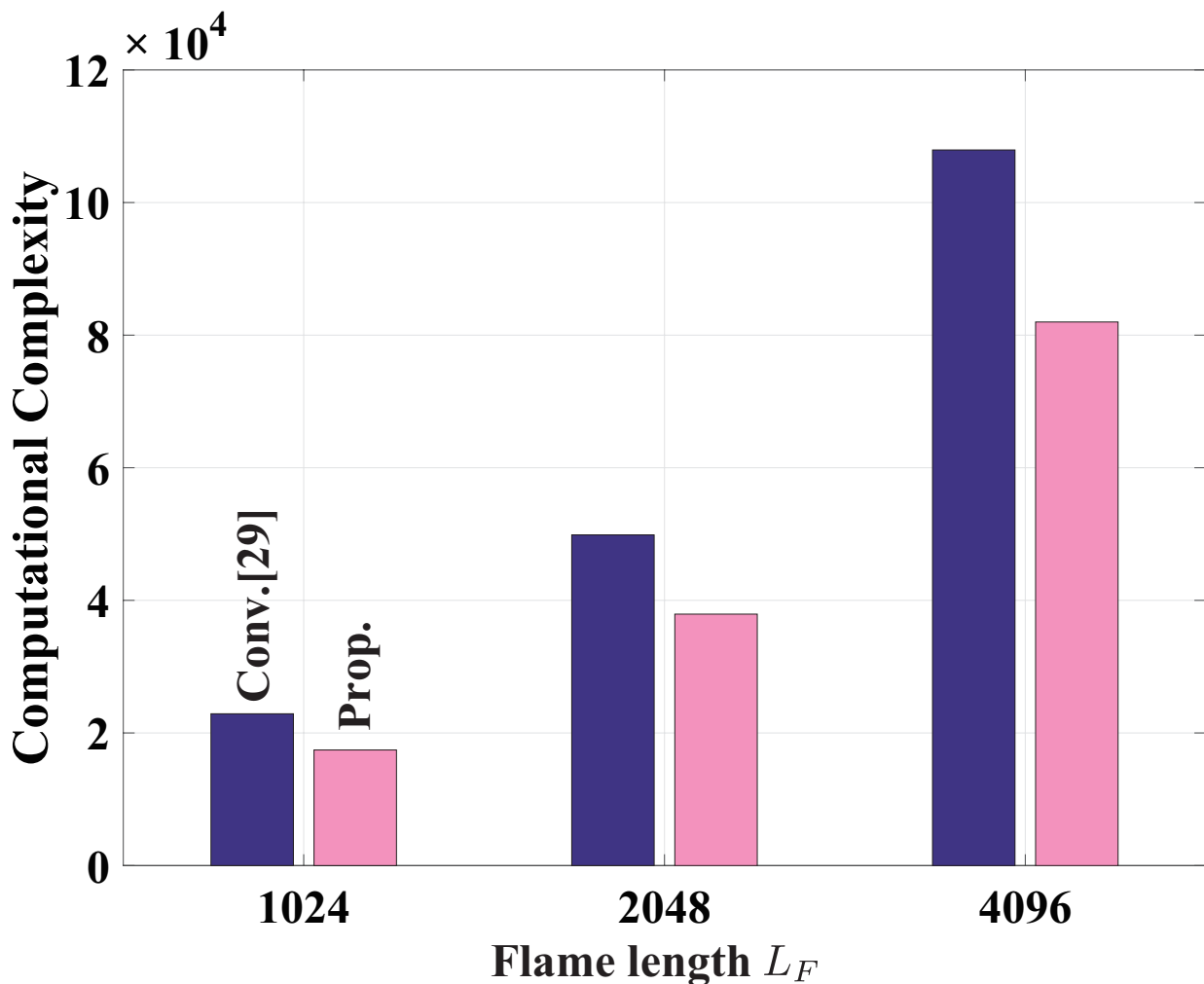


図 4.6: 従来手法と提案手法の演算量比較

発した企業の協力によりカラオケボックスのようにヴォーカルに合わせて歌詞の色が変化する機能を有している。

また著者らは、表 4.2 中の Step2 のカラオケ生成アルゴリズムにおいて、ヴォーカル信号の分散値 $\sigma_d^2(n)$ に適当な変数 ($\beta = 0 \sim 1$) を掛けることにより、ヴォーカルの抑圧具合を調節する機能を実装した。これによりユーザが段階的にカラオケの練習ができるように工夫した。加えてマイクにエコー機能を搭載することでカラオケボックスと同じ雰囲気を楽しめるようにしている。

アプリケーションのダウンロードサイト中のコメントでも好意的な意見が多いことから、“いつでも・どこでも・だれとでも”気軽に楽しめるカラオケアプリケーションを開発したといえる。



図 4.7: スマートフォンに実装したカラオケアプリケーション“カシレボ！JOYSOUND”

4.7 結言

本章では、ステレオ楽曲信号からステレオのカラオケ信号を生成する手法を提案した。提案手法は、ヴォーカル信号の分散値を算出した後に左右のチャンネルに対して独立した状態空間モデルを構成することにより、(i)ステレオ型のカラオケ信号が生成でき、(ii)IFFTの回数が減少することによる演算量の軽減ができる手法である。

提案手法の有効性は、(1)スペクトル評価によりステレオ感を有していること、(2)SNR評価が従来手法よりも改善されていること、(3)MOS評価によりカラオケ信号の劣化が少なくステレオ感を有していること、さらに(4)スマートフォンアプリケーションとして実装し多くの評価を得ていることから明らかにされている。

加えて、雑音抑圧法をカラオケ生成アルゴリズムに応用した場合、最終的に得られた生成アルゴリズムは最小分散推定法に帰着されることも示された。

以上より、提案手法はステレオ楽曲信号に含まれる伴奏信号を忠実に取り出すことが可能なことより、ステレオ感を維持したカラオケ信号生成法といえる。

今後の課題は、ボーカルの中央定位を仮定しないカラオケ信号生成法の検討などが挙げ

られる。

第5章 結論

本論文では、AR モデルを用いない状態量推定手法である雑音抑圧法を無線通信におけるチャンネル推定および、カラオケ音源の自動生成手法に適用した。

雑音抑圧法は、その状態空間モデルの構造、特に状態遷移行列の構造を考慮することで最小分散推定法に帰着されることを示した。

雑音抑圧法を MIMO MC-CDMA 通信方式におけるチャンネルゲイン推定に適用した場合、推定アルゴリズムは移動体の移動速度に依存しないことを計算機シミュレーションにより示した。また、雑音抑圧法を MIMO MC-CDMA 方式に応用した場合においてもチャンネルゲイン推定手法は最小分散推定法に帰着されることを示した。

さらに、カラオケ音源の自動生成アルゴリズムとして雑音抑圧法を応用した。雑音抑圧法により生成されたカラオケ信号はステレオ感が維持されたまま音質の劣化もほとんどないことが、計算機シミュレーションおよび聞き取り調査より示した。加えて、スマートフォンへの実装評価より実用的なカラオケ音源生成アルゴリズムであることを示した。

謝辞

本論文は筆者が東京理科大学工学研究科経営工学専攻博士課程における研究成果をまとめたものである。

本大学の博士課程入学から今日に至るまで研究や論文の執筆に関してご指導頂いた，東京理科大学工学部情報工学科 古川 利博教授には格別の感謝の意を申し上げます。進路に関して親身に御相談に乗って頂いたことも重ねて感謝の意を申し上げます。

また筆者の学部時代から今日に至るまで御指導頂いた，諏訪東京理科大学工学部コンピュータメディア工学科 田邊 造准教授にも深謝の意を申し上げます。企業との共同研究や電子工作を通して学生・一般の方々と触れあう機会を与えて頂き，研究者としてまた教える立場として今後必要な経験を積むことができた。

博士論文の審査にあたり，東京理科大学工学部情報工学科 浜田 知久馬教授，八嶋 弘幸教授，太原 育夫教授ならびに谷口 行信教授には異なる分野からの視点で御指摘，ご指導頂き感謝の意を申し上げます。

最後に高橋 智博助教を始め古川研究室，田邊研究室の院生・学部生の皆様にも研究を進める上で様々な御助言，ご討論頂き感謝の意を申し上げます。

参考文献

- [1] S.Haykin, 適応フィルタ理論 (科学技術出版, 東京, 2001), pp.340–363.
- [2] ルーエンバーガ著, 増淵 正美, 嘉納 秀明 共訳, 関数解析による最適理論 (コロナ社, 東京, 1973 年), pp.88–91.
- [3] 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, pp.818–829, Mar. 2008.
- [4] W. Kim, and H. Ko, “Noise variance estimation for Kalman filtering of noise speech,” *IEICE Trans. Inf. & Syst.*, vol.E84–D, No.1, pp.155–160, Jan. 2001.
- [5] 山内 雪路, スペクトラム拡散通信 高性能デジタル通信方式に向けて (東京電機大学, 東京, 2001 年).
- [6] 伊丹 誠, わかりやすい OFDM 技術 (オーム社, 東京, 2005 年).
- [7] 大鐘 武雄, 小川 恭孝, わかりやすい MIMO システム技術 (オーム社, 東京, 2008 年 6 月)
- [8] B.S. Chen, J.F. Liao, “Adaptive MC–CDMA Multiple Channel Estimation and Tracking Over Time–Varying Multipath Fading Channels,” *IEEE trans. Wireless Comm.*, Vol.6, No.6 Mar. 2007.
- [9] Z. WANG, C. HE, “A Channel Estimation Scheme for MIMO–MC–CDMA Downlink in High Mobility Environments,” *IEICE Trans. Commun.*, Vol.E88–B, No.3, Mar. 2005.
- [10] 吉田 晴人, 前田 文明, 笹森 文仁, 高畑 文雄, “高速フェージング環境下における OFDM 伝送を対象としたフェージング補償方式,” *信学論 (B)*, vol.J85-B, no.11, pp.1874–1883, Nov. 2002.

- [11] 神谷 幸宏, MATLAB によるデジタル無線通信技術 (コロナ社, 東京, 2008 年 12 月)
- [12] 松江 英明, 佐藤 明雄, 渡辺 和ニ, 守倉 正博, 高速ワイヤレスアクセス技術 (電子情報通信学会, 東京, 2004 年 3 月)
- [13] 高畑 文雄, 前原 文明, 笹森 文仁, デジタル無線通信入門 (培風館, 東京, 2007 年 10 月)
- [14] B.S.Chen, C.Y.Yang, and W.J.Liao, “Robust Fast Time-Varying Multipath Fading Channel Estimation and Equalization for MIMO-OFDM Systems via a Fuzzy Method,” *IEEE Trans. Vehicular Technology.*, vol.61, no.4, pp.1599–1609, May 2012.
- [15] P. Gupta and D.K.Mehra, “Kalman filter-based channel and ICI suppression for high-mobility OFDM systems,” in *Int’l.J.Commun.Syst.*, vol.21, pp.1075-1090, June 2008.
- [16] C.Min, N.Chang, J.Cha, and J.Kang, “MIMO-OFDM Down link Channel Prediction for IEEE802.16e Systems Using Kalman Filter,” in *Proc. of the IEEE Wireless Communications and NetWorking Conf.(WCNC2007)*, pp.943-947, Kowloon, March 2007.
- [17] B.S. Chen, J.F. Liao, “Adaptive MC-CDMA Multiple Channel Estimation and Tracking Over Time-Varying Multipath Fading Channels,” in *IEEETrans. Wireless Comm.*, vol.6, no.6, pp.2328–2337, Jun. 2007.
- [18] J.Yue, K.J.Kim, J.Gibson, and R.Iltis, “Channel estimation and data detection for MIMO-OFDM systems,” in *Proc. Global Tele commun.Conf.*, vol.2, pp.581–585, Dec. 2003.
- [19] H.Nooralizadeh and S.S.Moghaddam, “A novel shifted type of SLS estimator for estimation of Rician flat fading MIMO channels,” in *SignalProcessing*, vol.90, no.6, pp.1886–1893, 2010.
- [20] M.Biguesh and A.B.Gershman, “Training-Based MIMO Channel Estimation : A Study of Estimator Trade offs and Optimal Training Signals,” in *SignalProcessing*, vol.54, no.3, pp.884–893, March, 2006.
- [21] H.Hijazi, E.P.Simon, M.Lienard, and L.Ros, “Channel Estimation for MIMO-OFDM Systems in Fast Time-Varying Environments”, in *Int’l Symposium on Communications, Control and Signal Processing. (ISCCSP2010)*, pp.1-6, Nov. 2010.

- [22] Y.H.Kho, D.P.Taylor, “MIMO Channel Estimation and Tracking Based on Polynomial Prediction With Application to Equalization”, *IEEETrans.VehicularTechnology.*, vol.57, no.3, pp.1585–1595, May 2008.
- [23] K.E.Baddour and N.C.Beaulieu, “Auto regressive Modeling for Fading Channel Simulation”, in *IEEEtrans.WirelessComm.*, vol.4, no.4, pp.1659–1662, July 2005.
- [24] K.P.Bagadi, S.Das, “MIMO-OFDM Channel Estimation usign Pilot Carries”, in *Int’l Journal of Computer Applications*, vol.2, no.3, pp.81–88, May 2010.
- [25] 一般社団法人全国カラオケ事業者協会, “カラオケ歴史年表,” http://www.karaoke.or.jp/03nenpyo/03_02.php (参照 2015–10–05).
- [26] 谷井 章夫, 後藤 真孝, 片寄晴弘, “ミックスダウンデザインテンプレートの利用に関する提案,” 情報処理学会研究報告, 2003–MUS–51, pp.19–24, Aug. 2003.
- [27] 牧野 昭二, 荒木 章子, 向井 良, 澤田 宏, “成分分析に基づくブラインド音源分離,” 信学技報, EA2003–45, pp.17–24, Jun 2003.
- [28] N.R.Bhalani, J.Singh, M.Tiwari, “Karaoke Machine Implementation and validation Using Out of Phase Stereo Method,” 2012 Int’l.Conf. on Communication, Infomation & Computing Techology (ICCICT), 3pages 2012.
- [29] 株式会社神戸製鋼所. 池田陽平. 音響信号処理装置, 音響信号処理プログラム, 音響信号処理方法. 特開 2008-72600. Mar. 2008.
- [30] M.Kato, A.Sugiyama, M.Serizawa, “重み付き雑音推定と MMSESTSA 法に基づく高音質雑音抑圧,” 信学論 (A), vol.J87–A, no.7, pp.851–860 July 2004.
- [31] Y.Uemura, Y.Takahashi, H.Saruwatari, K.Kondo, “Automatic Optimization Scheme of Spectral Subtraction based on Musical Noise Assessment via Higher–Order Statistics,” 11th Int’l Workshop on Acoustic Echo and Noise Control (IWAENC2008), 4pages Sep. 2008.

目次

2.1	従来手法の状態空間モデル	6
3.1	MIMO MC-CDMA 通信方式 送信機ブロック図	17
3.2	MIMO MC-CDMA 通信方式 チャンネルモデル	18
3.3	MIMO MC-CDMA システム 受信機ブロック図	19
3.4	従来手法の NMSE 特性	25
3.5	n ブロック目の状態遷移モデル	27
3.6	$(n + 1)$ ブロック目の状態遷移モデル	28
3.7	送受信アンテナ数の変化に対する演算量評価	34
3.8	移動速度 v の変化に対する従来手法と提案手法のチャンネルゲイン推定精度比較	35
3.9	遅延波数 L_h に対する従来手法と提案手法のチャンネルゲイン推定精度比較 . .	36
3.10	送信アンテナ数 P の変化に対する従来手法と提案手法のチャンネルゲイン推 定精度比較	37
3.11	従来手法と提案手法のシンボルエラーレート評価	38
4.1	ステレオ楽曲信号の構成	42
4.2	提案手法 Step1 の流れ	48
4.3	Tag A に対するカラオケ信号のスペクトログラム比較	55
4.4	従来手法と提案手法による音質に関する評価結果	58
4.5	従来手法と提案手法によるステレオ感に関する評価結果	59
4.6	従来手法と提案手法の演算量比較	62
4.7	スマートフォンに実装したカラオケアプリケーション“カシレボ!JOYSOUND”	63

表 目 次

2.1	ARモデルとカルマンフィルタを用いた雑音抑圧アルゴリズム	8
2.2	有色駆動源カルマンフィルタを用いた雑音抑圧法のアルゴリズム	10
2.3	最小分散状態推定アルゴリズム	12
3.1	従来手法のカルマンフィルタアルゴリズム	23
3.2	シミュレーション諸元 [9]	24
3.3	提案手法のチャンネルゲイン推定アルゴリズム	29
3.4	雑音抑圧法を用いたチャンネルゲイン推定アルゴリズム	32
3.5	従来手法と提案手法の演算量比較	33
4.1	従来手法と提案手法のカラオケ信号生成手順	46
4.2	提案手法のカラオケ生成アルゴリズム	53
4.3	シミュレーションに用いた楽曲信号	54
4.4	SNRによる従来手法と提案手法のカラオケ信号推定精度評価	56
4.5	音質におけるスコアとカテゴリ	58
4.6	ステレオ感におけるスコアとカテゴリ	59
4.7	OS評価に対するt検定	60
4.8	従来手法と提案手法の演算量比較	61