

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 3 日現在

機関番号：13501

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24560519

研究課題名(和文) ソフトコンピューティング技術を駆使した超指向性マイクロホンアレイシステムの構築

研究課題名(英文) Development of a superdirective microphone-array system based on the soft computing technology

研究代表者

小澤 賢司 (OZAWA, Kenji)

山梨大学・総合研究部・教授

研究者番号：30204192

交付決定額(研究期間全体)：(直接経費) 4,200,000円

研究成果の概要(和文)：本研究は、音響信号の空間分離収録技術に関して、ソフトコンピューティングの技法を駆使することにより格段の性能向上を図ることを目的とした。

まず、ニューラルネットワークを組み込んだマイクロホンアレイシステムについて、システムのパラメータおよびニューラルネットワークの学習パラメータを遺伝的アルゴリズムを用いて決定することで最適化を達成した。次に、複数周波数について指向性を獲得させるようシステムを拡張し、種々の周波数の正弦波に対して超指向性を達成した。さらに、入力信号である広帯域音を高速フーリエ変換を利用して正弦波に分解し、それらを並列処理した後で重ね合わせることで広帯域化を実現した。

研究成果の概要(英文)：This study deals with the technology of spatial recording of acoustical signals and aims to improve the spatial resolution by introducing the soft computing technology.

The optimization of a neural-network-based microphone-array system was achieved by determining the system parameters and the neural-network-learning parameters using a genetic algorithm. Then the system was expanded to learn super-directivities for multiple frequencies. As a result, the super-directivities for sinusoidal signals with various frequencies were yielded. Finally, we developed a wideband system in which an input complex tone is decomposed into sinusoidal components using the fast Fourier transform, and processed in parallel with independent array-processing units.

研究分野：音響情報工学

キーワード：マイクロホンアレイ ニューラルネットワーク 遺伝的アルゴリズム ソフトコンピューティング 指向性 時空間音圧分布 GPU 最適化

1. 研究開始当初の背景

ニューラルネットワーク (以下 NN と略記) を利用した非線形マイクロホンアレイが提案され、小規模ながら高い性能が得られることが知られていた。そのうち、森田ら<引用文献①>が提案した差分型システムは、所望の信号は非線形系を通さないという工夫によって非線形歪の発生が抑えられるので有望である。しかし、<引用文献①>では NN のパラメータについて言及されておらず、そこに示されたシステム特性が最適化されたものであるかは不明であった。さらに、<引用文献①>のシステムでは、指向特性が保証されるのは NN の学習に用いた周波数の正弦波に限定されるという課題が残されていた。

2. 研究の目的

本研究では、以下の3点を目的とした。

- (1) まずシステムの指向特性を最適化することを目的とし、その最適特性を達成するための NN の学習パラメータを遺伝的アルゴリズム (以下 GA と略記) を利用して効率よく探索する。
- (2) また、NN の学習に用いていない種々の周波数の正弦波についても鋭い指向特性を得るようにシステムを拡張する。
- (3) さらに、音声などの広帯域音にも適用できるように、システムの広帯域化を実現する。

3. 研究の方法

(1) 本研究で検討対象とした差分型システム<引用文献①>の構造を図1に示す。N個のマイクロホン直線状に配置したアレイ延長線上に Reference Microphone (以下 RM と略記) が設置されている。アレイの鉛直方向を指向方向 0° とすると、指向方向から入射された音 (目的音) は各マイクロホンに同時に到達するため、RM の出力を引き去れば信号成分は相殺され、NN には入力されない。NN は RM における 0° 以外からの到来音 (雑音成分) を想起再生するように学習させ、NN からの出力を RM 出力から引き去ることで、システム全体としては雑音抑制を行う。

以上を原理とするマイクロホンアレイについて、GA を用いた最適化を図ることとした。システムの構成は<引用文献①>に従い、マ

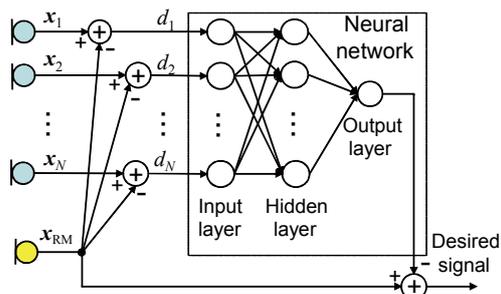


図1 従来のマイクロホンアレイ<引用文献①>のブロック図

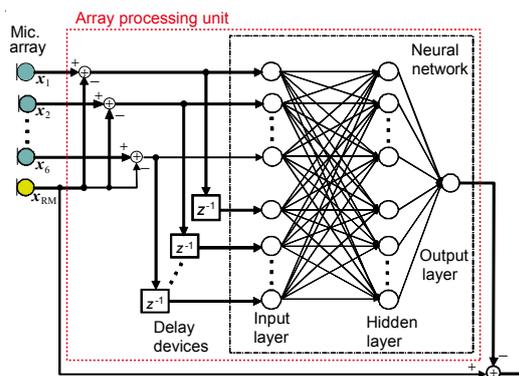


図2 提案したマイクロホンアレイのブロック図

イクロホンの間隔は 10 cm で、その数は $N=6$ とした。それゆえ、3層ニューラルネットワークの入力層のユニット数は6である。中間層と出力層のユニット数は、それぞれ6と1とした。各ニューロンからの出力はシグモイド関数により決定した。学習条件も<引用文献①>と同様であり、対象は周波数 1.7 kHz の正弦波とした。また、学習は単一音源に対して行うこととし、システムへの入射方向は、 $0^\circ \sim 90^\circ$ の間で 5° 間隔とした。学習には逆誤差伝搬法を用いた。

NN の学習を行う際には、複数のパラメータを設定する必要がある。そして、そのパラメータの設定により、学習結果であるシステムの特性が変化する。そこで、本研究では、NN 学習時のパラメータを変えることによって指向特性の最適化を図ることとした。最適なパラメータを効率よく探索するために、GA を用いることとした。

検討対象としたパラメータは、①逆誤差伝搬法における学習定数、②入力層-中間層の結合荷重の初期値が取り得る範囲、③中間層-出力層の結合荷重の初期値が取り得る範囲、④バイアスの初期値が取り得る範囲、⑤シグモイド関数の温度係数、⑥学習の収束判定を行うための二乗誤差値という6種類である。GA における個体を構成する各遺伝子は上記パラメータの値を並べた実数値コーディングとし、1世代の個体数を100、世代数の上限を100と設定した。最適性の指標となる評価関数としては、「入射角 0° についての出力パワー」/「入射角 $5^\circ \sim 90^\circ$ についての出力パワーの和」という比を用いた。これにより、メインロブの幅が 5° 未満という鋭い指向性を得ることを目標とした。

(2) 上記(1)の最適化がなされた場合であっても、NN の学習が単一周波数で行われているため、学習に用いていない周波数の入力については特性が劣化することは明らかである。この問題を解決するためには、システムの構成を見直すことが必要であると考えた。すなわち、従来のシステムは NN への入力パターンが「ある時刻においてマイクロホンアレイにより観測された音圧の空間分布」であり、これが問題であると考えた。そこで、図2に

示すように、観測信号を1サンプル遅延させた空間パターンも入力とすることで、「音圧の時空間分布」をNNへの入力パターンとするようシステムを拡張することとした。これにより、入力信号の周波数が異なることをNNの学習に反映させることが可能であるので、複数周波数パターンを同時に学習させることができる。

図2のシステムについて、(1)と同様にGAを利用した最適化を行った。学習条件等は基本的に(1)と同様であるが、ここでは中間層のユニット数というシステムのパラメータも最適化の対象に含めることとした。なお、学習対象としては(1)で用いた1700 Hzに加えて、850 Hz, 425 Hzも対象として3周波数の正弦波を同時に扱うことで広帯域化を図った。

(3) 上記(2)の広帯域化は、入力信号が種々の周波数を有する場合に有効であるが、ただし正弦波に代表されるような狭帯域信号に限定されるという課題がある。そこで、図3に示すように、広帯域な入力信号を狭帯域成分に分解した上で並列処理することによりシステムの広帯域化を図ることとした。帯域通過フィルタとしては、1/3オクターブ帯域通過フィルタを用いた場合と、高速フーリエ変換(FFT)を用いて正弦波成分にまで分解する場合の2通りを取り上げた。

なお、(2)のシステムでは、マイクロホンの間隔を10 cmとしていたため、空間の標本化定理によって1.7 kHzを上限とするという制約があった。そこで、広帯域化のために、マイクロホン間隔を100 cm, 10 cm, 1 cmとした3種のアレイを準備し、それぞれ170 Hz以下/170~1700 Hz/1.7~17 kHz以上の3帯域に分けて信号を取り出すアレイを切替えることとした。ただし、NNにおける結合荷重は共通のものを用いることとした。

ここで提案したシステムでは、分解された周波数成分ごとに独立してアレイ信号処理を行うため、並列処理が可能である。そこで、並列処理を実行するのに好適なアーキテク

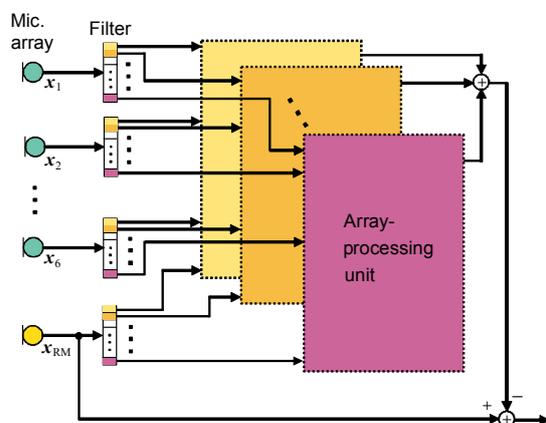


図3 アレイ信号処理の並列化により広帯域化したマイクロホンアレイのブロック図

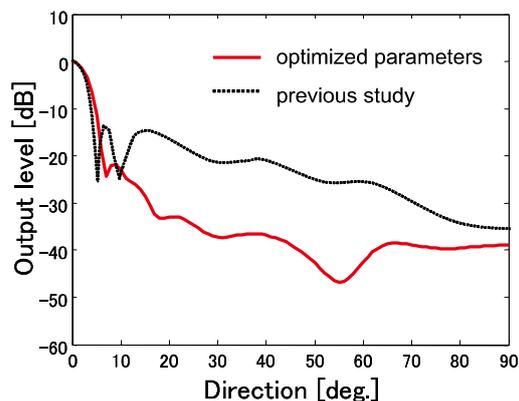


図4 1.7 kHz 正弦波入力に対する指向特性

チャを有する画像信号処理用デバイスであるGPU(Graphics Processing Unit)を用いて実装することで高速化を図ることとした。本研究では、GPUとしてNVIDIA社製GeForce GTX 580を用いた。

4. 研究成果

(1) GAを利用して探索した結果である最適パラメータを適用して学習したNNを用いたシステムについて、1.7 kHz正弦波に対する指向特性を、<引用文献①>の結果と比べて図4に示す。図から、本研究の最適化によってメインロブの特性をほとんど変えることなく、サイドロブが大幅に低減されていることが分かる。

学習に用いた正弦波以外の信号に対する指向性を検討するため、<引用文献①>と同じAM音、FM音を信号として指向性を観測した。AM音(搬送波周波数:1.7 kHz, 信号周波数:30 Hz, 変調度:15%)については、<引用文献①>よりサイドロブについて5 dB程度の改善が見られた。ただし、学習に用いた1.7 kHzから外れた周波数の入力に対しては指向性が著しく劣化するという問題が残った。

(2) 上記の問題を解決するために提案した図2のシステムについて、学習に用いた3種の正弦波に対する指向特性を図5に示す。いずれの周波数についても、一様な鋭いメインロブを確保し、指向方向外である5~90°については出力が抑圧されていることが分かる。その抑圧量は周波数が高いほど大きい傾向があるが、これは周波数が高いほど波長が短いため時空間パターンにおける音圧の変化が大きいことに因るものと考えた。

学習に用いた周波数以外の正弦波信号に対する指向性を検討するため、20~1700 Hzの範囲で10 Hzごとに指向性を調べた結果を図6に示す。図から分かるように、400 Hz以上の周波数については、学習に用いた周波数以外でも鋭い指向特性が得られている。また、指向方向以外の抑圧量は、やはり周波数が高いほど大きい傾向が見られる。

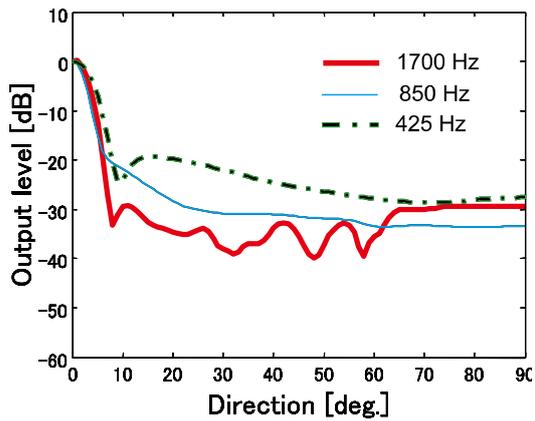


図 5 NN の学習に用いた 3 周波数の正弦波入力に対する指向特性

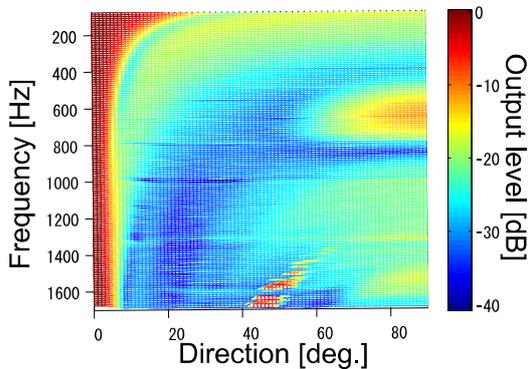


図 6 NN の学習に用いていない周波数の正弦波入力に対する指向特性

次に、搬送波周波数 850 Hz、信号周波数 30 Hz、変調度 15%の AM 音、FM 音を信号として狭帯域音に対する指向性を観測した。その結果、図 5 の 850 Hz 正弦波の場合に比べて若干の抑制量の減少はあるが、同様に鋭い指向性が得られた。ただし、広帯域な入力に対しては指向性が著しく劣化した。

(3) 上記を踏まえて広帯域化を図った図 3 のシステムについて、まず 1/3 オクターブ帯域通過フィルタを用いてシステムを実装した。これにより、高調波複合音のように周波数成分が疎な広帯域音については十分な指向性を得ることができたが、白色雑音のような広帯域音については全く指向性を得ることができなかった。

次に、入力信号を 1024 ポイントのブロックごとに FFT し、各周波数成分ごとに逆 FFT を行うことで正弦波成分に分解した時間波形を各ユニットで処理させた。17 kHz までに帯域制限した白色雑音を入力信号として用いた場合の指向性を図 7 に示す。10° 以上の角度では、約 20 dB の減衰となっており、広帯域音についても指向性を獲得することに成功した。

GPU 上に実装したシステムに、実際に 2 秒の音データを入力し処理するまでの実行時間を測定した。その結果、CPU (Intel core i7

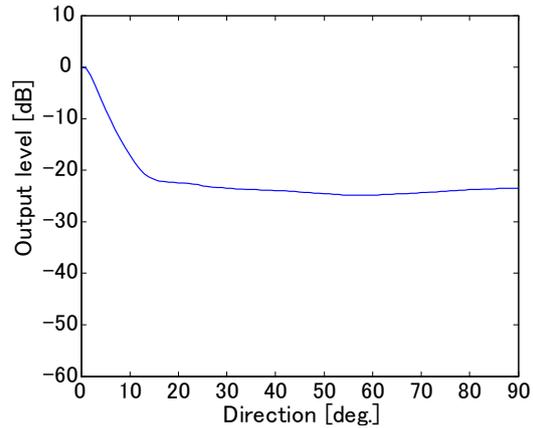


図 7 白色雑音を入力とした場合の指向特性

960) のみを用いて処理した場合には 174 秒であったのに対して、GPU を利用した場合には 28 秒であり約 6 倍の高速化を実現した。ただし、実時間処理には到達していないので、今後の GPU の性能向上を期待するとともにプログラミングの工夫も必要である。

以上、本研究では、3 年間にわたり着実に研究を進め、当初に掲げた 3 点の目的を達成したものと考えている。

<引用文献>

- ① 森田亘、小畑秀文、非線形多層構造を持つ超指向性マイクロホン・アレイ・システム、日本音響学会誌、49 巻、1992、28-33

5. 主な発表論文等

[雑誌論文] (計 1 件)

- ① ISEKI, Akihiro, KINOSHITA, Yuichiro, OZAWA, Kenji, Optimization of neural-network-based superdirective microphone-array system using a genetic algorithm, Acoustical Science and Technology, 査読有、36 巻、2015、掲載決定、<https://www.jstage.jst.go.jp/browse/ast>

[学会発表] (計 5 件)

- ① 秋鹿勇祐、小澤賢司、井関晃広、木下雄一朗、森勢将雅、周波数成分の分解による非線形マイクロホンアレイの広帯域化、日本音響学会 2015 年春季研究発表会、2015 年 3 月 16 日、中央大学 (東京)
- ② ISEKI, Akihiro, OZAWA, Kenji, KINOSHITA, Yuichiro, Neural network-based microphone array learning of temporal-spatial patterns of input signals, The 2014 IEEE 3rd Global Conference on Consumer Electronics (GCCE 2014)、2014 年 10 月 7 日、幕張メッセ (千葉)

- ③ OZAWA, Kenji、ISEKI, Akihiro、KINOSHITA, Yuichiro、Genetic algorithm optimization applied to a neural-network-based super-directive microphone array system、The 1st International Conference on Engineering and Applied Sciences Optimization (OPT-i 2014)、2014年6月5日、Kos (Greece)
- ④ 井関晃広、小澤賢司、木下雄一郎、入力信号の時空間パターンの学習による非線形マイクロホンアレイの特性改善、日本音響学会2014年春季研究発表会、2014年3月10日、日本大学(東京)
- ⑤ 井関晃広、小澤賢司、木下雄一郎、非線形マイクロホンアレイにおける指向特性の最適化、日本音響学会2013年春季研究発表会、2013年3月13日、東京工科大学(東京)

[その他]

ホームページ(発表文献の書誌情報リスト)
<http://www.ccn.yamanashi.ac.jp/~ozawa/>

6. 研究組織

(1) 研究代表者

小澤 賢司 (OZAWA, Kenji)
山梨大学・大学院総合研究部・教授
研究者番号：30204192

(2) 研究分担者

木下 雄一郎 (KINOSHITA, Yuichiro)
山梨大学・大学院総合研究部・准教授
研究者番号：70452133

(3) 連携研究者

なし

(4) 研究協力者

井関 晃広 (ISEKI, Akihiro)
秋鹿 勇祐 (AKISHIKA, Yusuke)
森勢 将雅 (MORISE, Masanori)