

テーパ窓を用いた特異値分解と極推定による 1 次元空間上の複数音源の位置推定*

金井 浩**

(東北大学工学部電気工学科)

永田仁史・安倍正人・牧野正三・城戸健一***

(東北大学応用情報学研究センター)

(平成元年 1 月 31 日受理)

内容概観 多数センサによる 1 次元空間上の複数音源の位置推定法が幾つか提案されている。しかし、特に低い S/N 下では、音源位置を高精度に推定することが難しい。最小 2 乗法によって雑音パワーを最小化することによって得られた under-determined 正規方程式を解いても、実音源の周囲に依然大きなリップルがあり、音源位置を高精度に決定できない。そこで本論文では、筆者が先に開発したテーパ窓を用いた特異値分解と、スペクトル極推定法によるマルチパルス駆動系列推定法に基づいて、複数音源位置を高精度に推定するための手法を提案する。最後にシミュレーション実験により理論的有効性を確かめる。本手法の基礎的原理は、多方面への応用が期待できる。

1. ま え が き

本論文では、1 次元空間上の複数の音源位置の高精度推定のための新しい手法を提案する。従来、音源位置推定に関しては、相関法に基づく方法^{1)~3)}が提案されている。しかし、この手法は音源波形を推定しないために、音源信号間に相関がある場合には複数の音源位置を同時に推定できない⁴⁾。また、複数の音源がある場合に、多数のセンサを用いて系統的に音源位置を推定することが困難である。

そこで、クロススペクトル法を用いた手法が提案された⁴⁾。しかし、複数の音源がある場合には、1 対の音源とセンサ間の伝達関数の単純な補償では、他の音源の放射音の影響を大きく受ける。更にセンサの出力信号が雑音を含む場合には、雑音成分を最小化する定式化が必要となる。

各音源位置がすべて既知であれば、通常の最小 2 乗法を用いて付加雑音のパワーの最小化によって導出された優決定系 (over-determined) 正規方程式から、各音源

の出力音を推定できる。しかし、音源位置推定問題においては、音源位置が未知であるから、仮想的に設定する音源⁴⁾の数をセンサ数よりも十分多く設定する必要がある。従って、雑音パワーの最小化によって導出される正規方程式は under-determined となり、各仮想音源の放射音を推定することが難しくなる。

そこで我々は、雑音下での複数音源の位置推定法を提案した⁵⁾。最小 2 乗法で導出される under-determined 正規方程式に対して、特異値分解を利用して最小ノルム解を得る。一般にセンサ数が仮想音源数に対して非常に少ないため、特異値分解では非常に低次での切り捨てが必要となる。この低次切り捨てによって、音源位置はリップルによってぼかされ、高精度の音源位置推定が困難となる。そこで、特異値分解の低次項切り捨てにおいて、リップルをある程度抑制するためのテーパ窓を用いた特異値分解による推定法を提案した^{5)~7)}。本論文では、全極型伝達系を駆動するマルチパルス系列の高精度推定法⁷⁾を基に、上記の手法で得られた各仮想音源位置のスペクトル系列の逆フーリエ変換系列に対して極推定法を適用し、リップルの影響を受けずに高精度に音源位置を決定するための手法を提案する。最後にシミュレーション実験によって、本手法の推定精度を示す。

2. 原 理

Fig. 1 に示すように、1 次元空間 l 上に M_0 個の実音源 (図中×印) を想定し、各々の音源から放射される

* A method to locate multiple sound sources in one-dimensional space using tapered singular-value-decomposition and pole estimation.

** Hiroshi Kanai (Department of Electrical Engineering, Faculty of Engineering, Tohoku University, Sendai, 980)

*** Yoshifumi Nagata, Masato Abe, Shozo Makino and Ken'iti Kido (Research Center for Applied Information Sciences, Tohoku University, Sendai, 980)

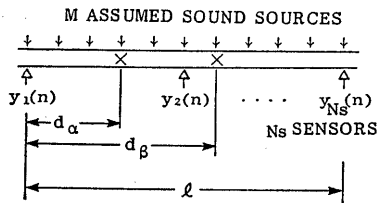


Fig. 1 The relation between the M_0 actual sound sources, the L sensors, and the M assumed sound sources in the one-dimensional space. They are represented by 'x', '▲', and '↓', respectively. In the figure, d_k , ($k=1, 2, \dots, M_0$) denotes the distance between the k -th actual sound source and the left edge of the one-dimensional space.

音が同じ 1 次元空間 l 上を伝搬して N_s 個のセンサによって受信されるものとする。 M_0 個の実音源の位置を決定するために、同じ 1 次元空間上に M 個の仮想音源を想定し、そのおのおのからの放射音 $x_i(n)$, ($i=1, 2, \dots, M$) を推定する。もし i 番目の仮想点上に実音源がある場合には、推定された放射音 $x_i(n)$ は、実音源から放射された音と理想的には一致する。従って、 i 番目の仮想音源の放射音の推定値は大きいパワーを持ち、音源として決定できる。 M 次元ベクトル $\mathbf{x}(n)$

$$\mathbf{x}(n) = [x_1(n), x_2(n), \dots, x_M(n)]^T \quad (0 \leq n < N)$$

によって、 M 個の仮想音源から放射される N 点の長さを持つ信号を表すものとする。ここで、 T は、行列・ベクトルの転置を表す。また、 N_s 個のセンサによって受信された長さ N の信号を N_s 次元ベクトル $\mathbf{y}(n)$ で表す。

$$\mathbf{y}(n) = [y_1(n), y_2(n), \dots, y_{N_s}(n)]^T \quad (0 \leq n < N)$$

i 番目の仮想音源点から j 番目のセンサへの伝達関数のインパルス応答を $h_{ij}(n)$ で表すとき、 $\mathbf{x}(n)$ と $\mathbf{y}(n)$ の関係は次式で表される。

$$\mathbf{y}(n) = \begin{bmatrix} \sum_{i=1}^{M} \sum_{k=0}^{N-1} x_i(n-k) h_{i1}(k) \\ \sum_{i=1}^{M} \sum_{k=0}^{N-1} x_i(n-k) h_{i2}(k) \\ \dots\dots\dots \\ \sum_{i=1}^{M} \sum_{k=1}^{N-1} x_i(n-k) h_{i,N_s}(k) \end{bmatrix} + \mathbf{n}(n) \quad (1)$$

$$(0 \leq n < N)$$

ここで、 $\mathbf{n}(n)$ を各センサで受信した信号に混入した、互いに無相関な白色雑音系列を表す N_s 次元ベクトルとする。

$$\mathbf{n}(n) = [n_1(n), n_2(n), \dots, n_{N_s}(n)]^T \quad (0 \leq n < N)$$

更に $h_{ij}(n)$ の N 点離散的フーリエ変換 $H_{ij}(k)$

$$H_{ij}(k) = \sum_{n=0}^{N-1} h_{ij}(n) \exp(-j2\pi kn/N) \quad (2)$$

からなる次の $N_s \times M$ マルチチャネル伝達行列 $H(k)^{8)}$ を考える。

$$H(k) = \begin{bmatrix} H_{11}(k) & H_{12}(k) & \dots & H_{1M}(k) \\ H_{21}(k) & H_{22}(k) & \dots & H_{2M}(k) \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ H_{N_s1}(k) & H_{N_s2}(k) & \dots & H_{N_sM}(k) \end{bmatrix}$$

$$(k=0, 1, 2, \dots, N-1)$$

また、 $X_i(k)$, $N_j(k)$, $Y_j(k)$ により、おのおの $x_i(n)$, $n_j(n)$, $y_j(n)$ のフーリエ変換の k 番目の周波数成分を表す。更に、 $\mathbf{X}(k)$, $\mathbf{N}(k)$, $\mathbf{Y}(k)$ によってこれらのスペクトルからなるベクトルを表そう。

$$\mathbf{X}(k) = [X_1(k), X_2(k), \dots, X_M(k)]^T$$

$$\mathbf{N}(k) = [N_1(k), N_2(k), \dots, N_{N_s}(k)]^T$$

$$\mathbf{Y}(k) = [Y_1(k), Y_2(k), \dots, Y_{N_s}(k)]^T \quad (0 \leq k < N)$$

以上の行列とベクトルを用いることによって、(1)式は、周波数領域では次のように表すことができる。

$$\mathbf{Y}(k) = H(k) \cdot \mathbf{X}(k) + \mathbf{N}(k), \quad (k=0, 1, \dots, N-1)$$

$$(3)$$

ただし、(1)式の線状畳み込みは因果律を満たすが、(3)式は巡回畳み込みであるため、現実の観測信号の近似的な表現となる。

この式を用いて、 N_s 個のセンサによる観測信号のスペクトル $\mathbf{Y}(k)$ から、各仮想音源上のスペクトル $\mathbf{X}(k)$ を推定することによって、実音源の位置とその信号を推定する。そのためには、 M 個の仮想音源の位置をあらかじめ設定する必要がある。例えば、1次元空間上の仮想音源点とセンサ間のすべての組み合わせに関する伝達関数を、あらかじめ設定できる場合には、その $H_{ij}(k)$ を用いればよい。しかし、本論文では i 番目の仮想点と j 番目のセンサ間の距離 r_{ij} によって決まる距離減衰と位相回転を用いて⁴⁾、次のように伝達関数 $H_{ij}(k)$ を設定する。

$$H_{ij}(k) = \exp(-\alpha r_{ij}) \exp(-j2\pi r_{ij} f_s k / cN) \quad (4)$$

α, f_s, c は、おのおの減衰定数、A/D 変換の標準化周波数、音速を表す。

N_s チャネルの観測信号 $\mathbf{y}(n)$ のスペクトル $\mathbf{Y}(k)$ から、 M_0 個の実音源の位置を決定するためには、 M 個の各仮想音源点上の中から、大きなパワーを放射している M_0 個の音源位置を求める必要がある。本論文では、従来の手法^{4), 5)} とその問題点に関して述べた後、新しい手法を提案する。

2.1 一対のセンサ・仮想音源間の伝達関数の補償に基づく音源位置推定⁹⁾

j 番目のセンサの出力信号のスペクトル $Y_j(k)$ に対し、 i 番目の仮想音源から j 番目のセンサへの伝達関数 $H_{ij}(k)$ の逆特性 $H_{ij}(k)^{-1}$ を掛ける。各センサの出力から同様に求めた N_s 個のスペクトル $\{H_{ij}(k)^{-1} \cdot Y_j(k)\}$, ($j=1, 2, \dots, N_s$) を加算することによって、 i 番目の仮想点のスペクトル $X_i(k)$ が次のように推定できる。

$$\hat{\mathbf{X}}(k) = H^{-1}(k) \cdot \mathbf{Y}(k) \quad (5)$$

ここで、 $H^-(k)$ は伝達関数の逆特性を表す $M \times N_s$ 行列である。

$$H^-(k) = \begin{bmatrix} H_{11}(k)^{-1} & H_{12}(k)^{-1} & \cdots & H_{1N_s}(k)^{-1} \\ H_{21}(k)^{-1} & H_{22}(k)^{-1} & \cdots & H_{2N_s}(k)^{-1} \\ \cdots & \cdots & \cdots & \cdots \\ H_{M1}(k)^{-1} & H_{M2}(k)^{-1} & \cdots & H_{MN_s}(k)^{-1} \end{bmatrix}$$

($k=0, 1, 2, \dots, N-1$)

もし、 i 番目の仮想点上に実音源がある場合には、仮想点の推定されたスペクトル $\hat{X}_i(k)$ 、すなわち $\hat{X}(k)$ の i 番目の要素は大きなパワーを持つから、この手法によって実音源の位置を同定できる⁴⁾。従って、各周波数 k ごとに推定した i 番目の仮想音源のスペクトル $\hat{X}_i(k)$ のパワー $W_i(k)$ を適当な周波数帯域で加算して、 i 番目の仮想音源からの放射パワー P_i が推定できる。

$$P_i = \sum_k W_i(k) = \sum_k \hat{X}_i(k) \hat{X}_i^*(k) \quad (1 \leq i \leq M) \quad (6)$$

各仮想音源ごとに求められた放射パワー P_i から、実音源の位置が決定できる。

しかし、(3)式の $Y(k)$ を(5)式の $\hat{X}(k)$ へ代入することによって、 $\hat{X}(k)$ は次のように表される。

$$\hat{X}(k) = H^-(k) \cdot H(k) \cdot X(k) + H^-(k) \cdot N(k) \quad (7)$$

ここで、第1項の $H^-(k) \cdot H(k)$ は単位行列とはならない。従って、 i 番目の仮想点の放射音の推定値 $\hat{X}_i(k)$ は、 i 番目以外の仮想音源の放射音の影響を受ける。更に、低い S/N 下においては、雑音によって発生する第2項も無視できなくなる。

2.2 $H(k)$ の特異値分解・テーパ窓を用いた音源位置の推定法⁹⁾

(3)式の雑音項 $N(k)$ のパワーを最小化することによって、 $X(k)$ の最小2乗解 $\hat{X}(k)$ が次式で与えられる。

$$\hat{X}(k) = (H(k)^H \cdot H(k))^{-1} H(k)^H \cdot Y(k) \quad (8)$$

ここで、 H は複素転置を表す。一般に最小2乗法を適用する場合、観測ベクトル $Y(k)$ の次元数 N_s は、未知ベクトル $X(k)$ の次元数 M よりも大きく、正規方程式は over-determined となる。しかし、音源位置推定問題においては、仮想音源の数 M は、センサの数 N_s よりも大きい。従って、 $M \times M$ 行列 $(H(k)^H \cdot H(k))$ の階数は、 N_s 以下となり、(8)式により直接 $\hat{X}(k)$ を推定することができない。

そこで、このように under-determined な場合において、我々は論文⁹⁾の中で、特異値分解を用いて $X(k)$ を推定する次の手法を開発した。 k 番目の周波数に関するマルチチャネル伝達特性を表す $N_s \times M$ 行列 $H(k)$ は、次のように特異値分解できる⁹⁾。

$$H(k) = Q_1(k) \Sigma(k) Q_2(k)^H \quad (9)$$

ここで $Q_1(k)$ と $Q_2(k)$ は、おのおの $H(k)H(k)^H$, $H(k)^H H(k)$ の固有ベクトルからなる $N_s \times N_s$, $M \times M$ ユニタリ行

列、 $\Sigma(k)$ は $H(k)$ の特異値 $\{\sigma_i(k)\}$ からなる次の $N_s \times M$ 対角行列である。

$$\Sigma(k) = \begin{bmatrix} \sigma_1(k) & \cdots & 0 & 0 \cdots 0 \\ 0 & \sigma_2(k) & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \sigma_{N_s}(k) & 0 \cdots 0 \end{bmatrix}$$

= $\text{diag} [\sigma_1(k), \sigma_2(k), \dots, \sigma_{N_s}(k)]$
($k=0, 1, 2, \dots, N-1$)

ただし、 $\sigma_1(k) \geq \sigma_2(k) \geq \dots \geq \sigma_{N_s}(k)$ 。

$H(k)$ の階数を R ($R \leq N_s < M$) と置き、 $(R+1)$ 次以降の微小な特異値を切り捨てて得られる、 $N_s \times M$ 対角行列 $\Sigma_R(k)$ を次式で表す。

$$\Sigma_R(k) = \text{diag} [\sigma_1(k), \sigma_2(k), \dots, \sigma_R(k), 0, 0, \dots, 0]$$

実際には、切り捨て次数 R は、例えば次式で定義される $\rho(R; \{\sigma_i\})$ を閾値 T_0 と比較することによって決定できる¹⁰⁾。

$$\rho(R; \{\sigma_i\}) = \sqrt{\sum_{i=1}^R |\sigma_i|^2 / \sum_{i=1}^{N_s} |\sigma_i|^2} \leftrightarrow T_0 \quad (10)$$

$\Sigma_R(k)$ の $M \times N_s$ 逆行列 $\Sigma_R(k)^{-1}$ を次式で定義する。

$$\Sigma_R(k)^{-1} = \text{diag} [\sigma_1(k)^{-1}, \sigma_2(k)^{-1}, \dots, \sigma_R(k)^{-1}, 0, 0, \dots, 0]$$

以上から、(8)式の $\hat{X}(k)$ は、次式で推定できる。

$$\hat{X}(k) = Q_2(k) \Sigma_R(k)^{-1} Q_1(k)^H Y(k) \quad (11)$$

ところで(3)式の $Y(k)$ と(9)式の $H(k)$ を(11)式の $\hat{X}(k)$ へ代入する。

$$\hat{X}(k) = Q_2(k) \Delta_R(k) Q_2(k)^H X(k) + Q_2(k) \Sigma_R(k)^{-1} Q_1(k)^H N(k) \quad (11')$$

ここで、 $\Delta_R(k) = \Sigma_R(k)^{-1} \Sigma(k)^H$ は、 $M \times M$ 行列のうち、左上 $R \times R$ 行列が単位行列で、残りの要素がすべて零からなる行列である。

$$\Delta_R(k) = \text{diag} [\delta_1(k), \delta_2(k), \dots, \delta_M(k)]$$

$$\delta_i(k) = \begin{cases} 1 & (1 \leq i \leq R) \\ 0 & (R < i \leq M) \end{cases}$$

(11')式の第1項 $Q_2(k) \Delta_R(k) Q_2(k)^H$ は、 $Q_1(k)$ と $Q_2(k)$ がユニタリ行列であるから、低次通過フィルタ⁵⁾⁻⁷⁾に相当する。また、第2項 $Q_2(k) \Sigma_R(k)^{-1} Q_1(k)^H$ は、雑音の影響を抑制する高次切り捨てフィルタを示す。従って、(11)式で $\hat{X}(k)$ を推定することにより、雑音成分を抑圧して、音源からの放射音のスペクトルを推定できる。

しかし、(11')式第1項での次数 R による急峻な切り捨ては、低域通過フィルタリングによって発生するギブスの現象¹¹⁾と同様、実音源の周りにリップルを生じる^{6), 7)}。そこで、リップルを抑圧し実音源の位置を明確化するために、上記の特異値分解において、次のようにテーパ窓を用いる。

逆行列 $\Sigma(k)^{-1}$ の対角成分を R 次で切り捨てるための行列 $\Delta_R(k)$ の対角成分である矩形型系列 $\{\delta_i\}$ に、テーパ窓関数 $\{w_i\}$ を畳み込み、得られる系列を $\{\lambda_i\}$

とする。具体的には、 δ_i , ($i=1, 2, \dots, M$) とその対称な系列 δ_i , ($i=M, M-1, \dots, 1$) からなる $2M$ 点の系列 ($\delta_1, \delta_2, \dots, \delta_{M-1}, \delta_M, \delta_M, \delta_{M-1}, \dots, \delta_1$) に対して、テーパ窓関数 $\{w_i\}$ を巡回畳み込みし、得られた $2M$ 点の系列 ($\lambda_1, \lambda_2, \dots, \lambda_{M-1}, \lambda_M, \lambda_{M+1}, \dots, \lambda_{2M}$) の最初の M 点を改めて $\{\lambda_i\}$ とおく。次に、この $\{\lambda_i\}$ を (i, i) 要素を持つ $M \times M$ 対角行列 $A(k)$ を求める。(11)式において用いた行列 $\Delta_R(k) = \Sigma_R(k)^{-1} \Sigma(k)$ の代わりに、 $A(k) \Sigma(k)^{-1}$ を用いることにより、 $\hat{X}(k)$ は次式で推定できる。

$$\hat{X}(k) = Q_2(k) A(k) \Sigma(k)^{-1} Q_1(k)^H Y(k) \quad (12)$$

ここで、

$$A(k) = \text{diag} [\lambda_1(k), \lambda_2(k), \dots, \lambda_M(k)]$$

例えばテーパ窓として、窓の長さ $(2N_w + 1)$ 点の次の Hamming 窓^{11), 12)}を用いる⁵⁾⁻⁷⁾。

$$w_i = 0.54 + 0.46 \cos(2\pi i / 2N_w)$$

$$(-N_w \leq i \leq N_w) \quad (13)$$

(12)式を用いて、 k 番目の周波数の M 個の仮想音源の放射音のスペクトルの推定値 $\hat{X}(k)$ が得られ、(6)式から i 番目の仮想点から放射される音のパワー P_i が計算できる。

このようにマルチチャネル伝達関数の一般化逆特性とテーパ窓を用いることによって、実音源位置とその放射音の信号波形を推定できる。

2.3 極推定法を用いた音源位置推定の高精度化

2.2 節の手法において用いたテーパ窓によっても、リップルを完全に抑制できない^{6), 7)}。そこで本節では、線形予測分析による極推定法¹³⁾⁻¹⁶⁾を拡張し、音源位置推定をより高精度に推定するための方法を提案する。

Fig. 1 のように、1次元空間 l 上に M_0 個の実音源があり、 p 番目の実音源の l 上左端からの距離を d_p 、その実音源の放射音スペクトルの k 番目の周波数成分を $D_p(k)$ と置く。1次元空間上に十分細かい幅 Δd で等間隔に設定した M 個の仮想音源のスペクトル $X_1(k)$, $X_2(k)$, \dots , $X_M(k)$ は、これを周波数間隔 Δd で標準化したスペクトルと考えると、 d_1, d_2, \dots, d_{M_0} に対応する周波数でおおの $D_1(k), D_2(k), \dots, D_{M_0}(k)$ の成分を持つ。この線スペクトルパターンに M 点の離散的逆フーリエ変換を施し、得られる系列を $x_q(k)$ と置く。

$$x_q(k) = \sum_{m=1}^M X_m(k) \cdot \exp(j2\pi(m-1)(q-1)/M) \quad (14)$$

$$(q=1, 2, \dots, M) \quad (k=1, 2, \dots, N/2-1)$$

更に $X_m(k)$ に関する線スペクトル構造を考えると、次式が得られる。

$$x_q(k) = \sum_{p=1}^{M_0} D_p(k) \cdot \exp(j2\pi d_p \cdot (q-1)/(M \cdot \Delta d)) \quad (15)$$

この式は、振幅 $D_p(k)$ 、周波数 $d_p/(M \cdot \Delta d)$ 、($p=1, 2, \dots, M_0$) からなる、実音源数 M_0 個の複素正弦波の和と

なっている。この系列 $x_1(k), x_2(k), \dots, x_M(k)$ に対し、線形予測分析¹³⁾⁻¹⁶⁾による極推定を行うことによって、各実音源位置を2.2節の手法に比べて更に高精度に推定できる。以上に基づく音源位置推定法を以下で具体的に述べる。

まず、(12)式で推定された各仮想音源上のスペクトル $\hat{X}_m(k)$, ($m=1, 2, \dots, M$) に対して、(14)式の逆フーリエ変換を施して、系列 $x_q(k)$, ($q=1, 2, \dots, M$) を求める。次にこの系列に対して、 L 次の複素係数 b_j , ($j=1, 2, \dots, L$) を用いた線形予測分析を行う。精度を上げるために、前向き予測誤差と後向き予測誤差¹³⁾を導入し、更に対象とする周波数帯域全体にわたって、同時に線形予測誤差パワーの最小化を行う。そのため $x_q(k)$, ($q=1, 2, \dots, M$) から構成される L 次元ベクトル $\mathbf{s}_f(q; k)$ と $\mathbf{s}_b(q; k)$ を定義する。

$$\mathbf{s}_f(q; k) = [x_{q-1}(k), x_{q-2}(k), \dots, x_{q-L}(k)]^T$$

$$\mathbf{s}_b(q; k) = [x_{q-L+1}^*(k), x_{q-L+2}^*(k), \dots, x_q^*(k)]^T$$

$$(q=L+1, L+2, \dots, M)$$

ここで、 $*$ は複素共役を表す。 $\mathbf{s}_f(q; k)$ と $\mathbf{s}_b(q; k)$ を用いて、周波数 k に関する前向き予測誤差 $e_f(q; k)$ と後向き予測誤差 $e_b(q; k)$ を定義する。

$$e_f(q; k) = x_q(k) + \sum_{j=1}^L b_j \cdot x_{q-j}(k) = x_q(k) + \mathbf{s}_f(q; k)^T \mathbf{b} \quad (16a)$$

$$e_b(q; k) = x_{q-L}^*(k) + \sum_{j=1}^L b_j \cdot x_{q-L+j}^*(k) = x_{q-L}^*(k) + \mathbf{s}_b(q; k)^T \mathbf{b} \quad (16b)$$

ここで、

$$\mathbf{b} = [b_1, b_2, \dots, b_L]^T$$

次に周波数 k に関する、前向き・後向き線形予測誤差パワーの和 $\varepsilon(k)$ を次式で定義する。

$$\varepsilon(k) = \sum_{q=L+1}^M \{ |e_f(q; k)|^2 + |e_b(q; k)|^2 \} \quad (17)$$

この誤差パワーを、対象とする周波数帯域で加算して得られる誤差パワー E_T を最小化するように線形予測係数ベクトル \mathbf{b} を求める。今、ナイキスト周波数までの直流成分以外の全周波数帯域 ($0 < k \leq N/2-1$) を利用できるものとすれば、 E_T は次式で定義できる。

$$E_T = \sum_{k=1}^{N/2-1} \varepsilon(k) \rightarrow \text{最小化} \quad (18)$$

次に E_T を行列とベクトルで表す。まず、 $2(M-L) \times L$ 行列 $S(k)$ を次式で定義する。

$$S(k) = [\mathbf{s}_f(L+1; k), \mathbf{s}_f(L+2; k), \dots, \mathbf{s}_f(M; k), \mathbf{s}_b(L+1; k), \mathbf{s}_b(L+2; k), \dots, \mathbf{s}_b(M; k)]^T$$

$$(k=1, 2, \dots, N/2-1)$$

更に、 $2(N/2-1)(M-L) \times L$ 行列 S_T を、 $S(k)$ を用いて次式で定義する。

$$S_T = [S(1)^T, S(2)^T, \dots, S(N/2-1)^T]^T$$

一方, $2(M-L)$ 次元ベクトル $S(k)$ を次式で表す。

$$S(k) = [x_{L+1}(k), x_{L+2}(k), \dots, x_M(k), \\ x_1(k)^*, x_2(k)^*, \dots, x_{M-L}(k)^*]^T \\ (k=1, 2, \dots, N/2-1)$$

$S(k)$ を用いて $2(N/2-1)(M-L)$ 次元ベクトル S_T を次式で定義する。

$$S_T = [S(1)^T, S(2)^T, \dots, S(N/2-1)^T]^T$$

行列 S_T とベクトル $S_T \mathbf{b}$ を用いて, (18)式の全周波数帯域における誤差パワー E_T は, 次のように表すことができる。

$$E_T = (S_T + S_T \mathbf{b})^H (S_T + S_T \mathbf{b}) \rightarrow \text{最小化} \quad (18')$$

この誤差パワーを \mathbf{b} に関して最小化することによって得られる正規方程式は次のように表される。

$$(S_T^H S_T) \mathbf{b} = -S_T^H S_T \quad (19)$$

ここで, $2(N/2-1)(M-L) \times L$ 複素行列 S_T を特異値分解する。

$$S_T = P_1 \Gamma P_2^H$$

ここで, P_1, P_2 は, おおの $S_T S_T^H, S_T^H S_T$ の固有ベクトルからなる $2(N/2-1)(M-L) \times 2(N/2-1)(M-L), L \times L$ ユニタリ行列である。また, $2(N/2-1)(M-L) > L$ と仮定すると, Γ は S_T の特異値 $\{\eta_i\}$ からなる $2(N/2-1)(M-L) \times L$ 対角行列である。

$$\Gamma = \text{diag} [\eta_1, \eta_2, \dots, \eta_L]$$

(10)式の $\rho(P; \{\eta_i\})$ とある閾値 T_1 との比較により, 決定した行列 S_T の階数を P と置く。以上から(19)式の最小ノルム解 \mathbf{b} は, 次式で与えられる。

$$\mathbf{b} = -P_2 \Gamma_P^{-1} P_1^H S_T \quad (20)$$

ただし, Γ_P^{-1} は, 次の $L \times 2(N/2-1)(M-L)$ 行列である。

$$\Gamma_P^{-1} = \text{diag} [1/\eta_1, 1/\eta_2, \dots, 1/\eta_P, 0, \dots, 0]$$

また, 2.2 節の推定ではリップル抑圧のためにテーパ窓を用いていたが, 本節ではリップル抑圧を必要としないため, 特異値分解においてテーパ窓を利用しない。得られた \mathbf{b} の各要素を係数とする多項式 $B(z)$ を考える。

$$B(z) = 1 + \sum_{j=1}^L b_j z^{-j} \\ = \prod_{j=1}^L (1 - \beta_j z^{-1}) = 0 \quad (21)$$

$B(z)$ を解いて求めた L 個の根 $\{\beta_j\}$ のうち, M_0 個の極は, (15)式の系列 $x_q(k)$, ($q=1, 2, \dots, M$), ($k=1, 2, \dots, N/2-1$) に含まれる M_0 個の複素正弦波の周波数に対応し, 残りの $(L-M_0)$ 個の極は雑音成分を表す¹⁶⁾。このように, 前向き・後向き予測誤差を導入し, 利用可能な周波数帯域全体にわたって同時に最適なパラメータ \mathbf{b} を求めることによって, 各仮想音源点上のスペクトルパターンを, 真の音源位置に線スペクトルがあるモデルと整合することができるため, under-determined 正規方程式に起因するリップルの影響を抑制し実音源位置を明

瞭に決定できる。

3. シミュレーション実験

提案した方法の有効性と精度を調べるために, 計算機によるシミュレーション実験を行う。2.1~2.3 節で述べる各処理は, 大型計算機 IBM 3081-KX 6 上で, 倍精度実数型 (指数部 7 bit, 小数部 56 bit) で計算を行った。また, 各実験において音速 c は 400 [m/s]⁴⁾, 減衰定数 α は 0.01 [1/m]⁴⁾, 標準化周波数 f_s は 2.4 [kHz] とする。 N_s 個のセンサで受音した各信号の長さ N は 64 点とする。また, Fig. 1 の1次元空間 l の長さは $l=10$ [m] とし, 両端等での反射は考えないものとする。

また, 2.2 節の手法における処理上のパラメータである(13)式のテーパ窓幅 N_w と(10)式の行列 $\Sigma(k)$ の階数選択時の閾値 T_0 に関しては, 論文⁷⁾の実験の最適値から $T_0=95\%$, $N_w=4$ と設定する。

実験 1

1次元空間 l 上の2個の実音源から互いに無相関な白色雑音 $x_a(n), x_b(n)$ が放射されているものとする。それらの音を l 上に等間隔に配置した $N_s=11$ 個のセンサによって観測する。各センサで受音する長さ $N=64$ 点の信号には, 互いに無相関な $S/N=20$ dB の白色雑音を付加する。各センサで受音した信号 $y_i(n)$, ($i=1, 2, \dots, N_s, n=0, \dots, 63$) は, ナイキスト周波数 $f_N=f_s/2$ の $1/2$ までの低周波帯域成分のみ含むものとする。すなわち, 受音された信号の最小波長は, $\lambda_{\min}=c/(f_s/4)=2/3$ [m] であるから, l 上に $\lambda_{\min}/2$ 幅で等間隔に $M=$

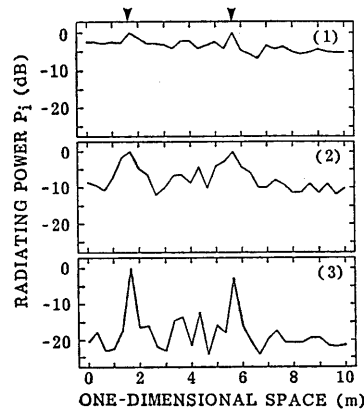


Fig. 2 The estimated power P_i of the sound radiating from each assumed sound source on the one dimensional space using the proposed method. The number M_0 of the actual sound sources is equal to 2 and their positions are represented by '▼' ($S/N=20$ dB, $N_s=11$, $L=26$, $T_1=0.9$, and $4f=f_N/2$). (1) Results obtained by the method in Sect. 2.1. (2) Results obtained by the method in Sect. 2.2. (3) Results obtained by the proposed method in Sect. 2.3.

31 個の仮想音源を設定する。

2.1~2.3 節の三つの手法によって推定した仮想音源上の音圧パワー P_i , ($i=1, 2, \dots, M$) をおのおの Fig. 2 (1)(2)(3) に示す。2.3 節の手法においては、各仮想音源上の各周波数 k ごとの振幅パターン $X_i(k)$ を、理想的には各実音源位置に極 β_j を対応させた (21) 式の $B(x)$ を分母に持つ極モデル $1/B(x)$ と整合させる。従って、(21) 式から振幅パターン $X_i(k)=1/B(x)$ を決定できる。更に、(6) 式から各点の音圧パワー P_i を決定できる。真の音源位置は、▼印で示してある。いずれの手法によっても、真の音源近くのパワーは大きくなっているが、Fig. 2 (1)(2) に比較し、Fig. 2 (3) の本手法によって、リップルを抑制し、真の音源位置をより明確に同定できる。

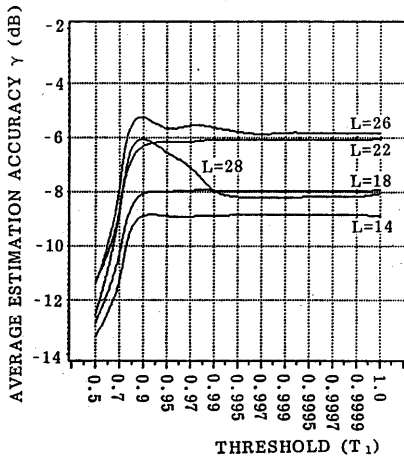


Fig. 3 The relation between the threshold T_1 and the average estimation accuracy γ for various values of the order L used in Eqs. (16)~(21). ($S/N=20$ dB, $M_s=2$, $N_s=11$, and $\Delta f=f_N/2$.)

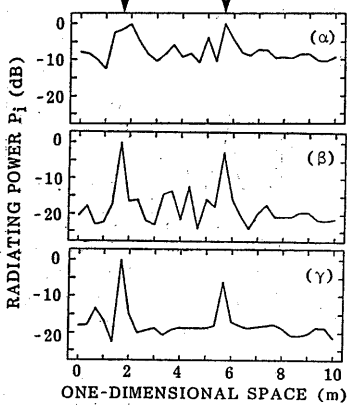


Fig. 4 The radiating power P_i estimated by the proposed method using various frequency band Δf . ($S/N=20$ dB, $M_s=2$, $N_s=11$, $L=26$, and $T_1=0.9$.) (α) $\Delta f=f_N/4$. (β) $\Delta f=f_N/2$. (γ) $\Delta f=f_N$.

実験 2

2.3 節で提案した手本法のパラメータの最適値を決定するために、次のような定量的評価実験を行う。(20) 式の Γ の階数 P を決定するための閾値 T_1 , (16)~(21) 式の極推定時の分析次数 L の 2 種のパラメータの値を変化させたときの、次式で定義される音源 1 個当たりの位置の平均推定精度 γ を Fig. 3 に示す。

$$\gamma = \frac{E\{\sum_{i \in \{\text{実音源点}\}} P_i\} / \text{実音源数}}{E\{\sum_{i \in \{\text{仮想音源点}\}} P_i\}} \quad (22)$$

ここで、 $E\{\cdot\}$ は、集合平均を表し、シミュレーション実験では、32 回独立に実音源位置を設定して得られた処理結果の平均値とする。(22) 式の γ は、1 次元空間

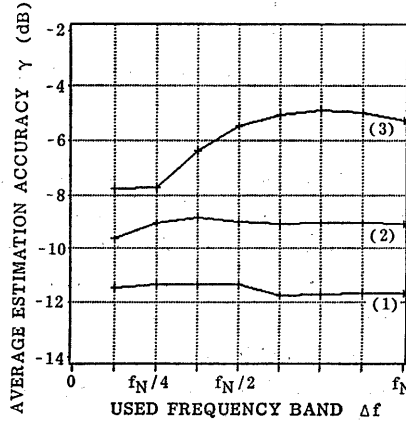


Fig. 5 The relation between the used frequency band Δf and the average estimation accuracy γ obtained by each method. ($S/N=20$ dB, $N_s=11$, $M_s=2$, $L=26$, and $T_1=0.9$.) (1) Results obtained by the method in Sect. 2.1. (2) Results obtained by the method in Sect. 2.2. (3) Results obtained by the proposed method in Sect. 2.3.

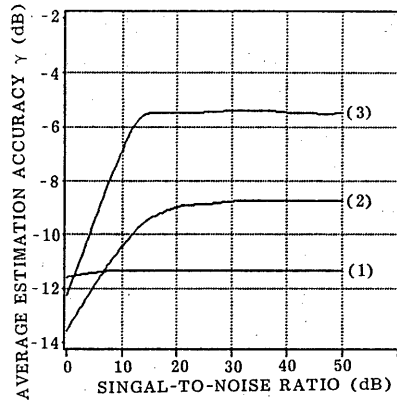


Fig. 6 The relation between the signal-to-noise ratio and the average estimation accuracy γ obtained by each method. ($N_s=11$, $M_s=2$, $L=26$, $T_1=0.9$, and $\Delta f=f_N/2$.) (1) Results obtained by the method in Sect. 2.1. (2) Results obtained by the method in Sect. 2.2. (3) Results obtained by the proposed method in Sect. 2.3.

l 上の全仮想音源の放射音圧パワーに対する、実音源1個当たりの平均音圧パワーを表す。Fig. 3の結果から、パラメータの最適値は、閾値 $T_1=90\%$ 、分析次数 $L=26$ となった。これらの値は、Fig. 2 及び以降の実験で利用される。

実験 3

上記実験1と同じ条件に関して、利用可能な周波数帯域 Δf の値を、 $f_N/4$ 、 $f_N/2$ 、 f_N と変化させたときの、本手法による各仮想点の音圧パワー P_i の推定値を Fig. 4 に示す。利用できる周波数帯域が広がるに従って、真の音源位置が明確化する。

また、2.1~2.3 節の手法に関して、 Δf を変化させて (23) 式の推定精度 γ を調べた結果を Fig. 5 に示す。このように従来の 2.1、2.2 節の手法に対して、本手法では、利用できる帯域が低周波に限られる場合においても、実音源位置が明確化し、良好な結果が得られている。

実験 4

Fig. 6 は、実験1と同じ条件で S/N を変化させたときの各手法の推定精度 γ を示す。本手法によって、

S/N=15 dB 以上において、実音源位置を高い精度で推定できる。

実験 5

Fig. 7 は、実験1と同じ条件で、センサ数を $N_s=5, 11$ と変化させた場合の、本手法による各仮想点の音圧パワー P_i の推定値を示す。センサ数が増加するに従って、音源位置が明確化している。

さらに、Fig. 8 にはセンサ数を変化させたときの推定精度 γ を示す。実音源数 M_0 が多くなるほど音源一つ当たりの推定精度は低下しているが、全体の傾向として、センサ数を10倍に増加することによって約6dBの精度の改善が得られている。

最後に、センサ数 $N_s=11$ 個、実音源数 $M_0=4$ の場合に関して、S/N を10 dB、20 dB と変化させた際に得られる 2.1~2.3 節の各手法による各仮想点の音圧パワー P_i の推定結果を Fig. 9 に示す。以上のように本手法の有効性が明らかとなった。

4. 結 論

本論文では、under-determined な正規方程式に対し

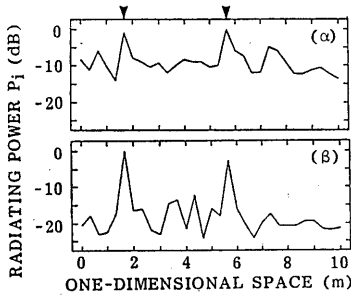


Fig. 7 The radiating power P_i estimated by using the various number N_s of the sensors based on the proposed method. (S/N=20 dB, $M_0=2$, $L=26$, $T_1=0.9$, and $\Delta f=f_N/2$). (a) $N_s=5$. (b) $N_s=11$.

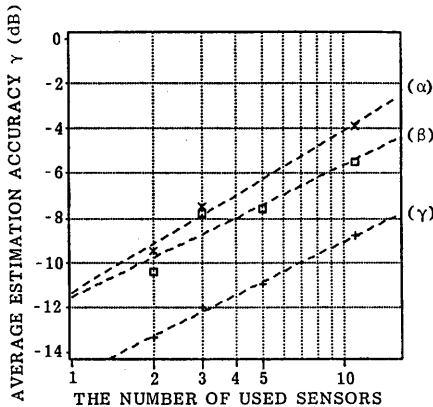


Fig. 8 The relation between the number N_s of the used sensors and the average estimation accuracy γ obtained by the proposed method for various number N_s of the actual sound sources. (S/N=20 dB, $L=26$, $T_1=0.9$, and $\Delta f=f_N/2$). (a) $M_0=1$. (b) $M_0=2$. (c) $M_0=4$.

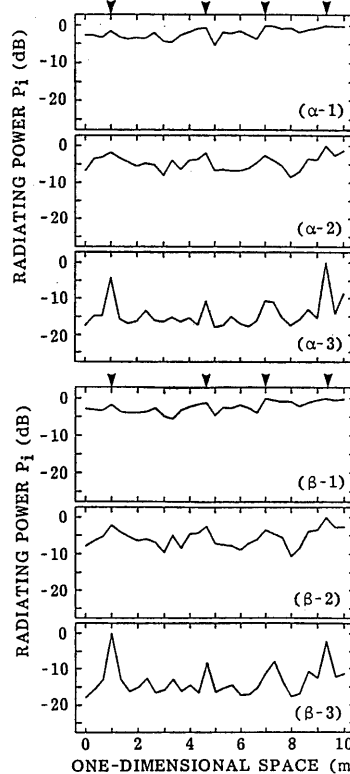


Fig. 9 The radiating power P_i estimated by each method when there are four actual sound sources. ($L=26$, $T_1=0.9$, and $\Delta f=f_N/2$). (a) S/N=10 dB. (b) S/N=20 dB. (1) the method in Sect. 2.1. (2) the method in Sect. 2.2. (3) the proposed method in Sect. 2.3.

てテーパ窓を用いた特異値分解と極モデルによる推定法を利用することによって、1次元空間上での複数音源の位置を高精度に推定するための新しい手法を提案した。また、要求される解像度を波長 $\lambda_0/2$ で表すとき、利用できる信号の最高周波数 f_{\max} が、波長 $\lambda_0/2$ に対応する周波数 $f_0=2c/\lambda_0$ よりも低い場合においても、高精度に複数音源の位置を推定できる。また、シミュレーション実験によって、本手法のパラメータの最適値及び本手法の限界についても示した。

本論文でなお未解決な課題として、次の4点が挙げられる。

(1) 本論文では、実音源数は既知であるとしたが、実際に未知である場合に、音圧パワー分布等から音源数も決定する必要がある。

(2) 本手法の有効性は、計算機シミュレーションで示されただけであり、実用化のためには、実際の音源を用いた実験的検討が必要である。本論文では、音源とセンサ間の伝達関数は、(4)式で簡略的に表されているが、伝達関数に関しては、個々の応用において実験的な検討が必要である。また、1次元空間内に反射点がある場合の処理に関する考察が必要である。

(3) 2.2節では、リップル抑圧のためにテーパ窓を用いたが、2.3節の特異値分解においてもテーパ窓を用いた方が効果があるのか、なお理論的実験的検討が必要である。

(4) (11)式又は(11')式において $X(k)$ を推定する際の特異値の切り捨て次数 R の設定法としては、本論文で用いた手法¹⁰⁾や論文¹⁷⁾の手法があるが、更に、テーパ窓を用いた場合における切り捨て次数 R 、テーパ窓の形状・幅の設定は非常に重要である。本論文ではこれらの点に関して、Fig. 3において実験的に検討しているが、なお理論的解析が必要である。

また、仮想音源数は非常に多くなるが、本手法は基本的に1次元空間ばかりでなく、2次元・3次元空間においても適用可能である。この点についても現在検討中である。

謝 辞

東北大学工学部電気工学科 中鉢憲賢教授、櫛引淳一助教授、竹田宏教授、佐藤光男助教授、情報処理教育センター 川添良幸助教授に感謝します。

文 献

- 1) "Special issue on time delay estimation," IEEE Trans. Acoust. Speech Signal Process. **ASSP-29**, 463-623 (1981).
- 2) M. T. Silvia, *Handbook of Digital Signal Processing Engineering Application*, D. E. Elliot, Ed. (Academic Press, New York, 1987), Chap. 10-11.
- 3) D. H. Youn and N. Ahmed, "Time delay estimation via coherence: An adaptive approach," J. Acoust. Soc. Am. **75**, 505-514 (1984).
- 4) A. Shima, J. Liu, H. Noto, M. Abe and K. Kido, "A new cross-spectral method to locate sound sources in one-dimensional space," J. Acoust. Soc. Jpn. (E) **9**, 1-11 (1988).
- 5) H. Kanai and K. Kido, "A new method to locate multiple sound sources in one-dimensional space based on estimation of each source sound," Am. Soc. Mech. Eng. Pressure Vessels & Piping, bound volume, **177**, 227-232 (1989).
- 6) 金井 浩, 城戸健一, "全極型伝達系を駆動するマルチパルス系列の推定 (特異値分解における Taper 窓の利用)," 機械学会論文集 (C) **54**, 2907-2914 (1988).
- 7) H. Kanai and K. Kido, "Estimation of multiple excitation pulses from response of all-pole transfer system using tapered singular-value-decomposition in low SNR cases," IEEE Trans. Acoust. Speech, Signal Process. (投稿中).
- 8) M. Yanagida, O. Kakusho and Y. Nomura, "Least-square method from source sound separation in multi-source environment," Proc. 11th ICA, Lyon-Toulouse, Paris, 167-170 (1983).
- 9) P. Lancaster and M. Tismenetsky, *The Theory of Matrices* (Academic Press, London, 1985).
- 10) J. A. Cadzow, "Spectral estimation: An over-determined rational model equation approach," Proc. IEEE **70**, 907-939 (1982).
- 11) A. V. Oppenheim and R. W. Schaffer (伊達 玄訳), デジタル信号処理 (コロナ社, 東京, 1975).
- 12) 城戸健一, デジタル信号処理入門 (丸善, 東京, 1986).
- 13) J. D. Markel, A. H. Gray, Jr., *Linear Prediction of Speech* (Springer-Verlag, Berlin, Heidelberg, 1976).
- 14) G. E. P. Box and G. M. Jenkins, *Time Series Analysis Forecasting and Control* (Holden-Day, Oakland California, 1976).
- 15) S. M. Kay and S. L. Marple, Jr., "Spectrum analysis a modern perspective," Proc. IEEE **69**, 1380-1419 (1981).
- 16) R. Kumaresan and D. W. Tufts, "Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise," IEEE Trans. Acoust. Speech Signal Process. **ASSP-30**, 833-840 (1982).
- 17) L. L. Scharf and D. W. Tufts, "Rank reduction for modeling stationary signals," IEEE Trans. Acoust. Speech Signal Process. **ASSP-35**, 350-355 (1987).