

# 指向性パターンに着目したマイクロフォンアレイにおける雑音除去 Noise Reduction by Microphone Array Using Directional Patterns

鈴木 優一  
Yuichi Suzuki

石田 義久†  
Yoshihisa Ishida

## 1. まえがき

近年、携帯電話や音声認識システムなどの利用が急増し、マイクロフォンで観測された信号から周囲の雑音を除去する技術はますます重要になってきている。また、状況によっては発話者が一人であることが多く、周囲の雑音環境下で発話者の音声のみを選択的に聴取する必要もある。

本論文では指向性パターンに着目したマイクロフォンアレイにおける雑音除去手法を提案する。マイクロフォンアレイ技術はアレイアンテナ技術に基づくものである[1]。本論文ではその中でも収束が早い方向拘束付電力最小化規範（以下 DCMP 規範）を用いている。電波は波長が短く帯域が狭いのに対し、音声信号は波長が長く帯域が広いためそのまま音声信号に適用することができない。そこで、音声信号に適用させるために、観測された信号を複数の狭帯域信号に分割して帯域毎に処理を行う必要がある。本論文で提案する手法では2素子マイクロフォンアレイで得られた指向性パターンより不要信号の到来方向を推定することで、帯域毎の計算量を削減し、全体での実行時間を短縮させることが可能である。

## 2. アレイアンテナ

$K$  素子より構成される直線状に配置された等間隔リニアアレイを考える。ある時刻  $t$  にアンテナに垂直な方向より角度  $\theta$  の方向から信号が到来したとき、第  $k$  素子により観測される信号  $x_k(t)$  は次式で与えられる。

$$x_k(t) = x_1(t - \tau_k) \quad (k=1, \dots, K) \quad (1)$$

$$\tau_k = \frac{d_k \sin \theta}{c}$$

式中  $c$  は音速、 $d_k (k=1, \dots, K)$  は基準点から第  $k$  素子までの間隔、 $\tau_k (k=1, \dots, K)$  は第  $k$  素子での遅延である。このとき観測信号  $x_k(t)$  がアレイ開口長  $|d_K - d_1|$  に対し十分狭帯域、すなわち

$$2\pi \Delta f \frac{|d_K - d_1|}{c} \ll 1 \quad (2)$$

ならば、 $x_k(t)$  は次式で書き表すことができる。

$$x_k(t) \approx x_1(t) e^{-j2\pi f_c \tau_k} \quad (k=1, \dots, K) \quad (3)$$

上式中  $\Delta f$  は観測信号の帯域幅、 $f_c$  は  $x_k(t)$  の中心周波数である。アレイ出力信号  $y(t)$  は各素子での観測信号に振幅と位相を調整する重みを掛け合わせ、それらを加算することにより次式で与えられる。なお式中の添え字  $*$ ,  $^H$ ,  $^T$  はそれぞれ複素共役、複素共役転置、転置を表す。

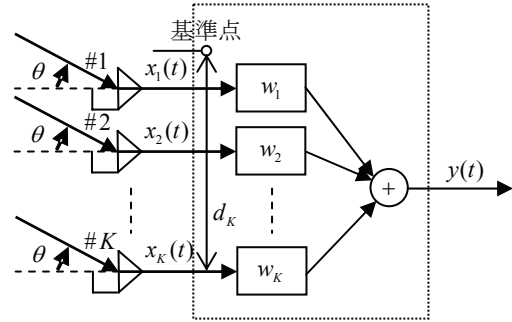


図1:  $K$  素子等間隔リニアアレイの構成図

$$y(t) = \sum_{k=1}^K w_k^* x_k(t) = \mathbf{w}^H \mathbf{x}(t) = \mathbf{x}^T(t) \mathbf{w}^* \quad (4)$$

なお上式の入力ベクトル  $\mathbf{x}(t)$ , 重みベクトル  $\mathbf{w}$  は次式で定義される。

$$\mathbf{x}(t) = [x_1(t) \ \dots \ x_K(t)]^T \quad (5)$$

$$\mathbf{w}(t) = [w_1(t) \ \dots \ w_K(t)]^T \quad (6)$$

図1に  $K$  素子等間隔リニアアレイの構成図を示す。

### 2.1 DCMP 規範

DCMP 規範は到来方向に基づく線形拘束下で出力電力を最小化する規範である。拘束数を  $N$  とすると DCMP 規範は以下のように定義される[2]。ただし拘束数  $N$  は素子数  $K$  よりも少なくなければならない ( $N < K$ )。

目的関数

$$\min_{\mathbf{w}} \left( P_{out} = \frac{1}{2} \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w} \right) \quad (7)$$

制約

$$\mathbf{C}^T \mathbf{w}^* = \mathbf{h} \quad (8)$$

式中  $\mathbf{C}$  は拘束行列、 $\mathbf{h}$  は拘束応答ベクトルでそれぞれ

$$\mathbf{C} = \begin{bmatrix} c_{11} & \dots & c_{1N} \\ \vdots & \ddots & \vdots \\ c_{K1} & \dots & c_{KN} \end{bmatrix} \quad (9)$$

$$\mathbf{h} = [h_1 \ \dots \ h_N]^T \quad (10)$$

$$c_{kn} = e^{-j2\pi f_c \tau_{kn}} \quad (k=1, \dots, K; n=1, \dots, N) \quad (11)$$

† 明治大学理工学部 (問合せ: 生体情報制御研究室)

e-mail: ishida@isc.meiji.ac.jp

$$\tau_{kn} = \frac{f_c(k-1)d \sin \phi_n}{c} \quad (12)$$

(k=1, \dots, K; n=1, \dots, N)

で与えられる。ただし、 $d$  は素子間隔である。 $\phi_n$  (n=1, \dots, N) は拘束方向を示し、一般的に  $\phi_n$  が所望信号の到来方向の場合  $h_n$  を 1 とし、 $\phi_n$  が不要信号の到来方向の場合  $h_n$  を 0 とする。最適な重みベクトル  $\mathbf{w}_{opt}$  は目的関数(7)、及び制約式(8)より次式で与えられる。

$$\mathbf{w}_{opt} = \mathbf{R}_{xx}^{-1} \mathbf{C} (\mathbf{C}^H \mathbf{R}_{xx}^{-1} \mathbf{C})^{-1} \mathbf{h}^* \quad (13)$$

式(13)中の  $\mathbf{R}_{xx}$  は観測信号  $\mathbf{x}(t)$  の自己相関行列で、次式により得られる。

$$\mathbf{R}_{xx} = E \left[ \mathbf{x}(t) \mathbf{x}(t)^H \right] \quad (14)$$

$E[\cdot]$  は期待値演算を表す。期待値演算は  $\mathbf{x}(t)$  のすべてのデータを必要とするため、自己相関行列  $\mathbf{R}_{xx}$  を適応的に求めることは不可能である。そこで  $\mathbf{R}_{xx}$  の推定が必要となる。この推定には最適化アルゴリズムの中でも収束の早い Sample Matrix Inversion (以下 SMI) アルゴリズムを用いる。SMI アルゴリズムは、時刻  $t$  での自己相関行列の逆行列  $\hat{\mathbf{R}}_{xx}^{-1}(t)$  を 1 サンプルごとに更新していくもので、次式で与えられる[3]。

$$\hat{\mathbf{R}}_{xx}^{-1}(t) = \frac{1}{\beta} \hat{\mathbf{R}}_{xx}^{-1}(t-1) + \frac{(1-\beta) \hat{\mathbf{R}}_{xx}^{-1}(t-1) \mathbf{x}(t) \mathbf{x}(t)^H \hat{\mathbf{R}}_{xx}^{-1}(t-1)}{\beta^2 + \beta(1-\beta) \mathbf{x}(t)^H \hat{\mathbf{R}}_{xx}^{-1}(t-1) \mathbf{x}(t)} \quad (15)$$

式中  $\beta$  ( $0 < \beta < 1$ ) は忘却係数で、過去のサンプルの重要性を決定するものである。得られた  $\hat{\mathbf{R}}_{xx}^{-1}(t)$  を用い、推定される時刻  $t$  での重みベクトル  $\hat{\mathbf{w}}(t)$  は

$$\hat{\mathbf{w}}(t) = \hat{\mathbf{R}}_{xx}^{-1}(t) \mathbf{C} (\mathbf{C}^H \hat{\mathbf{R}}_{xx}^{-1}(t) \mathbf{C})^{-1} \mathbf{h}^* \quad (16)$$

で与えられる。

## 2.2 音声信号への適応

1章で述べたように、マイクロフォンアレイ技術はアレイアンテナ技術に基づくものである。そのためそのまま音声信号に適応することができない。これは音声信号が電波に比べ帯域が広く、波長が長いためである。そのため観測された信号を複数の狭帯域信号に分割して帯域毎に処理を行う必要がある。この帯域分割には短区間フーリエ変換 (Short-Time Fourier Transform, 以下 STFT) [4]を用いる。STFT を用いることである周波数での時間信号を狭帯域信号とすることができ、窓関数を  $h(t)$  とおくと、短区間フーリエ変換及び逆変換は次式で定義される。

$$X_k(f, t) = \sum_p x_k(p) h(p-t) e^{-j2\pi f t} \quad (k=1, \dots, K) \quad (17)$$

$$x_k(t) = \frac{1}{h(t)} \sum_p \sum_f X_k(f, p) e^{j2\pi f(t-p)} \quad (k=1, \dots, K) \quad (18)$$

このとき DCMP 規範の制約式(8)は、

$$\mathbf{C}^T(f) \mathbf{w}^*(f, t) = \mathbf{h} \quad (19)$$

と書くことができる。ここで  $\mathbf{C}(f)$ 、 $\mathbf{w}(f, t)$  は周波数  $f$  での拘束行列及び重みベクトルである[5]。また重みベクトルの推定値は以下のように書き直すことができる。

$$\hat{\mathbf{w}}(f, t) = \hat{\mathbf{R}}_{xx}^{-1}(f, t) \mathbf{C}(f) (\mathbf{C}(f)^H \hat{\mathbf{R}}_{xx}^{-1}(f, t) \mathbf{C}(f))^{-1} \mathbf{h}^* \quad (20)$$

## 3. 提案法

本論文で提案する手法は 2 素子マイクロフォンアレイで得られた指向性パターンから不要信号の到来方向を推定し、帯域毎の計算量を削減することで全体での実行時間を短縮させるものである。そのため以下では素子数  $K$  を 2 に限定する。また所望信号の到来方向を  $\phi$  とする。

### 3.1 不要信号の到来方向推定

周波数  $f$  におけるマイクロフォンアレイの指向性パターンは以下の式で与えられる[6]。

$$D(\theta, f, t) = \left| \mathbf{w}^H(f, t) \mathbf{V}(\theta, f) \right|^2 = \left| \mathbf{V}^T(\theta, f) \mathbf{w}^*(f, t) \right|^2 \quad (21)$$

式中  $\mathbf{V}(\theta, f)$  はアレイ応答ベクトルであり次式で定義される。

$$\mathbf{V}(\theta, f) = [v_1(\theta, f) \quad v_2(\theta, f)]^T$$

$$v_k(\theta, f) = \exp\left(-j2\pi \frac{f(k-1)d \sin \theta}{c}\right) \quad (k=1, 2) \quad (22)$$

さらに、

$$w_1(f, t) = a_1(f, t) e^{j b_1(f, t)} \quad 0 \leq b_1 \leq 2\pi$$

$$w_2(f, t) = a_2(f, t) e^{j b_2(f, t)} \quad 0 \leq b_2 \leq 2\pi \quad (23)$$

とおくと式(21)は

$$D(\theta, f, t) = a_1(f, t)^2 + a_2(f, t)^2 + 2a_1(f, t)a_2(f, t) \cos\left(2\pi f \frac{d \sin \theta}{c} + b_1(f, t) - b_2(f, t)\right) \quad (24)$$

と書き換えることができる。DCMP アルゴリズムは拘束下で出力電力を最小化するように重みベクトルを決定するため、不要信号の到来方向で  $D(\theta, f, t)$  は最小となる。したがって上式中、

$$\cos\left(2\pi f \frac{d \sin \theta}{c} + b_1(f, t) - b_2(f, t)\right) = -1 \quad (25)$$

となるような角度  $\sigma_m$  ( $m=1, \dots, M$ ) を決定する。この指向性パターンを最小とする角度ベクトルは式(23)、式(24)により以下のように表すことができる。

$$\sigma(f, t) = [\sigma_1(f, t) \quad \dots \quad \sigma_M(f, t)]$$

$$\sigma_m(f, t) = \sin^{-1}\left(\frac{c}{2\pi f d} \left\{ (2n+1)\pi - b_1(f, t) + b_2(f, t) \right\}\right) \quad (26)$$

(m=1, \dots, M; n=0, \pm 1, \pm 2, \dots)

$M$  は最小点の数である。

角度  $\sigma_m$  が複数存在するのは、ナイキスト周波数の回り込みによるもので、マイクロフォンにより観測された信号の波長がアレイ開口長  $d_k$  よりも短い場合、素子間に 1 波長以上の波が存在することになり、重み  $w_k$  は位相を一意に決定することができないからである。したがって周波数  $f$  が高くなるにつれ波長が短くなるため、 $M$  は増加する。

上記のように観測された信号の波長がアレイ開口長  $d_k$  よりも長い場合には不要信号の到来方向は一意に推定できるが、それ以外の場合複数の最小点があるため到来方向を一意に推定することができない。また実際の到来方向と推定される到来方向の推定誤差  $\Delta\sigma_m(f, t)$  は周波数  $f$  が高くなるにつれ小さくなる。これは以下により証明することができる。

式(26)より  $b_1(f, t)$ ,  $b_2(f, t)$  の推定誤差をそれぞれ  $\Delta b_1(f, t)$ ,  $\Delta b_2(f, t)$  とすると角度の推定誤差は

$$\Delta\sigma_m(f, t) = \sin^{-1} \left( \frac{c}{2fd} \left( \frac{\Delta b_2(f, t) - \Delta b_1(f, t)}{\pi} + 2n + 1 \right) \right) \quad (27)$$

$(n = 0, \pm 1, \pm 2, \dots)$

となる。正確な位相を  $\bar{b}_1(f, t)$ ,  $\bar{b}_2(f, t)$  とおくと位相の推定誤差  $\Delta b_1(f, t)$ ,  $\Delta b_2(f, t)$  は及び誤差範囲は式(23)より、以下のよう定義される。

$$\begin{aligned} \Delta b_1(f, t) &= b_1(f, t) - \bar{b}_1(f, t) \quad (-2\pi \leq \Delta b_1(f, t) \leq 2\pi) \\ \Delta b_2(f, t) &= b_2(f, t) - \bar{b}_2(f, t) \quad (-2\pi \leq \Delta b_2(f, t) \leq 2\pi) \end{aligned} \quad (28)$$

よって式(27)は以下のように書き直すことができる。

$$\begin{aligned} \Delta\sigma_m(f, t) &= \sin^{-1} \left( \frac{c}{2fd} (A + 2n + 1) \right) \\ A &= \frac{\Delta b_2(f, t) - \Delta b_1(f, t)}{\pi} \quad (29) \\ &(-4 \leq A \leq 4; n = 0, \pm 1, \pm 2, \dots) \end{aligned}$$

このように  $A$  は一定な範囲内に収まり、しかも  $f$  が分母にあるため、到来方向の推定誤差  $\Delta\sigma_m(f, t)$  は周波数  $f$  が高くなるにつれて小さくなる。

したがって高精度で不要信号の到来方向  $\hat{\phi}(t)$  を推定するためには、まず指向性パターン<sup>1</sup>の最小点が一意に決定できる帯域内の任意の周波数  $f_L$  で推定された角度  $\sigma(f_L, t)$  を求める。次に、これよりも高い任意の周波数  $f_H$  によって推定された式(26)における角度ベクトル  $\sigma(f_H, t)$  を求め、その中で最も  $\sigma(f_L, t)$  の近傍の角度を推定角度  $\hat{\phi}(t)$  とする。

$$\hat{\phi}(t) = \arg \min_{\sigma} [\sigma(f_L, t) - \sigma_m(f_H, t)] \quad (30)$$

$(m = 1, \dots, M_H)$

なお式中  $M_H$  は周波数  $f_H$  での最小点の数である。

### 3.2 帯域毎の計算量削減

図2に本論文で提案するマイクロフォンアレイのブロック線図を示す。

各帯域では既に与えられた所望信号の到来方向  $\phi$  に関する拘束条件に、角度推定によって得られた不要信号の到来方向  $\hat{\phi}(t)$  に関する拘束条件を追加することで拘束行列、拘束応答値ベクトルは以下のように書き直すことができる。

$$\mathbf{C}(f, t) = [c_{kn}(f, t)] \quad (k = 1, 2; n = 1, 2) \quad (31)$$

$$\begin{aligned} \mathbf{h} &= [h_1 \quad h_2]^T \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{aligned} \quad (32)$$

$$\begin{aligned} \phi_1(t) &= \phi \\ \phi_2(t) &= \hat{\phi}(t) \end{aligned} \quad (33)$$

$$\begin{aligned} c_{kn}(f, t) &= \exp(-j2\pi f \tau_{kn}(f, t)) \\ \tau_{kn}(f, t) &= \frac{f(k-1)d \sin \phi_n(t)}{c} \quad (34) \\ &(k = 1, 2; n = 1, 2) \end{aligned}$$

ここで  $\mathbf{C}(f, t)$  は正方行列なので重みベクトル  $\mathbf{w}(f, t)$  は式(20)の最適化アルゴリズムを用いることなく次式による行列演算で求めることができる。

$$\mathbf{G} = \mathbf{C}^{-1}(f, t) \quad (35)$$

$$\mathbf{w}(f, t) = \{\mathbf{G}^T \mathbf{h}\}^* \quad (36)$$

さらに  $\mathbf{G}$  を次式のように分割する。

$$\begin{aligned} \mathbf{G} &= \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} \\ \mathbf{g}_1 &= [g_{11} \quad g_{12}] \\ \mathbf{g}_2 &= [g_{21} \quad g_{22}] \end{aligned} \quad (37)$$

このとき式(37)は簡略化でき、重みベクトル  $\mathbf{w}(f, t)$  は次式により求めることができる。

$$\begin{aligned} \mathbf{w}(f, t) &= \left\{ \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix}^T [h_1 \quad h_2]^T \right\}^* \\ &= \mathbf{g}^H \end{aligned} \quad (38)$$

## 4. シミュレーション

シミュレーションは素子数を2、素子間隔を0.05[m]とし、所望信号が0[度]、不要信号が59.59[度]から到来していると仮定して、サンプリング周波数8[kHz]で計算機上で作成したものを5秒間用いて行った。音声は、男性・女性各一名ずつの会話音声で、所望信号を男声として行った。また音速は330[m/s]とし、所望信号に関する拘束条件として  $\phi = 0$  [度]、SMI アルゴリズムの忘却係数を  $\beta = 0.9994$  とした。STFT パラメータとしてフレーム長を32, 64, 128, 256、フレーム周期をそれぞれ8, 16, 32, 64とした。

既存法として、観測信号をSTFTを用いて帯域分割し、帯域ごとにDMCPアルゴリズムを適用して重みベクトルを推定し、出力信号を得る手法を用いた。5秒間の音声の処理時間、および出力Signal to Noise Ratio (以下出力SNR)とで提案法との比較を行った。図3に分割する帯域数毎の出力SNR、表1に分割する帯域数毎の実行時間を示す。なお既存法のSTFTパラメータは提案法で用いたものと同じである。

また出力SNRは所望信号  $s(t)$ 、出力信号  $y(t)$  から次式により定義される。

$$\begin{aligned} \text{SNR}(t) &= 10 \log_{10} \sum_{\tau} |s(t+\tau)|^2 \\ &\quad - 10 \log_{10} \sum_{\tau} |s(t+\tau) - y(t+\tau)|^2 \end{aligned} \quad (39)$$

なお、シミュレーションに用いた音声信号および不要信号の到来方向を変更しても、ほぼ同様なシミュレーション結果が得られた。紙面の都合により詳細はいずれ稿を改めて述べたい。

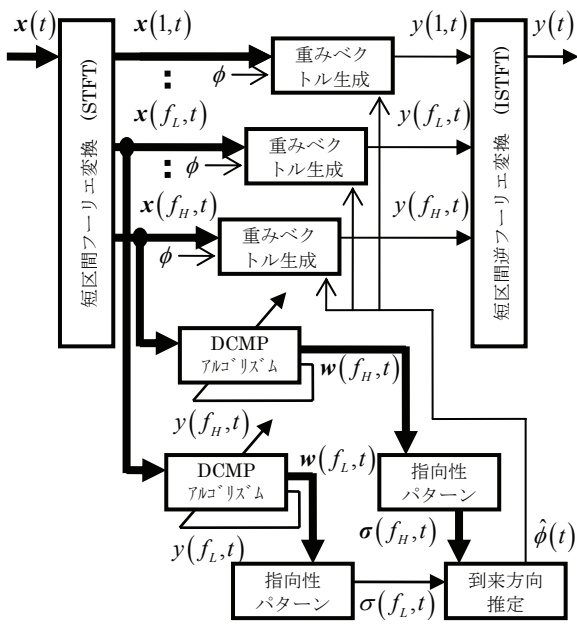


図2：マイクロフォンアレイのブロック線図

表1：帯域数毎の実行時間の比較

帯域数	既存法	提案法
32	11.7512 [秒]	7.2274 [秒]
64	11.6807 [秒]	5.0118 [秒]
128	11.7779 [秒]	3.9414 [秒]
256	11.9825 [秒]	3.4916 [秒]

5. 結論

図3から分かるようにすべての帯域数で出力 SNR の改善がみられた。これは不要信号の到来方向推定により、予め与えられた所望信号の到来方向に関する拘束条件に、不要信号の到来方向に関する拘束条件を追加できたためと考えられる。またすべての帯域において実行時間が短縮された。これは既存法ではすべての帯域で計算量の多い最適化アルゴリズムを用いて重みベクトルを推定しているのに対し、提案法では角度推定に用いた2帯域でしか最適化アルゴリズムを使用しておらず、各帯域では逆行列演算により重みベクトルを推定しているためと考えられる。

謝 辞

本研究に対して有益なご助言をいただいた査読者の方々に謝意を表します。また、日頃ご指導いただいている村上隆啓氏（日本学術振興会特別研究員・東京農工大）に感謝します。

文 献

- [1] 菊間 信良：アダプティブアンテナ技術，オーム社。
- [2] K. Takao, M. Fujita, and T. Nishi, *An Adaptive Antenna Array under Directional Constraint*, IEEE Trans. Antennas & Propag., vol. Ap-24, no. 5, pp. 662-669, 1976.
- [3] 菊間 信良：アレーアンテナによる適応信号処理，科学技術出版。
- [4] Cohen. L, *Time-Frequency Analysis*, Prentice-Hall PTR, Englewood Cliffs, New Jersey.
- [5] T. Murakami, K. Kurihara, and Y. Ishida, *Directionally Constrained Minimization of Power Algorithm for Speech Signals*, INTERSPEECH 2005, pp. 2333-2336, Lisbon, Portugal.
- [6] Y. Suzuki, Y. Okuma, T. Murakami, and Y. Ishida, *Reduction of Computing Time for Microphone Array by Directional Patterns*, International Conference of Signal Processing 06', Guilin.

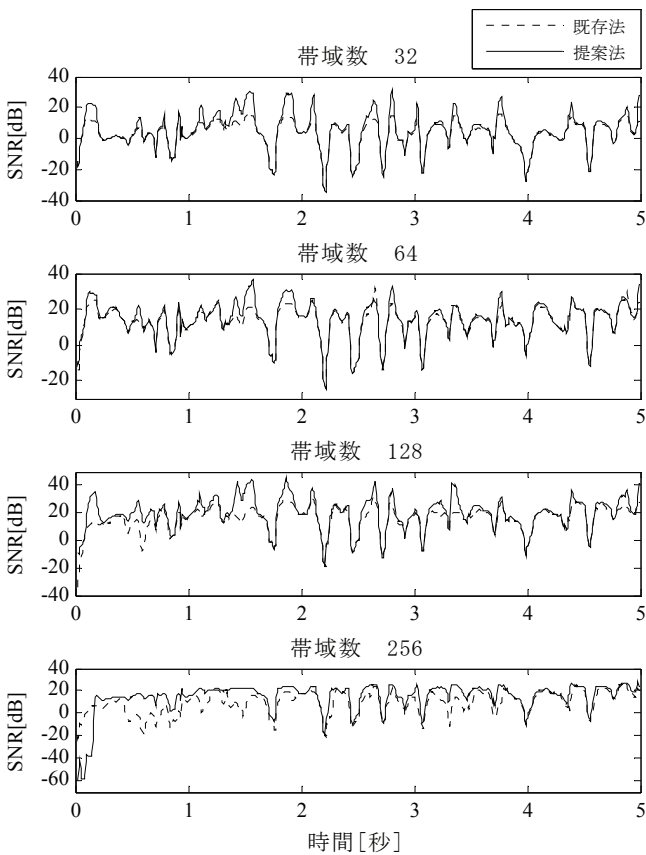


図3：帯域数毎の出力 SNR の比較