

話者追尾2チャンネルマイクロホンアレーに関する検討

永田 仁史[†] 安倍 正人[†]

Two-Channel Adaptive Microphone Array with Target Tracking

Yoshifumi NAGATA[†] and Masato ABE[†]

あらまし 本論文では、音声認識や携帯電話などへの高品質な音声入力を目的とした2チャンネル(ch)適応マイクロホンアレーについて述べる。このマイクロホンアレーは、適応ビームフォーマを二つ並列動作させて話者方向を追尾し、適応処理で問題となっていた目的信号除去による性能低下を防止する。追尾に必要な話者方向は、目的信号を抑圧するように制約をおいたビームフォーマから検出しており、従来困難だった2chによる追尾処理を可能にした。また、従来マイクロホンアレーで問題となっていた空間的なエリアシングによる性能低下の問題に対し、アレーを構成するマイクロホンの指向性を利用して影響を小さくする方法を提案した。移動する音源の信号抽出精度について計算機シミュレーションにより検討した結果、2chでは従来困難だった移動音源の信号抽出をSNR15dBまで改善できることを確認できた。

キーワード マイクロホンアレー、指向性、適応ビームフォーマ、エリアシング、一般化サイドローブキャンセラ、話者追尾

1. ま え が き

ハンズフリーの音声認識や携帯機器への音声入力を目的とし、複数のマイクロホンの信号から自動的に雑音方向の感度を低くするように指向性を形成して雑音を抑圧する適応マイクロホンアレーが研究されている[5]~[7],[9]。

適応マイクロホンアレーは、少数のマイクロホンで高い性能を得られるため、パソコンや携帯機器など、ハードウェア上の制約から音声入力のチャンネル数が限られる場合に有利である。このため、筆者らは、高雑音下での音声認識や携帯電話などへの音声入力を目的とし、パソコン等のステレオ音声入力を利用可能な2chの適応マイクロホンアレーを開発している。

適応マイクロホンアレーの処理方式としては、Frostのビームフォーマ[2]、Griffiths-Jimの一般化サイドローブキャンセラ(Griffiths-Jim Generalized Sidelobe Canceller: GJGSC)[4]、パイロット信号法[1],[5]などが知られている。これらのビームフォーマは、一般に平面波の伝搬モデルを仮定し、目的信号の到来方向を既知として処理するが、信号の伝搬状況

が仮定している伝搬モデルからずれた場合や、想定している目的信号の方向と実際の目的信号の到来方向とがずれた場合に目的信号の除去が生じて性能が低下するため、これを回避する種々のロバスト化手法が提案されてきた[3],[6],[8],[9],[11]。

2chのアレーにおけるロバスト化手法としては、補聴器の雑音抑圧を目的としたGreenbergらの方法[6]がある。この方法では、目的信号の除去が、目的信号の対雑音比(SNR)が高いほど生じやすいことに着目し、SNRの指標として2ch間の相互相関と出力パワーを用いて適応フィルタの収束速度を制御している。到来方向を反映した相互相関に基づいて適応の停止/継続を制御し、到来方向のずれを許容するように処理しているが、方向のずれを大きく許容するように設定すると、適応の停止時間が多くなりフィルタの収束が遅くなるという問題がある。

多チャンネルアレーにおけるロバスト化手法としては、到来方向のずれを許容するようにフィルタの指向性に拘束を設ける方法[3],[8]がある。この方法は、先の方法のような適応速度の制御がないため、収束速度に問題はないものの、拘束条件によりフィルタの自由度が減少して雑音除去性能が低下するため、もともと自由度の小さい少数マイクロホンのアレーに適用するのは困難である。

[†] 岩手大学工学部、盛岡市
Department of Computer and Information Science, Iwate
University, Morioka-shi, 020-0066 Japan

これに対し、GJGSCにおいて、目的信号を除去するブロッキングフィルタをリーク適応フィルタとすることにより、不完全な目的信号除去を改善し、到来方向のずれの影響を抑える手法が開発されている [11]. この方法では、フィルタの自由度が損なわれることはないものの、固定ビームフォーマ出力をリーク適応フィルタの参照信号としており、マイクロホン数が少ない場合参照信号のSNRが低下するため、2ch アレーへの適用には問題がある。

以上の方法は、到来方向のずれを許容するための手法であり、ずれによる周波数特性のひずみ等の問題に関しては考慮されていない。これに対し、信号の到来方向を追尾する方法が提案されている。この方法は、フィルタの自由度を減少させずに到来方向のずれを小さくでき、また、ずれに起因する周波数特性のひずみを小さくできる点でも有利である。中でも、Affes らの方法 [9] は、少ない演算量で目的音を追尾できるが、追尾開始時に音源方向の初期値を与える必要がある。しかし、2ch のアレーにおいては、妨害音と目的音の2音源のある場合に音源方向を2ch の入力から推定する必要があり、従来の音源方向推定法では対処できなかった。

そこで、本論文では、追尾方式の利点を生かすため、2音源、2ch 入力の場合でも音源方向を推定し、推定した音源方向にビームフォーマの向きを逐次再設定することにより追尾を行う2ch の話者追尾マイクロホンアレーを提案する。音源方向の推定は、適応ビームフォーマのつくる指向性上の死角を検出することによって可能とした。

更に、2ch マイクロホンアレーにおいて問題となる空間的なエリアシングの問題に対し、マイクロホンの指向性を利用してエリアシングを抑える方法を提案する。

以下、本論文では、まず、2ch の適応ビームフォーマの空間的エリアシングについて検討し、これを避けるためのマイクロホン配置について述べる。続いて、2ch 話者追尾マイクロホンアレーの構成と処理について述べた後、話者が移動する場合の追尾性能と目的音抽出性能について計算機シミュレーションにより検討する。

2. 2ch マイクロホンアレーの空間的エリアシング

適応ビームフォーマは種々の方式が提案されている

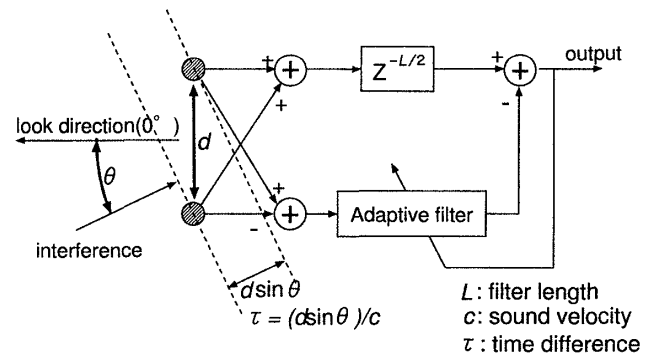


図1 2ch Griffiths-Jim GSCの構成
Fig.1 Two-channel Griffiths-Jim GSC.

が、ここでは、演算量と性能の点で有利なGJGSCに関して検討した。図1に2chの場合のGJGSCの構成を示す。なお、以下では、混乱を避けるため、ビームフォーマにおいて想定している目的音の到来方向を入力方向と呼ぶことにする。

図1に示すように、2chのGJGSCは、二つの入力信号の差の信号を参照信号とし、入力信号を加算した信号を希望応答として適応ノイズキャンセルを行う。図は、マイクロホンの並びに垂直な方向を入力方向とするブロードサイドアレーの場合を示している。図に示すように、妨害音は時間差 τ を伴って到来する一方、目的音は二つのマイクロホンに同時に到達するため、2ch間の差をとることによって目的音の成分が除去され、理想的には妨害音成分だけが残る。入力信号の差をとる処理は、GSCにおいて目的信号を抑圧するブロッキングフィルタに相当する。

ブロッキングフィルタは、目的音だけを除去するのが目的であるが、これによって妨害音成分の一部が消去されることがある、この場合、適応フィルタの参照信号から妨害音成分の一部が欠落することになるので、その成分に関しては、適応ノイズキャンセラが働かない。このような妨害音成分の消去は、妨害音成分の2ch間の位相差が 2π の整数倍になるとき、すなわち、図1において、妨害音の伝搬時間差 τ と周波数 f が

$$2\pi f\tau = 2n\pi \quad (n = 1, 2, \dots) \quad (1)$$

となる場合に生じ、マイクロホン間隔を大きくするほど妨害音を除去できない帯域が多く現れる。エリアシングを回避するには、マイクロホン間隔を利用帯域の上限の周波数の波長以下に設定すればよいが、2chの場合、マイクロホン間隔を狭めると低域の抑圧性能低

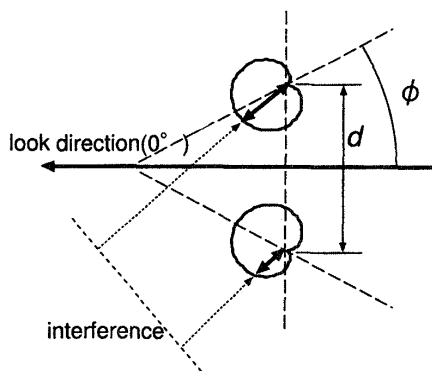


図2 指向性マイクロホンの配置
Fig. 2 Microphone arrangement.

下と、音源追尾の際の音源方向推定における分解能低下を招く。このため、マイクロホン間隔を比較的大きく保ったまま空間的エリアシングを回避する必要がある。

そこで、指向性マイクロホンを図2に示すように、指向性の軸の角度を互いに傾けて配置する。この場合、目的方向から到来する信号の振幅は各チャンネルで等しい一方で、目的方向以外からの信号はチャンネルごとに振幅が異なっており、式(1)が成り立つ周波数であっても差分により完全には消去されないため、エリアシングの軽減が期待できる。

3. 話者追尾マイクロホンアレー

3.1 話者追尾マイクロホンアレーの構成

話者の方向に合わせてビームフォーマの入力方向を変えるには、入力音声から話者方向を推定する必要がある。2ch 入力の場合、従来の方向推定法により推定できる音源数はたかだか1個であるが、2ch マイクロホンアレーは、目的音のほかに1個の妨害音がある場合これを除去できることが特徴であるため、音源が2個の条件で音源方向の推定を行う必要がある。そこで、一般的な方向推定法ではなく、ビームフォーマのフィルタの特性を利用して音源方向を推定する。

ビームフォーマは、妨害音方向に感度が低くなるように自動的に指向性を形成するため、指向性の上で極小値を示す方向を妨害音方向とすることができる。通常、ビームフォーマは、入力方向に対して感度を一定に保つように制約をおいているので目的音の方向に死角はできないが、別のビームフォーマを並列処理して入力方向を妨害音方向に設定すれば目的音方向を求めることが可能となる。二つのビームフォーマを使った

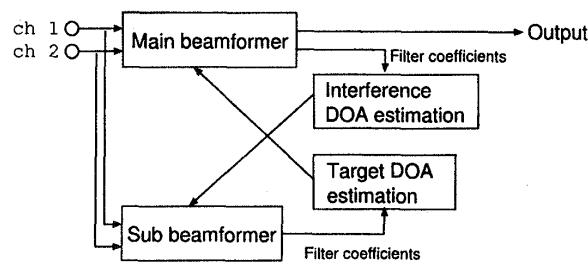


図3 話者追尾マイクロホンアレーの構成
Fig. 3 Two-channel adaptive microphone array with target tracking.

2ch マイクロホンアレーの構成を図3に示す。

図3に示すように、マイクロホンから入力した信号を二つのビームフォーマで並列に処理する。ビームフォーマの一方は妨害音を抑圧して目的信号を出力する主ビームフォーマ（主BF）、もう一方は、妨害音方向に入力方向を設定し、目的音を抑圧する副ビームフォーマ（副BF）である。

この構成において、主BFのフィルタから妨害音方向を推定して副BFの入力方向とし、副BFのフィルタから目的音方向を推定して主BFの入力方向とする。方向推定と入力方向の設定を短い時間間隔で繰り返すことにより話者方向を追尾する。

3.2 音源方向の推定

音源方向の推定は、ビームフォーマの指向性上の最小値を求めることにより行う。指向性に関しては、計算量を考慮し、GSCの適応フィルタを一般的な空間フィルタに変換した後、周波数領域で方向ごとの応答を求めて計算した。計算の手順を以下に述べる。

(a) GSCと一般の空間フィルタとの関係[10, pp.227-235]を用いてGSCのフィルタから空間フィルタへ変換する。2chの場合は次式となる。

$$h_0(i) = \begin{cases} w(i) & (i \neq L/2) \\ w(i) + 1 & (i = L/2) \end{cases} \quad (2)$$

$$h_1(i) = \begin{cases} -w(i) & (i \neq L/2) \\ -w(i) + 1 & (i = L/2) \end{cases} \quad (3)$$

ここで、 i はタップ番号、 $w(i)$ はGSCのフィルタ、 $h_0(i)$ は0ch目の空間フィルタ、 $h_1(i)$ は1ch目の空間フィルタである。

(b) フィルタをFFTにより周波数領域に変換する。

$$H_m = \text{FFT}[h_m] \quad (m = 0, 1) \quad (4)$$

(c) 次式により、追尾範囲にわたる指向性 $D(\theta)$ を計算する。

$$D(\theta) = \sum_{f_{min}}^{f_{max}} |S^T(f, \theta)H(f)|^2 \quad (5)$$

ここで、 $H(f) = \{H_0(f), H_1(f)\}^T$ は周波数領域の空間フィルタベクトル、 $S(f, \theta)$ は、想定する音源方向に対応した遅延ベクトルであり、2ch のブロードサイドアレイの場合、

$$S(f, \theta) = \{1, e^{-j \cdot 2\pi f d \sin(\theta)/c}\}^T \quad (6)$$

である。ここで、 d はマイクロホン間隔、 c は音速である。

(d) 指向性 $D(\theta)$ 上の最小値をとる方向を探索し、音源方向とする。

3.3 適応フィルタのアルゴリズム

追尾処理は、ビームフォーマの入力方向と目的音の到来方向とのずれには対処できるが、音声の場合は音源が理想的な点音源とはみなせないため、仮定した伝搬モデルからのずれによる目的信号除去が生ずる。

そこで、GJGSC の適応フィルタとして、ビームフォーマの入力パワーと出力パワーの和に基づいた正規化 [6] によって修正した学習同定法を用いることとした。この場合、適応フィルタの更新は k を離散時間として次式で表される。

$$W_k = W_{k-1} + \frac{2\mu e_k}{\alpha e_k^2 + (1-\alpha)|X_k|^2} \quad (7)$$

$$e_k = d_k - W_k^T X_k \quad (8)$$

ここで、 $W_k = \{w_k(0), w_k(1), \dots, w_k(N-1)\}^T$ はフィルタ係数、 $X_k = \{x_k, x_{k-1}, \dots, x_{k-N+1}\}^T$ は参照信号ベクトル、 d_k は希望応答、 μ はステップサイズ、 α は正規化パワー中の出力パワーに関する重みである。

3.4 ビームフォーマの入力方向の更新

ビームフォーマの入力方向の設定は、 0° 方向から信号が到来したのと等価になるように入力信号を遅延させることにより行う。遅延を与えるフィルタは、sinc 関数を有限で打ち切った FIR フィルタ [12, pp.215-216] を用いた。

音源追尾の際は、雑音の影響を軽減するため、次式に示すように、現在の瞬時の推定値にビームフォーマの出力パワーに基づく重みをかけて過去の推定値と

の間で平滑化し、現在時刻の音源方向とするようにした。

$$\theta_m(n) = \theta_m(n-1) * \beta + \phi_m(n) * (1-\beta)$$

$$\theta_s(n) = \theta_s(n-1) * \beta + \phi_s(n) * (1-\beta) \quad (9)$$

ここで、 $\theta_m(n)$ 、 $\theta_s(n)$ は各々主 BF、副 BF の入力方向、 β は忘却係数、 ϕ_m 、 ϕ_s は各々、各時刻で推定された目的音、妨害音の方向である。

4. 計算機シミュレーション

以上に述べたマイクロホンアレイの有効性を確かめるため、計算機シミュレーションを行った。シミュレーションでは、図 2 に示すマイクロホン配置において、マイクロホン間隔 d を 16 cm とし、サンプリング周波数は 11 kHz、適応フィルタのフィルタ長は 50 点とした。マイクロホンの信号は、先に述べた sinc 関数のフィルタにより、音源方向に対応した遅延を与えて生成し、背景雑音としてチャンネル間に無相関な白色雑音を各マイクロホン信号に加えた。

また、予備実験により、適応フィルタのステップサイズ $\mu = 0.1$ 、正規化パワー中の出力パワーの割合 $\alpha = 0.9$ 、ビームフォーマの入力方向更新の際の忘却係数 $\beta = 0.1$ とした。話者追尾の場合は、音源推定処理を 64 点おきに行い、そのつど主 BF と副 BF の入力方向を式 (9) により再設定した。方向推定の際の指向性は、 1° おきに値をずらしながら計算した。

4.1 指向性マイクロホンによるエリアシング軽減の効果

まず、2. に述べた指向性マイクロホンの効果を確かめるため、図 2 に示したマイクロホン配置において、マイクロホンの指向性の軸の向きを変えながらビームフォーマ収束後の妨害音方向の周波数特性を調べた。マイクロホンは単一指向性を仮定し、指向性の軸の向き ϕ を 0° 、 30° 、 60° 、 90° とした。なお、単一指向性を表す式としては、 θ を指向性の軸から測った角度として $(1+\cos\theta)/2$ を用いた。背景雑音はチャンネル間に無相関な白色雑音とし、この背景雑音に対して 25 dB の白色雑音が 50° の方向から妨害音として到来するものとした。なお、ビームフォーマは通常の GJGSC とした。図 4 にこの結果を示す。

ビームフォーマ収束後は妨害音方向に死角が形成されるため、妨害音方向のビームフォーマの利得は低くなるが、図からわかるように、軸の向きが 0° のときには、2.5 kHz、5 kHz 付近の利得が高くなっており、

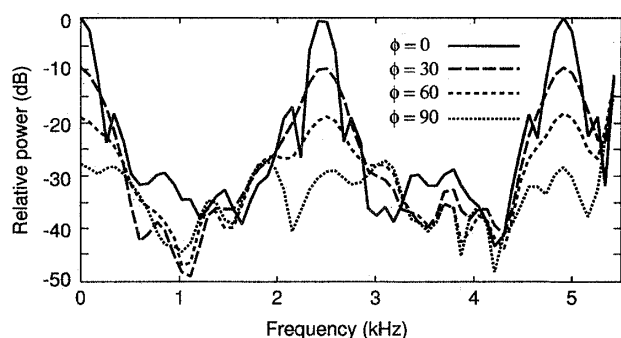


図4 妨害音方向に対するビームフォーマの周波数特性
Fig. 4 Frequency response of the beamformer at interference DOA.

妨害音の除去できない帯域ができていることがわかる。これに対し、指向性の軸を傾けた場合は上記の帯域の利得が減少し、軸の傾きを増すほど妨害音の残留成分が低減していることがわかる。

4.2 収束後の方向対感度特性

適応マイクロホンアレーは、状況によって指向性を変えるため、特定の場面設定に対する指向性だけでは全体的な特性を把握しにくい。様々な妨害音方向に対応した特性を調べるため、1個の音源を固定し、適応させて収束後のパワーを求め、この処理を様々な音源方向に関して行い、音源方向とパワーの関係を求めた。

シミュレーションは、マイクロホンが無指向性の場合と単一指向性で軸の傾き ϕ を 60° とした場合とに関して、また、音源が白色雑音の場合と音声の場合とに関して各々行った。また、背景雑音に対する音源の SNR は 25 dB とした。結果を図5に示す。

図5において、パワーの値は入力方向 0° に音源があるときの出力パワーで正規化してあり、(A), (B) は音源が白色雑音の場合、(C), (D) は音源が音声の場合である。また、各々の図において、曲線 a は話者追尾、学習同定法の修正ともしなし、曲線 b は話者追尾なしで学習同定法の修正あり、曲線 c は追尾範囲 $\phi = \pm 20^\circ$ で学習同定法の修正あり、曲線 d は追尾範囲 $\phi = \pm 40^\circ$ で学習同定法の修正あり、の場合の結果である。

図からわかるように、追尾処理、学習同定法の修正ともしなし (a) の場合は到来方向のずれに対する許容範囲が非常に狭いのにに対し、追尾なしで修正ありの場合 (b)、ずれの許容範囲が $\pm 8^\circ$ 程度に広がる。追尾する (c), (d) の場合は、設定した追尾範囲にわたって感度が一定となり、ずれの許容範囲を任意に設定で

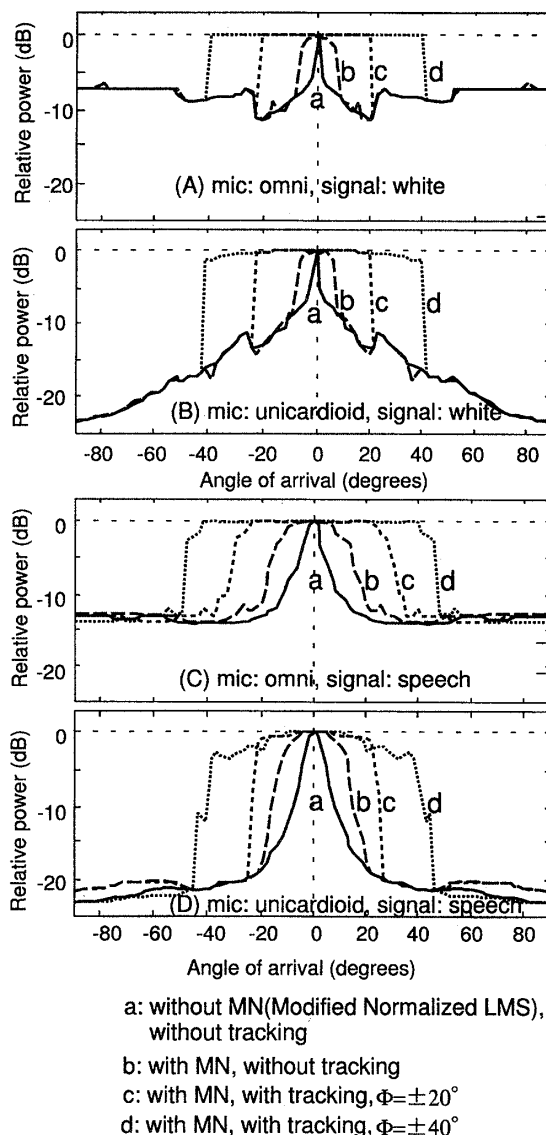


図5 収束後の方向対感度特性
Fig. 5 Output power after convergence vs. DOA.

きることをわかる。

また、音源が白色雑音の場合と音声の場合とを比較すると、音源が音声の場合の方が追尾範囲外の感度を若干低く抑えることができるが、白色雑音の場合の方が追尾範囲の境界が急峻である。

また、(A)と(B), (C)と(D)を比較すると、指向性マイクロホンを用いた場合の方が追尾範囲外の感度が 10 dB 以上低く、提案したマイクロホン設定の有効性が示されている。

4.3 目的音源追尾

前節は固定した音源に対する特性であったが、移動する音源に対する追尾性能を調べるため、音源を移動させて追尾軌跡を求めた。音源は図6に示す配置と

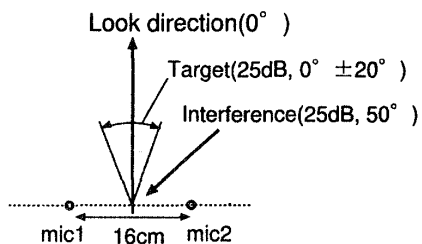


図6 音源配置
Fig. 6 Sound source arrangement.

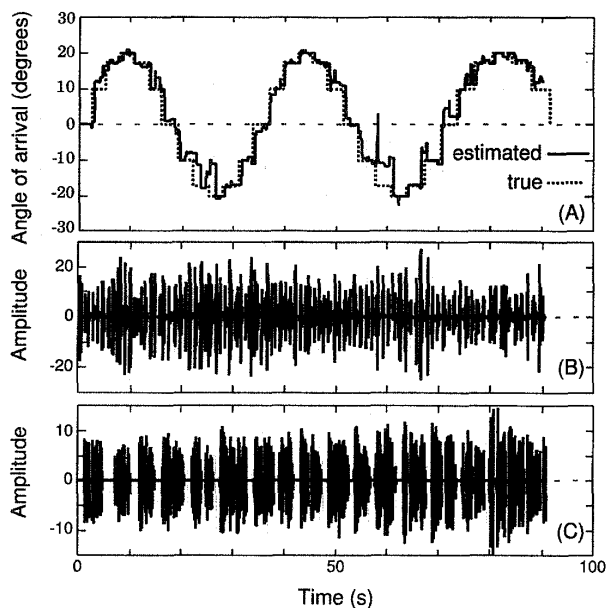


図7 目的音源追尾結果
Fig. 7 Result of target tracking.

し、目的音源の方向を入力方向 0° の周りに正弦波的に移動させた。

目的音源は単語発声、妨害音源は文章読上げの音声とし、各音源のSNRは、周囲雑音を基準として25dBとした。この場合、目的音源対雑音のトータルのSNRは -1.4 dBとなる。目的音の方向は、3単語の発声区間をセグメントとしてその間方向を固定し、セグメントごとに次式に従って方向を変化させた。

$$\theta = R \sin(kv) \quad (k = 0, 1, 2, \dots) \quad (10)$$

ここで、 k はセグメント番号、 v は方向移動の角速度、 R は方向変化の最大振幅であり、 $v = 30^\circ$ 、 $R = 20^\circ$ とした。

音源方向追尾結果を図7(A)に、目的音源と妨害音源の波形を図7(B)、(C)に示す。図7(A)において、実線は推定された音源の軌跡、点線は真の軌跡である。

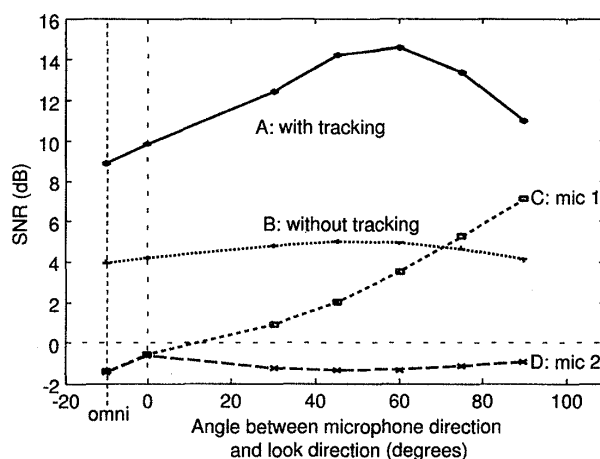


図8 移動音源の信号抽出精度
Fig. 8 Result of target signal extraction.

図からわかるように、追尾誤差は平均 2° 程度であり、前節でみたように、適応速度制御によって許容される範囲が $\pm 8^\circ$ 程度であることを考慮すると、追尾誤差による目的音除去は生じないことがわかる。

4.4 目的音抽出

最後に、目的音源が先と同じ条件で移動する場合についてマイクロホンの向き ϕ を 0° から 90° の間で変えながら目的音声の抽出精度について調べた。結果を図8に示す。

図8において、曲線(A)は追尾範囲を $\pm 30^\circ$ とし、曲線(B)は追尾しない場合に得られた結果であり、値は出力の目的音の原信号に対する精度である。図からわかるように、追尾しない場合は、精度が5dB以下であるのに対し、追尾した場合は9dB以上であり、 $\phi = 60^\circ$ のときに最大値14.7dBが得られている。

5. むすび

PCや携帯器など、マイクロホン数が2個に限られた場合でもロバストな処理ができる適応マイクロホンアレイを開発した。計算機シミュレーションによる検討の結果、二つのビームフォーマを用いた音源追尾と指向性の軸を傾けて配置した指向性マイクロホンの組合せにより、広い角度からの到来音を高いSNRで入力できることを確認し、従来雑音抑圧効果がほとんどなかった移動音源からの抽出性能を大幅に改善できた。従来の適応マイクロホンアレイでは、使用の際、話者位置を固定しなければならなかったが、このマイクロホンアレイでは話者位置の自由度を設定できるため使い勝手の向上が期待できる。今後、更に高雑音の環境

にも対応できるよう多チャンネル化を検討している。

謝辞 本研究の機会と有益な助言を与えてくださった株式会社東芝関西研究所の方々に感謝致します。

文 献

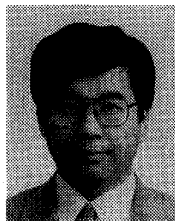
- [1] B. Widrow, P.E. Mantey, L.J. Griffiths, and B.B. Goode, "Adaptive antenna system," Proc. IEEE, vol.55, no.12, pp.2143-2161, Dec. 1967.
- [2] O.L. Frost, III, "An algorithm for linearly constrained adaptive array processing," Proc. IEEE, vol.60, no.8, pp.926-935, Aug. 1972.
- [3] K. Takao, M. Fujita, and T. Nishi, "An adaptive antenna array under directional constraint," IEEE Trans. Antennas & Propag., vol.AP-24, no.5, pp.662-669, Sept. 1976.
- [4] L.J. Griffiths and C.W. Jim, "An alternative approach to linearly constrained adaptive beamforming," IEEE Trans. Antennas & Propag., vol.AP-30, no.1, pp.27-34, Jan. 1982.
- [5] Y. Kaneda and J. Ohga, "Adaptive microphone-array system for noise reduction," IEEE Trans. Acoust., Speech & Signal Process., vol.ASSP-34, pp.1391-1400, Dec. 1986.
- [6] J.E. Greenberg and P.M. Zureck, "Evaluation of an adaptive beamforming method for hearing aids," J.A.S.A., vol.91, no.3, pp.1662-1676, March 1992.
- [7] S. Oh, V. Viswanathan, and P. Papamichalis, "Hands-free voice communication in an automobile with a microphone array," IEEE Proc. of ICCASP'92, pp.I-281-284, 1992.
- [8] K.M. Buckley and L.J. Griffith, "An adaptive generalized sidelobe canceller with derivative constraints," IEEE Trans. Antennas & Propag., vol.AP-34, pp.311-319, March 1986.
- [9] S. Affes and Y. Grenier, "A signal subspace tracking algorithm for microphone array processing of speech," IEEE Trans. Speech & Audio Process., vol.5, no.5, pp.425-437, Sept. 1997.
- [10] S. Haykin, "Adaptive filter theory," Prentice-Hall, 1996.
- [11] 宝珠山治, 杉山昭彦, "ブロッキング行列にリーク適応フィルタを用いたロバスト一般化サイドローブキャンセラ," 信学論 (A), vol.J79-A, no.9, pp.1516-1524, Sept. 1996.
- [12] 大賀寿郎, 山崎芳男, 金田 豊, "音響システムとデジタル処理," 電子情報通信学会, 1995.

(平成 10 年 9 月 18 日受付, 12 月 9 日再受付)



永田 仁史 (正員)

昭 59 東北大・工・電子卒。平 2 同大大学院情報工学専攻博士課程了。工博。同年(株)東芝入社, 研究開発センター勤務。平 6 同社関西研究所, 平 9 岩手大学工学部講師。音声認識, デジタル音響信号処理, 並列計算機用コンパイラの研究に従事。日本音響学会, 情報処理学会各会員。



安倍 正人 (正員)

昭 56 東北大大学院電気及び通信工学専攻博士課程了。工博。昭 58 東北大学情報処理教育センター助手。平 1 東北大学大型計算機センター助教授。平 8 岩手大学工学部情報工学科教授。デジタル信号処理, 並列計算機アーキテクチャに関する研究に従事。IEEE, ACM, 米国音響学会, 日本音響学会, 日本騒音制御学会, 日本機械学会各会員。