

音源定位・音源分離に関する研究成果

安藤 繁, 篠田裕之, 小野順貴

情報理工学系研究科システム情報学専攻

Abstract — 本グループでは、実世界で活用可能な音源定位センサと音源分離システムの実現を目指す。すなわち、視覚センサと協調し、三次元環境の視覚的・音響的变化を高感度かつ高安定に検出し、その環境中で行動するロボットや環境自体の目と耳となる人型センサの実用機を開発する。また、音源定位センサをマイクロ化し、それらをネットワークを介して連携させることにより、広がりをもつ音環境を面的・並列的に把握するセンサネットワークシステムの基礎技術を確立する。これにより、三次元環境中での音環境理解と選択的聴取、すなわち「カクテルパーティー効果」の工学的実現を目指す。

ステレオやエンジン音に置き換えれば自動車内の音環境に、サイレンや警報音や燃え盛る燃焼音に置き換えれば火災現場の音環境に、そのまま置き換わり、それぞれに対する関心事も、環境安全、悪環境でのコミュニケーション、災害救助などに容易に置き換えることができる。このようなことから、上記の複雑に混合した音環境のセンシング能力、複雑に混合した音環境における対象音の選択能力の実現は、多くの分野の技術に大きな革新をもたらすことは明らかである。

以下では、それぞれの要素技術ごとに、本年度の主要な研究成果をまとめる。

1 はじめに

立食式パーティーのように、大きな部屋に多数の人々が移動しながら立ち話をするような状況を想定する。本人は多くの人たちととりとめない話をしながらも、自分の上司がどのような人とどんな話をしているかを気にかけている。これから始まる重要なプロジェクトがどのようなものか、自分がその主要メンバーになれるかが重大な関心事だからである。このような音環境は、一般に

- 1) 吸収性だが反射は無視できない天井や壁
 - 2) 再帰反射や独特の周波数特性をもつ室内形状
 - 3) 外界との音の出入りや外乱音を許す窓の存在
 - 4) バンド演奏や館内放送のような大きな環境音などからなり、大変に複雑な要因がからみあう。
- しかし「カクテルパーティー効果」と呼ばれるように、人間は非常にたやすくこのような音環境の把握と選択的聴取を行うことができる。上記の環境は、バンド演奏や館内放送を広告の音楽や雑踏や自動車騒音に置き換えれば街中の音環境に、カー

2 時空間勾配法による音源定位センサ

2.1 四耳型音源定位システムの開発

音源定位とは、生物が音源の方位や距離を知覚する能力のことである。この能力は瞬時的なものと経時的なものとに二つに分けて考えられる。瞬時的な能力では、人間は両耳時間差と両耳強度差を手がかりとして、単一音源ならば数 10ms の時間で 1 度以下という優れた定位能力をもつ。経時的な能力では、カクテルパーティー効果として知られるように、多数の分布した音源の空間配置を知覚したりそれぞれを聞き分けることができる。

人間の耳をセンサとして見たときの第一の特色は、その配置の対称性と、対象とする音波の波長に対してその間隔がはるかに小さいという点である。これは対称性と差動構造、測定対象との一体化を重視するセンサの構造とよく共通点をもつ。物理場にセンサを一体化させる考え方を音場に対して適用して実現したのが、図 1 に示す視覚聴覚融合人型センサ [5]–[8] である。このシステムは、



(a) SmartHead Boy



(b) 音源定位視線移動システム SL-3000

図 1. 視覚聴覚融合型人型センサとして研究室独自に本年度新たに開発した”SmartHead Boy” と、音源定位と視線移動機能に絞って実用化を目指し、昨年度に共同開発した SL-3000。単チャンネルのアナログ信号処理だった従来機に対して複数帯域分割の全デジタル処理化され、騒音や多重反射環境における定位性能の改善が図られている。

5ms の時間窓で近接 4 点音場の時空間勾配を観測し、波面の法線方向から音源の方位を、法線方向での波の振幅減衰率から音源までの距離を検出する。これらの結果を、画面上と計算機のメモリ上に、単一音源か 2 音源かの分類や音源定位の信頼度とともに累積する。複数音源における音場の時間空間勾配に関する確率モデルに従い、パルゼン推定とノンパラメトリックカルマンフィルタによって上位システムがその注意を集中すべき経時的に最も明瞭な音源を選択する。

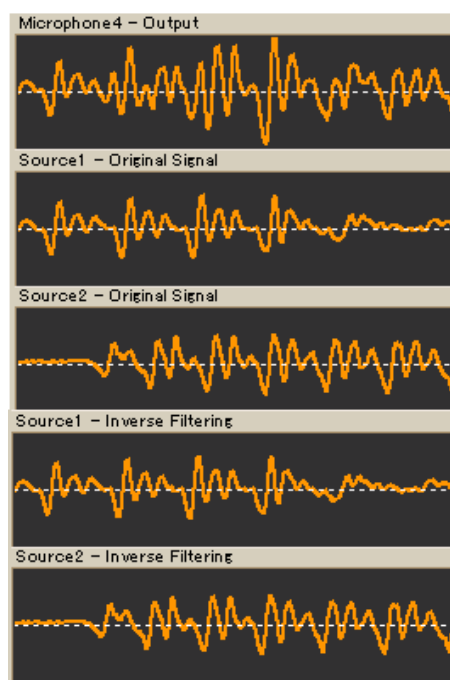


図 2. 音源定位に基づく音源分離の原理確認実験結果。上から、マイクロフォン採取音（2 音源の混合音）、2 音源それぞれの音源音、研究手法によるそれぞれの分離音を示す。

2.2 音源定位に基づく音源分離

本年度より新たに実施した研究テーマである。音源定位により、複数の音源について、それらを追従しつつ方位と距離の情報が得られる。この中から着目する 1 個の音源と、排除したい 1 個の音源を選択し、それらの方向と距離の情報を与えて、一方を他方から分離する最適フィルタを設計するというのがこの原理である。

従来は、音源の定位の段階と分離の段階を明瞭に分けた研究は行われていなかった。着目音と妨害音の波形を与えて、それらを分離する最適フィルタを学習的最小二乗法により漸近的に構成するというアプローチが主要であった。この方法は、収束まで長時間（数秒程度）を要し、音源の移動に対応が難しこと、分離後の音響信号に特有な周波数特性を生じさせ、その後の音声認識の精度を大幅に低下させるなどの問題をかかえていた。本方法は、瞬時的および中期的な音源定位・音源分布形成の段階と、音源の移動に追従できるだけの比較的長期的な音源分離の段階とを分離させるこ

とにより，環境の変化に適応しやすく，上位システムによる機動的な音源選択が可能な実用的なシステムを可能にする。それと同時に，音源定位で用いる近接四点マイクロフォンを音源分離に同時に利用するため，分離された音の周波数特性が広い範囲でフラットにできるため，後段の音声認識の性能向上も可能にすると期待される。現在は，理論の構築をほぼ終了し，図2の示すように，数値実験による諸性能の確認と，実験室において採取した音データを用いたオフラインの音源定位・音源分離実験の段階にある。

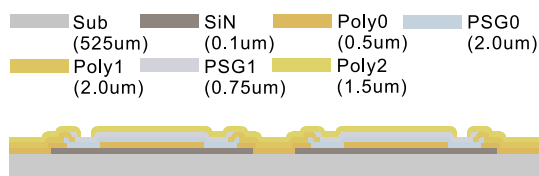


図3. Poly-Si₃N₄層，PSG2層のプロセス用に設計したセンサ構造。振動板と対向電極間は2.75μm。

3 マイクロ音源定位センサ

本研究の目的は，非常に小さいながら鋭い方向感覚を有するヤドリバエの鼓膜を模倣することにより，従来のマイクロフォンアレイ技術では困難であったマイクロ音源定位センサを実現することである。昨年度は，ヤドリバエの鼓膜をモデル化し，これを2次元的に拡張したジンバル型振動板を有するセンサを開発している。

- 1) ジンバル構造振動板，2×2に分割された対向電極，構造を保つための周辺ビーム等を有する静電型センサ構造を設計した。(図3)
- 2) Si表面微細加工のマルチチップサービスを利用し，直径1~3mmのセンサを試作した。(図4)
- 3) 試作したセンサで，ジンバル構造体が形成されていることを確認した。(図5)

REFERENCES

[1] S. Ando and A. Kimachi, IEEE Trans. Electron Devices, vol.50, no.10, pp.2059-2066, 2003.

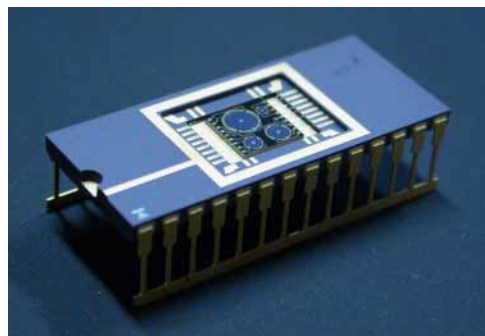


図4. Si表面微細加工で試作したセンサのパッケージ後の写真。サイズの異なる4つのセンサを同一チップ内に試作した。チップサイズは6.8mm×6.8mm。

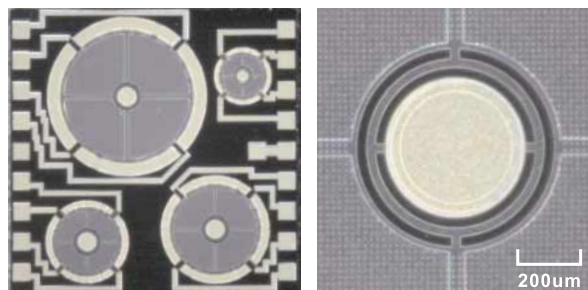


図5. 試作したセンサのチップ写真。全体図(左)。ジンバル部拡大図(右)。

[2] Takaaki Nara and Shigeru Ando, Inverse Problems, vol.19, no.2, pp.355-370, 2003.

[3] N. Ono, A. Saito, and S. Ando, Trans. IEEJ, vol.123-E, no.3, pp.92-97, 2003.

[4] N. Ono *et al.*, Proc. 12th Int. Conf. Solid State Sensors and Actuators, Boston, 2003.

[5] 安藤, 篠田, 小川, 光山, 計測自動制御学会論文集, vol.29, no.5, pp.520-528, 1993.

[6] S. Ando, Proc. Int. Workshop Machine Vision Applications, Kawasaki, pp.417-422, 1994.

[7] S. Ando, Trans.IECE on Information and Systems, vol.E78-D, pp.1621-1629, 1995.

[8] S. Ando, Sensors and Materials, vol.7, no.3, pp.213-231, 1995.

[9] S. Ando and H. Shinoda, IEEE Control Systems, vol.15, no.1, pp.61-69, 1995.

[10] S. Ando *et al.*, IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control, vol.48, no.4, pp.1031-1045, 2001.

[11] N. Ono *et al.*, Tech. Digest 19th Sensor Symp., pp.351-354, Kawasaki, May 2001.