音源の方位情報を用いた複数音源の分離

黄 捷* 大西 昇* 杉江 昇*

Separating Sound Sources Using their Orientation Obtained by Localization Each

Jie Huang Noboru OHNISHI Noboru SUGIE

We propose a method for sound source separation. The only informations we need for this method is temporal disparities of arrivals of sound sources at every pair of microphones. For this purpose we linked our separating system to a localizing system. The localizing system can localize sound sources using temporal disparities at the onsets that are not mixed with ongoing portions of other sound sources. Several experiments were carried out in an anechoic chamber. We could separate two sound sources (speech sounds), located with azimuth difference of 38° , resulting in 25 dB average attenuation ratio, by a DSP system. The separable azimuth difference was more than 15° . And 5° localization error caused about 6 dB reduction in attenuation ratio.

Key Words : Sound Source Separation, Sound Source Localization, Separable Azimuth Difference

1. ま え が き

われわれ人間は雑踏の中で特定の話相手の声を聞き分けることができる。このことを心理学ではカクテルパーティー効果という。音声認識システムにおいて各音源からの音声を分離することが可能であれば、音声認識システムは多数の人間の会話を同時に聞き分け、それらに同時に応答することが可能となる。ロボットシステムにおいても複数の音信号を同時に受けとり、それぞれに対して、カメラを向ける、接近するなどの適切な行動を起こすことができる。また、雑音源も音源と見なせば、目的とする音源と雑音源を分離することによって S/N 比の改善をはかることもできる。しかし、現在のところ、フィルタに方向性を持たせ、目的音源方向以外の音をカットする指向性マイクロホンの研究は多くなされてきたが、カクテルパーティー効果の重要な手がかりの1つである音源方位の定位と関連づけることによって自動的に複数音源分離を行なう研究は少ない。金田らは目的信号のない状況をコンピュータに学習させ、環境の雑音の状態によって自動的にそれらを最大限に押さえるようなフィル

タの設定を行なうシステムを開発したり¹⁾。しかし、このシステムはあらかじめ目的信号無しの環境の学習が必要で、変化する環境には対応できない。また、多数の目的信号がある場合ではそれらを分離することができない。

一般的に観測点(マイクロホン)の数が音源の数より大きい場合では、音源から観測点までの伝達関数がわかれば、行列式の逆演算により観測点の波形から各音源の波形を求めることができる²⁾。しかし、伝達関数が未知の場合には、音源分離は難しい。伝達関数を求めるためには音源の位置、環境の反射状態などを知らなければならぬ。反響のない環境つまり自由音場では音源とマイクロホンとの相対位置がわかれば、伝達関数を知ることができる。このことから音源分離において音源の位置情報を知ることが重要な意味を持つことがわかる。

複数音源定位を行なうためには、何らかの方法で各音源が分離する時刻あるいは周波数領域を見いだす必要があると考えられる。我々は従来から音源定位について研究してきた^{3,4)}。単一音源の場合では、3つのマイクロホンの出力間の時間差により、音源方向を一義に定位することができる。また、音の onset (立上り部)を優先的に利用して音源定位すれば、直接音と反響音との経路差により反響音の影響を除去することができる。複数音源

原稿受付 1990年7月23日

* 名古屋大学工学部

の場合は、音源の断続性という仮定により、ほかの音源の ongoing (持続部) に埋もれない音の onset が確率的に存在することがわかる。このような onset は単一音源しか含まないのでこれを利用して特定の音源を定位することができる。また、onset の数が多数あれば、多数の onset において音源定位をし、統計的な処理を行なうことによって複数の音源を同時定位できる。

ところで、森田らは人間の音声を対象にフォルマント周波数による分離を試みた⁵⁾。彼らは各音源となる人間の音声のフォルマント周波数は任意の時刻においてお互いに異なるという仮定のもとで、フォルマント周波数ごとに音源定位をし、その情報を用いて音源分離を試みた。彼らの方法が周波数領域での分離定位であるのに対し、我々の方法は時間領域での分離定位といえる。そのいずれが有効であるかは音源の性質によって決定される。フォルマント周波数による分離はフォルマント周波数のある音源に限定されるし、onset による分離は音源の onset の数によって限定される。また、これらの方法は二者択一的ではなく、音源によって両者を組み合わせて音源定位を行なうのも有効と思われる。本論文ではわれわれはまず複数音源定位の結果を利用した複数音源の分離方法について述べ、次に音源分離実験およびその結果の評価実験について述べる。

2. 音源分離の方法

この章では、まず onset による複数音源の方位定位について述べ、次に音源の方位情報を用いた複数音源の分離と安定アルゴリズムの実現について述べる。

2.1 複数音源の方位定位⁴⁾

人間の会話などは単音あるいは音素の連続として考えられる。また、個々の単音あるいは音素の間は連続的になっている場合もあるが、無音部分によって隔たれる場合も多い。特に呼吸を伴う場合は、呼吸段落と言われる長い無音部分が続く。一般的に連続的にしゃべる人間の音声は約 30% の無音部分を含む。このように人間の音声、あるいは無音部分を多く含む一般音では、ある音源が別の音源の無音部分にその onset が生じる可能性は十分高く、このような onset は 2 つ以上の音源からの音が同時に始まりさえしなければ、単一音源からの音しか含まないと考えられる。このような onset をわれわれは複合音の onset と呼ぶことにする。従って、onset を多く含む音源では複数の音源が存在しても、その複合音の多数の onset でそれぞれを定位できる。以下ではその具体的な方法について簡単に述べる。

まず、正三角形（一辺の長さが 13.5cm）の頂点に配置されたマイクロホンセットからの信号を複数のバンド

パスフィルタセットおよび A/D コンバータを通してコンピュータに入力する。時間差の検出方法として処理時間の短いゼロクロス法を用いる。バンドパスフィルタを通した信号は準正弦波と考えられる。マイクロホン間の波形の対応は許容される最大時間差内に複数サイクルが含まれるような高周波領域では一義的には決まらない。そこで、まず時間差のいくつかの候補値を決める。単一の onset だけでは、一つの音源しか定位できないから、複数の onset を検出して、時間差候補値を求め、ヒストグラムを作る。 $h_{ij}^{f_n}(\Delta\tau)$ をマイクロホン $i, j (i, j \leq 3)$ 間の成分周波数 f_n (f_n はバンドパスフィルタの中心周波数) の時間差 $\Delta\tau$ ($|\Delta\tau| \leq \text{IMD}$, IMD はマイクロホン間を音が通過するのに要する時間である) における時間差候補値の度数とする。周波数成分についての合成ヒストグラムを以下のようにつくる。

$$h_{ij}(\Delta\tau) = \min(h_{ij}^{f_1}(\Delta\tau), h_{ij}^{f_2}(\Delta\tau), \dots, h_{ij}^{f_n}(\Delta\tau), \dots) \quad (1)$$

さらに、方位 θ にある音源がマイクロホン i, j 間に生ずる時間差を $\Delta\tau_{ij}(\theta)$ とする。音源方位についての合成ヒストグラムを以下のように合成する ($0 \leq \theta < 2\pi$)。

$$H(\theta) = \min(h_{12}(\Delta\tau_{12}(\theta)), h_{23}(\Delta\tau_{23}(\theta)), h_{31}(\Delta\tau_{31}(\theta))) \quad (2)$$

以上で得られるヒストグラム $H(\theta)$ は音源の方位に対するヒストグラムで、 $H(\theta)$ の度数から音源の存在する方位についての情報が得られる。音源方位を決める方法として、まず、ヒストグラムをピークと谷によっていくつかのブロックに分割する。ブロックの数から音源の数、各ブロックの平均角度から音源の方位が求まる。

2.2 複数音源の分離

音源の数が N 、マイクロホンの数が M の場合を考える。音源 j の信号を $x_j(t)$ 、マイクロホン i での信号を $y_i(t)$ 、それらのラプラス変換をそれぞれ $X_j(s)$ 、 $Y_i(s)$ 、音源 j からマイクロホン i までの伝達関数を $G_{ij}(s)$ とすると、

$$Y_i(s) = \sum_j G_{ij}(s) X_j(s) \quad (3)$$

と表すことができる。また、

$$\mathbf{Y} = (Y_1(s), Y_2(s), \dots, Y_M(s)),$$

$$\mathbf{X} = (X_1(s), X_2(s), \dots, X_N(s))$$

とし、 $\mathbf{G}$ を $G_{ij}(s)$ を要素とする $M \times N$ の行列とすると、

$$\mathbf{Y} = \mathbf{G}\mathbf{X} \quad (4)$$

と表すことができる。 $\mathbf{Y}$ はマイクロホンの波形で既知である。 $\mathbf{Y}$ から $\mathbf{X}$ を求めるためには音源からマイクロホンまでの伝達関数を知る必要がある。反響のない環境（つまり自由音場）では、音源 j からマイクロホン i までの伝達関数は $G_{ij}(s) = \frac{K}{\tau_{ij}} e^{-s\tau_{ij}}$ と表すことができる。

ここで、 K は定数である。 τ_{ij} は音源 j からマイクロホン i までの音の伝達時間で、これは両者の間の距離に比例する。音源からマイクロホンまでの距離がマイクロホン間距離より十分大きければ、距離の差によって生じる音のマイクロホン間での振幅の差は無視することができ、 $G_{ij} = K'e^{-s\tau_{ij}}$ と表すことができる。ここで、 $K' \approx \frac{K}{\tau_{ij}}$ である。また、実際に各マイクロホンの波形から求められるのはマイクロホン間の相対時間差である。ここで、 $\tau_j^{\min} = \min(\tau_{1j}, \tau_{2j}, \dots, \tau_{Mj})$ とし、

$$G_{ij}' = \frac{G_{ij}}{K' \exp[-s\tau_j^{\min}]} = \exp[-s(\tau_{ij} - \tau_j^{\min})],$$

$$X_i' = K' X_i \exp[-s\tau_j^{\min}]$$

とする。 G' を G_{ij}' 要素とする行列とすると、

$$Y = G' X' \quad (5)$$

となる。ここで、 $\tau_{ij} - \tau_j^{\min}$ は0かマイクロホン間の時間差である。従って、音源定位システムにより G' を知ることができる。行列 G' のランク R が音源の数 N より大きければ、行列の逆演算から X' を求めることができる²⁾。特に $R = M = N$ の場合では

$$X' = G'^{-1} Y \quad (6)$$

と表すことができる。ここで、 G'^{-1} は G' の逆行列である。

音源、マイクロホンの数が2の場合では、上式より

$$\begin{aligned} & \begin{bmatrix} K' X_1 \exp[-s\tau_1^{\min}] \\ K' X_2 \exp[-s\tau_2^{\min}] \end{bmatrix} \\ &= \frac{1}{(\exp[-s(\tau_{11} - \tau_1^{\min} + \tau_{22} - \tau_2^{\min})] - \exp[-s(\tau_{21} - \tau_2^{\min} + \tau_{12} - \tau_1^{\min})])} \\ & \times \begin{bmatrix} \exp[-s(\tau_{22} - \tau_2^{\min})] \\ -\exp[-s(\tau_{21} - \tau_1^{\min})] \\ -\exp[-s(\tau_{22} - \tau_2^{\min})] \\ \exp[-s(\tau_{11} - \tau_1^{\min})] \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \end{aligned}$$

となる。 $\Delta\tau_j = \tau_{2j} - \tau_{1j}$ とし、 $\Delta\tau_1 > \Delta\tau_2$ の場合では

$$X'' = K' \begin{bmatrix} X_1 \exp[-s(\tau_{11} + \tau_{22} - \tau_2^{\min})] \\ X_2 \exp[-s(\tau_{11} + \tau_{22} - \tau_1^{\min})] \end{bmatrix}$$

$\Delta\tau_1 < \Delta\tau_2$ の場合では

$$X'' = K' \begin{bmatrix} -X_1 \exp[-s(\tau_{12} + \tau_{21} - \tau_2^{\min})] \\ -X_2 \exp[-s(\tau_{12} + \tau_{21} - \tau_1^{\min})] \end{bmatrix}$$

とおくと、

$$\begin{aligned} & \begin{bmatrix} X_1'' \\ X_2'' \end{bmatrix} = \frac{1}{(1 - \exp[-s|\Delta\tau_2 - \Delta\tau_1|])} \\ & \times \begin{bmatrix} \exp[-s(\tau_{22} - \tau_2^{\min})] \\ -\exp[-s(\tau_{21} - \tau_1^{\min})] \\ -\exp[-s(\tau_{12} - \tau_2^{\min})] \\ \exp[-s(\tau_{11} - \tau_1^{\min})] \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \quad (7) \end{aligned}$$

と統一した式を得ることができる。明らかに X'' は X を定数倍し、時間的に遅らせたものとなっている。上式から $\Delta\tau_1 = \Delta\tau_2$ の場合、つまり音源1と音源2が同方向

Fig. 1 Difference in temporal disparities.

あるいはマイクロホン対に対して前後対称の位置にある場合は音源1と音源2とを分離することはできないことがわかる。ただし、音源1、音源2がマイクロホン対の前後対称位置にある場合は、別のマイクロホン対を選ぶことによって $\Delta\tau_1 = \Delta\tau_2$ となるのを避けることができる。また、時間差の差 $|\Delta\tau_2 - \Delta\tau_1|$ の値は大きいほうが音源分離しやすいこともわかる。したがって、2つの音源を分離する場合、2つのマイクロホンで十分な場合もあるが、2つの音源がマイクロホン対の前後対称、あるいはそれに近い方位にある時は時間差の差の値が小さくなり、音源分離ができなくなることが発生する。上述現象を防ぐためには必要最小限のマイクロホン数よりも1つ多くマイクロホンを使う必要がある。本研究では音源定位で使う3つのマイクロホンをそのまま音源分離でも使い、時間差の差の値が最大となるように常に最適のマイクロホンペアを選ぶことができる。Fig. 1に示すように、音源1, 2の方位がそれぞれ $\pm \frac{\theta}{2}$ で、マイクロホンペアを結ぶ直線の方角を α とする。音源1, 2のマイクロホンペアに生じる時間差の差は近似的に

$$\begin{aligned} |\Delta\tau_2 - \Delta\tau_1| &= \frac{d}{C} \left| \cos\left(\alpha - \frac{\theta}{2}\right) - \cos\left(\alpha + \frac{\theta}{2}\right) \right| \\ &= \frac{2d}{C} \sin(\alpha) \sin\left(\frac{\theta}{2}\right) \quad (8) \end{aligned}$$

である。ここでは、 d はマイクロホン間隔で、 C は音速である。上式から時間差の差は $\alpha = \frac{\pi}{2}$ 、つまりマイクロホンペアを連結する直線が音源方位の中心線と垂直になる時に最大になることがわかる。マイクロホン数が3の場合ではマイクロホンペアのなす角度は $\frac{\pi}{3}$ であるから最適のマイクロホンペアを選ぶことによってつねに $\frac{2d}{C} \sin\left(\frac{\pi}{3}\right) \sin\left(\frac{\theta}{2}\right)$ 以上の時間差の差を得ることができる。

Fig. 2 Experimental setup.

2.3 アルコリズムの安定化

式(7)は $s=0$ の時に極を持ち、不安定である。人間の耳はある程度のひずみがあってもそれを感知することができないか、あるいは、感知できてもさほど気にはならない。そこで、1より小さい正の定数 k を選んで、以下のようにする。

$$\begin{aligned} \begin{bmatrix} X_1'' \\ X_2'' \end{bmatrix} &= \frac{1}{(1 - \kappa \exp[-s|\Delta\tau_2 - \Delta\tau_1|])} \\ &\times \begin{bmatrix} \exp[-s(\tau_{22} - \tau_2^{\min})] \\ -\exp[-s(\tau_{21} - \tau_1^{\min})] \\ -\exp[-s(\tau_{12} - \tau_2^{\min})] \\ \exp[-s(\tau_{11} - \tau_1^{\min})] \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \quad (9) \end{aligned}$$

κ の大きさを適切に選べば、人間の耳では気にならない程度にひずみを押え、式を安定化させることができる。

上式を時間領域での表現に戻すと、

$$\begin{cases} x_1''(t) = y_1(t - (\tau_{22} - \tau_2^{\min})) \\ \quad - y_2(t - (\tau_{12} - \tau_2^{\min})) + \kappa x_1''(t - |\Delta\tau_2 - \Delta\tau_1|) \\ x_2''(t) = y_1(t - (\tau_{21} - \tau_1^{\min})) \\ \quad - y_2(t - (\tau_{11} - \tau_1^{\min})) + \kappa x_2''(t - |\Delta\tau_2 - \Delta\tau_1|) \end{cases}$$

Fig. 3 Diagram of sound source separating system.

となる。

3. 音源分離実験および結果の評価

3.1 音源分離実験

音源分離実験はスピーカから発生した二つの音源について行なわれた。音源1はラジオで収録した天気予報のアナウンスで、音源2はやはりラジオで収録した男女2人の対談である。時間はともに約20秒である。音源とマイクロホン (m_1, m_2, m_3 はマイクロホン) は Fig. 2 のように配置され、音源1と音源2は約38度分かれておかれる。実験が行なわれた環境は無響室で、大きさが $5.5 \times 5.5 \times 3.0 \text{ m}^3$ で、カットオフ周波数が200 Hzで、暗騒音 (background noise) が20 dBである。特性の校正された三つの無指向性マイクロホン (リオン社 NA-20 型) は正三角形 (辺の長さが13.5 cm) の頂点に配置された。また、音源1の方位角 θ_1 を0度から120度まで15度ずつ変化させて実験を行なった ($\theta_2 = \theta_1 + 38^\circ$)。Fig. 3に複数音源分離システムのブロック図を示す。まず、複数音源定位システムによって音源の onset が検出され、時間差ヒストグラムから複数音源の方位が決定される。次に三つのマイクロホンペアから $|\Delta\tau_2 - \Delta\tau_1|$ が最大となるマイクロホンペアを選ぶ。音源定位システムの定位結果によって、音源分離システムの遅延素子の定数が決定される。なお、音源分離システムは DSP (TMS 320 C 25) によるデジタルフィルタによって実現される。

Fig. 4に音源1と音源2、複合音および分離後の音源1と音源2の波形の一部を示す。

3.2 音源分離結果とその評価

3.2.1 音源定位の結果の評価

音源分離の度合を表すものとして、分離した二つの音源の相関係数をとる方法が考えられる。音源1を s_1 、音源2を s_2 とすると、音源分離の結果はそれぞれ $s_1 + \kappa_2 s_2$ 、 $s_2 + \kappa_1 s_1$ と表すことができる。ここで、 κ_1 と κ_2

を減衰比と呼ぶことにする。 $0 < \kappa_1, \kappa_2 < 1$ で $\kappa_1 s_1$ と $\kappa_2 s_2$ はそれぞれ残留音源を示す。相関係数を以下のよう表すことができる。

$$\rho(s_1 + \kappa_2 s_2, s_2 + \kappa_1 s_1) = \frac{(s_1 + \kappa_2 s_2)(s_2 + \kappa_1 s_1)}{\sqrt{(s_1 + \kappa_2 s_2)^2 (s_2 + \kappa_1 s_1)^2}} \quad (11)$$

簡単のために $\kappa_1 = \kappa_2$ とし、分母の κ_1 と κ_2 項が無視できるほど小さいと仮定する。分子部分は音源 1 と音源 2 が完全独立であるなら $\overline{s_1 s_2} = 0$ となり、

$$\rho = \frac{\kappa(\overline{s_1^2} + \overline{s_2^2})}{\sqrt{\overline{s_1^2} \overline{s_2^2}}} \quad (12)$$

となる。さらに $\overline{s_1^2}, \overline{s_2^2}$ は音のレベルによるから、 $\overline{s_1^2} = \overline{s_2^2}$ となるように正規化を行えば

$$\kappa = \frac{\rho}{2} \quad (13)$$

と表すことができる。Table 1 に音源の方位 θ_1, θ_2 、分離した音源の相関係数 ρ およびデシベル表示に変換した κ の値を示す。表からわかるように 2 つの音源を 23 dB から 33 dB の減衰比で分離できたことがわかる。

以上の結果が人間の聴覚特性に一致するものかどうかを調べるために、聴覚テストを行ない、Table 1 の結果が人間の聴覚と明らかな差がないことが証明できた。聴覚テストとはテスト音源 1 (s_1) とテスト音源 2 (s_2) を以上の結果で得られた減衰比に従って複合音をつくり、二人の被験者に分離後残留音源とのレベルがマッチするかどうかを判断させるテストである。なお、3.2.3 で述べる聴覚テストは二人の被験者に音源 1 と音源 2 のレベルを調節させ、テストの対象となる音源とのレベルがマッチしたところの相対レベルを分離音源の減衰比とするものである。

3.2.2 分離可能な音源方位差の限界

音源分離の難易を決定するのは時間差の差 $|\Delta\tau_2 - \Delta\tau_1|$ の値である。この時間差の差を決定するのは音源の方位差と音源とマイクロホンとの配置である。音源の方位差が大きければマイクロホンペアを選ぶことによっ

Fig. 4 Wave form of separated sound source.

て、時間差の差をつねに $\frac{2d}{C} \sin\left(\frac{\pi}{3}\right) \sin\left(\frac{\theta}{2}\right)$ 以上に選ぶことができる。しかし、音源の方位差が小さければ、時間差の差を大きくとることはできない。従って、音源分離は音源の方位差によって決定される。われわれはいろいろな音源方位差について実験する代わりに、音源分離実験と同じ条件設定の中で、いろいろな時間差の差の場合について実験を行ない（つまり、時間差の差を最大に選ぶルールを無視する）、それから分離できる最小限の音源方位差を推定する。Table 2 にいくつかの時間差

Table 1 Relation between correlation coefficient ρ and attenuation ratio κ .

No.	$\theta_1, \theta_2 (^{\circ})$	ρ	$\kappa(\text{dB})$
1	0,38	0.10	26
2	15,53	0.13	23
3	30,68	0.09	26
4	45,83	0.12	24
5	60,98	0.07	29
6	75,113	0.04	33
7	90,128	0.14	23
8	105,143	0.14	23
9	120,158	0.12	24

Table 2 Relation between difference in time disparity $|\Delta\tau_2 - \tau_1|$ and correlation coefficient ρ .

No.	$ \Delta\tau_2 - \tau_1 (\mu\text{s})$	ρ	$\theta_1^{\circ}, (\text{microphone pair})$
1	14	0.38(X)	75,(2,3)
2	20	0.51(X)	15,(3,1)
3	45	0.38(Δ)	0,(2,3)
4	45	0.30(Δ)	120,(1,2)
5	51	0.40(Δ)	60,(2,3)
6	80	0.24(O)	90,(2,3)
7	85	0.33(O)	30,(3,1)
8	109	0.20(O)	105,(1,2)
9	113	0.28(O)	45,(2,3)

Table 3 Relation between azimuth error and reduction in attenuation ratio κ_e .

error of azimuth (°)	$\frac{\kappa_e}{\kappa}$ (dB)
5	6
10	12

の差 $|\Delta\tau_2 - \Delta\tau_1|$ と分離した音源の相関係数 ρ および音源 1 の方位 θ_1 とマイクロホンの番号を示す。Table 2 から分離音源の相関係数が Table 1 のものと比べてかなり大きいことがわかる。また、時間差の差が小さくなるにつれて分離の減衰比が落ち、同時に音のひずみが増大し、ノイズが増えることが聴覚テストによって明らかにされた。結果は 3 段階で評価され、×は分離の失敗を示し、△はある程度は分離できているが、結果がよくないことを示し、○は明らかな分離効果があったことを示す (Table 2)。以上の結果から明らかな音源分離効果が期待できるのは音源の方位差が約 125 度 (時間差の差約 88 μ s) 以上の場合と考えることができる。

3.2.3 音源定位の誤差による影響

本システムは音源定位の結果により音源分離を行なうもので、音源定位の誤差は音源分離に直接影響する。音源定位の誤差による影響を調べるために聴覚テストを行ない、分離減衰比の変化を調べた。Table 3 にその結果を示す ($\theta_1 = 0^\circ$, $\theta_2 = 38^\circ$, 音源 1 についての聴覚テストである)。結果として定位誤差 5 度の時は減衰比が約 6 dB 落ち、定位誤差 10 度の時はさらに約 6 dB 落ちることがわかった。無響室では音源定位誤差は 5 度以内であるので定位誤差による影響は 6 dB 以内であることが

わかる。

4. ま と め

本研究は onset による複数音源定位システムをベースに複数音源分離の方法を提案し、実験によりその有効性を示した。音源分離の減衰比は平均 25 dB, 分離できる音源の方位差は 15 度以上で、音源定位の誤差 (5 度以内) による減衰比の変化は 6 dB 以内であった。

謝辞

無響室を使わせていただいた名古屋市工業研究所の皆様に深く感謝します。なお、この研究の一部はサウンド技術振興財団、放送文化基金および文部省科学研究賞補助金重点領域研究 (1)「脳の計算理論の研究」の助成によるものです。

参 考 文 献

- 1) Y. Kaneda and J. Ohga: "Adaptive microphone-array system for noise reduction", IEEE Trans. Acoust. Speech and Singnal Process, **ASSP-34**, 6, pp. 1391-1400 (Dec. 1986).
- 2) 柳田益造, 角 所 収, "一般逆たみ込み行列を用いた音源波形分離の定式化", 昭和 56 年日本音響学会講演論文集 2-7-7, pp. 499-500 (1981)
- 3) 黄 捷, 大西 昇, 杉江 昇, "生体に示唆を得た音源定位システム…反響のある環境での単一音源定位…", 電子情報通信学会論文誌 **AJ 71-A-10**, pp. 1780-1789 (Oct. 1988).
- 4) 黄 捷, 大西 昇, 杉江 昇, "時間差ヒストグラムを用いた複数音源定位システム", 日本ロボット学会誌 Vol.9, No.7, pp. 29-38 (1991).
- 5) 村井伸行, 森田龍弥, 藤井克彦, "複数話者の音声分離", 第 29 回自動制御連合講演会論文集 pp. 611-612 (1986).