

MUSIC 法に基づく空中音響センシング手法の検討

Study of airborne acoustic sensing based on MUSIC method

干場功太郎^{*1}, 土屋健伸^{*1}

Kotaro HOSHIBA^{*1}, Takenobu TSUCHIYA^{*1}

神奈川大学 工学部 電気電子情報工学科^{*1}

Department of Electrical, Electronics and Information Engineering

Faculty of Engineering, Kanagawa University^{*1}

{hoshiba, kenshin}@kanagawa-u.ac.jp

Abstract

本稿では、ロボットなどの自律動作などのための周辺環境認識を目的に、空中音響センシング手法についての検討を行った。従来手法である合成開口法は、機器を小型化、低周波数化すると分解能の高い計測ができないという問題点があった。そこで、高い分解能で音源の到来方向を検出することのできる MUSIC (Multiple Signal Classification) 法に基づく、MUSIC-SONAR (MUSIC based Sound Navigation and Ranging) を提案した。MUSIC-SONAR では、MUSIC 法を時間方向の分解能を確保できるよう拡張し、ターゲットからの反射波の到来方向および伝搬時間を検出し、ターゲットの位置を特定する。計算機シミュレーションによる評価を行った結果、合成開口法では計測ができない状況でも、高分解能な計測が可能なのことがわかり、有用性が確認できた。

1 はじめに

現在、ロボットなどの自律動作や車の自動運転のための周辺環境認識を目的に、さまざまな二次元・三次元のリモートセンシング技術の開発が行われている。それらはカメラ・レーザー・赤外線など主に電磁波が用いられている。特に近年では、LIDAR (Light Detection and Ranging) と呼ばれる技術を用いた手法が多く用いられている [1, 2]。しかし、LIDAR の問題点として

- デバイスが高価である
- 光速の信号を処理するため、処理コストが高い
- 指向性の狭いレーザーを用いているため、機械的走査が必要である

といった点が挙げられる。

そこで、音響信号を用いたセンシング手法を用いることを考える。音響信号を用いる場合、

- デバイスを安価に構築することができる
- 光速と比べ音速は遅いため、処理コストが低い
- 容易に指向性を広くすることができるため、機械的走査を行わなくても広範囲を計測することができる

といった特徴があるため、電磁波を用いる手法の問題点を解決できる可能性がある。

1.1 関連研究

音響信号を用いた計測は古くから用いられている。基本原理はパルスエコー法と呼ばれる、ターゲットからの反射波の伝搬時間からターゲットまでの距離を計測する手法である。水中や体内における音響計測では、パルスエコー法を応用し、合成開口法と呼ばれる、複数の送受波器を用いて二次元、三次元計測を行う手法が実用されている [3, 4]。しかし、空中での音響計測では主に単純な一次元の距離計測しか行われていない [5]。その理由の一つとして、空中で合成開口法によるセンシングを行う場合、機器を小型化、低周波数化すると分解能の高い計測ができないという問題点がある。例として、Fig. 1 に合成開口法による計測結果を示す。(a) が直径 0.5 m のセンサ、送信信号の周波数として 50 kHz を用いた場合、(b) が直径 0.1 m のセンサ、送信信号の周波数として 20 kHz を用いた場合の結果である。(a) では各受波器で得られたターゲットからの反射波が合成開口法により 1 点に交わり、ターゲットの検出ができていることがわかる。しかし、ロボットへ搭載するには大きく、また周波数が高く、組み込み化するには処理コストが大きい。そこで小型化、低周波数化すると、(b) のように 1 点で交わらなくなり、分解能が低くなることわかる。

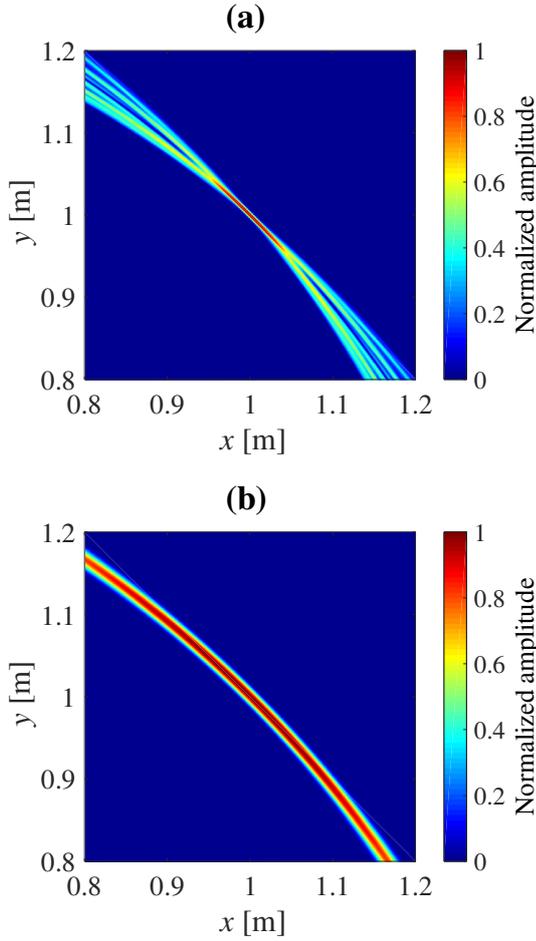


Fig. 1: Sensing by synthetic aperture method using: (a) a sensor at a diameter of 0.5 m and at a frequency of 50 kHz, (b) a sensor at a diameter of 0.1 m and at a frequency of 20 kHz.

1.2 本研究の目的

前節の問題点を解決するため、MUSIC (Multiple Signal Classification) 法 [6] に基づく空中音響センシング手法を提案する。MUSIC 法を用いた音源定位手法では、高い分解能で音源の方向を検出できることがわかっている [7]。また、iGEVD-MUSIC (MUSIC based on incremental generalized eigenvalue decomposition) 法 [8] や、iGSVD-MUSIC (MUSIC based on incremental generalized singular value decomposition) 法 [9] といった、ノイズ耐性の高い手法が提案されているため、ロボット上などノイズが大きい場所でも計測できる可能性がある。ロボット聴覚オープンソースソフトウェア HARK (Honda Research Institute Japan Audition for Robots with Kyoto University)¹ [10] では、すでにリアルタイムで動作する MUSIC 法が実装されており、提案手法の実装も容易であると考えられる。しかし、MUSIC 法では音の到来方向は高分解能で検出で

¹<https://www.hark.jp/>

きるが、伝搬時間方向の分解能が低いといった問題点がある。パルスエコー法では伝搬時間から距離を計測するため、時間方向に高い分解能が必要である。そこで本稿では、時間方向に高い分解能を確保できるように MUSIC 法を拡張した空中音響センシング手法を提案する。計算機シミュレーションにより、提案手法の評価および考察を行う。

2 音響センシング手法

本章では、本稿で提案する空中音響センシング手法について述べる。

2.1 MUSIC 法

音源定位等において多く用いられる、一般的な MUSIC 法のアルゴリズムを示す。 M チャンネル入力音響信号の f フレーム目をフーリエ変換して得られる $Z(\omega, f)$ から、以下のように相関行列 $R(\omega, f)$ を定義する。

$$R(\omega, f) = \frac{1}{T_R} \sum_{\tau=f}^{f+T_R-1} Z(\omega, \tau) Z^*(\omega, \tau) \quad (1)$$

ここで、 ω は周波数ビン番号、 T_R は相関行列の計算に用いるフレーム数、 Z^* は Z の共役転置である。MUSIC 法では、こうして得られた $R(\omega, f)$ を固有値展開して固有ベクトルを計算する。

$$R(\omega, f) = E(\omega, f) \Lambda(\omega, f) E^*(\omega, f) \quad (2)$$

ただし、 $\Lambda(\omega, f)$ は降順に並んだ固有値を対角成分に持つ行列、 $E(\omega, f)$ は $\Lambda(\omega, f)$ に対応する固有ベクトルを並べた行列である。これと、センサに対する音の到来方向 θ に対応した伝達関数 $G(\omega, \theta)$ を用いて MUSIC 空間スペクトル $P(\omega, \theta, f)$ を計算する。

$$P(\omega, \theta, f) = \frac{|G^*(\omega, \theta) G(\omega, \theta)|}{\sum_{m=L+1}^M |G^*(\omega, \theta) e_m(\omega, f)|} \quad (3)$$

ただし、 L は目的音源数、 e_m は、 E に含まれる m 番目 ($1 \leq m \leq M$) の固有ベクトルを表す。このようにして得られた $P(\omega, \theta, f)$ を、音源方向を推定するために以下のように ω 方向に平均する。

$$\bar{P}(\theta, f) = \frac{1}{\omega_H - \omega_L + 1} \sum_{\omega=\omega_L}^{\omega_H} P(\omega, \theta, f) \quad (4)$$

なお、 ω_H 、 ω_L は使用する周波数ビンの上限と下限に対応したインデックスである。 $\bar{P}(\theta, f)$ に対して閾値処理、ピーク検出を行い、得られたピークに対する θ を音源方向として検出する。

2.2 MUSIC-SONAR

前節で述べた MUSIC 法は、フレーム単位でフーリエ変換を行うため、伝搬時間方向の分解能が低くなる。しかし、

パルスエコー法では、ターゲットからの反射波の伝搬時間からターゲットまでの距離を算出するため、伝搬時間方向に高い分解能が必要である。また、送信周波数の帯域を小さくすれば、周波数方向の情報は不要である。そこで、MUSIC法をパルスエコー法に対応させるため、以下のよう
に拡張する。

M チャンネル入力音響信号に対し、直交検波により得られる複素信号 $z(t)$ から、以下のように相関行列 $R'(t)$ を定義する。

$$R'(t) = \frac{1}{t_{R'}} \sum_{\tau=t}^{t+t_{R'}} z(\tau)z^*(\tau) \quad (5)$$

ここで、 $t_{R'}$ は相関行列の平均化に用いる時間幅である。また、 t は伝搬時間 (サンプル点) であり、サンプル点単位の相関行列を算出することができる。また、直交検波では、解析周波数成分のみが算出されるので、 R' は t のみの関数となる。こうして得られた $R'(t)$ から、MUSIC法同様、固有ベクトルを算出し、空間スペクトル $P'(\theta, t)$ を得る。

$$R'(t) = E'(t)\Lambda'(t)E'^*(t) \quad (6)$$

$$P'(\theta, t) = \frac{|G'^*(\theta)G'(\theta)|}{\sum_{m=L+1}^M |G'^*(\theta)e'_m(t)|} \quad (7)$$

このように、サンプル点単位での空間スペクトルが得られ、伝搬時間方向に高い分解能を持つことができる。得られた $P'(\theta, t)$ に対して閾値処理、ピーク検出を行い、ターゲットからの反射波の到来方向および伝搬時間を検出し、ターゲットの位置へ変換を行う。以降、本手法を MUSIC-SONAR (MUSIC based Sound Navigation and Ranging) と呼ぶ。

3 評価実験

3.1 実験方法

計算機シミュレーションにより、音響信号を作成し、ターゲットの検出精度の評価実験を行った。設定条件を Fig. 2 に示す。送受波器としては、送波器を中心とした同心円上に等間隔に 16 チャンネルの受波器を設置したものを設定した。送波器、受波器はそれぞれ無指向性のものを仮定した。また、送波器から受波器までの距離は、ロボットに搭載することを想定し、50 mm とした。これらの送受波器一式を以降センサとする。センサの中心を $x = 0m, y = 0m$ の位置に設定し、センサの中心から距離 r 、方位角 θ の位置にターゲット (点反射体) を仮定し、音線法によりターゲットからの反射波を作成した。今回、送波器から受波器への直接波は考慮していない。 r を 1~3 m、 θ を $-180^\circ \sim 180^\circ$ と変化させ、各位置での計測精度について評価を行う。送信信号は sin 波 3 波にハニング窓を適応したものをを用いた。送信周波数は 1, 5 kHz の 2 種類について検証

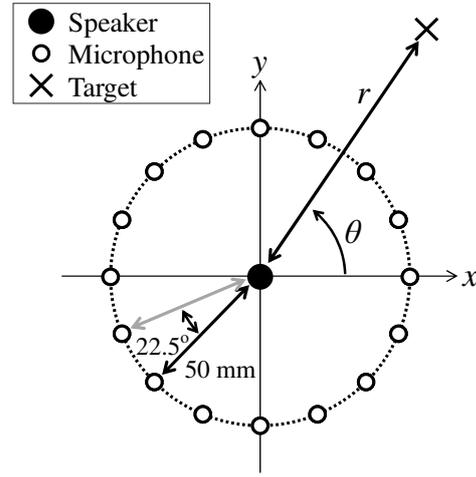


Fig. 2: Setting of the coordinate system of the simulation.

を行った。シミュレーションにより作成した音響信号を MATLAB にて合成開口法および MUSIC-SONAR にて処理を行い、比較を行う。MUSIC-SONAR で用いる伝達関数 (Eq. 7 における G) については、 θ が $-180^\circ \sim 180^\circ$ の範囲を 1° 刻み、 d が 1 m に対応するものを幾何計算にて算出した。

3.2 実験結果

実験結果について述べる。ターゲットが $r = 2m, \theta = 45^\circ$ ($x = 1.41m, y = 1.41m$) の位置にあった場合の処理結果を Fig. 3 に示す。(a) が合成開口法、(b) が MUSIC-SONAR にて処理を行った結果である。(c) には (b) の拡大図を示す。また、(i) は送信周波数が 5 kHz、(ii) は 1 kHz の場合の結果を示している。Fig. 2 に示す座標系に従ってプロットされており、カラーマップで各点における反射波の振幅の大きさを表している。合成開口法を用いた場合、ターゲットからの反射波が 1 点で交わらず、振幅の高い範囲がセンサを中心に方位方向に円状に広がっていることがわかる。これは、送信周波数が 5, 1 kHz の場合、波長がそれぞれ約 70 mm, 350 mm であり、センサが十分な開口幅を確保できていないためである。一方、MUSIC-SONAR を用いた場合、ターゲット位置に鋭いピークを確認することができ、方位方向への広がりは少ないことがわかる。このように合成開口法では高精度な計測を行うことができない条件でも、MUSIC-SONAR 法では高精度に計測が行える可能性があることがわかった。しかし、送信周波数が 5 kHz の場合、距離方向への広がり少ないが、1 kHz の場合、距離方向に 2 m 程度広がっていることが確認できる。これは送信信号の信号長によるものであり、信号長が長くなるほど、距離方向への広がりが大きくなり、計測の分解能が低下すると考えられる。そのため、実装には必要な分解能を考慮し、送信信号を設計する必要がある。

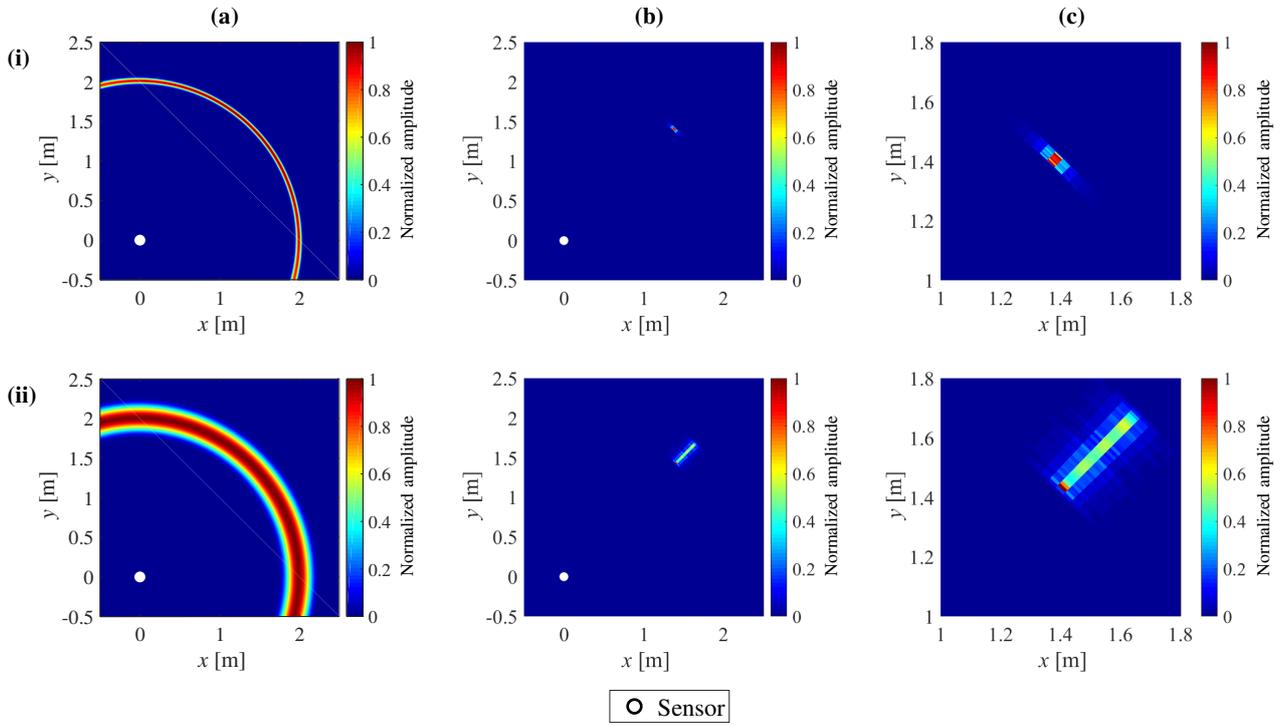


Fig. 3: Results of the simulation processed by: (a) synthetic aperture method, (b) MUSIC-SONAR. (c) Enlarged view of (b). Transmitted frequency was: (i) 5 kHz, (ii) 1 kHz.

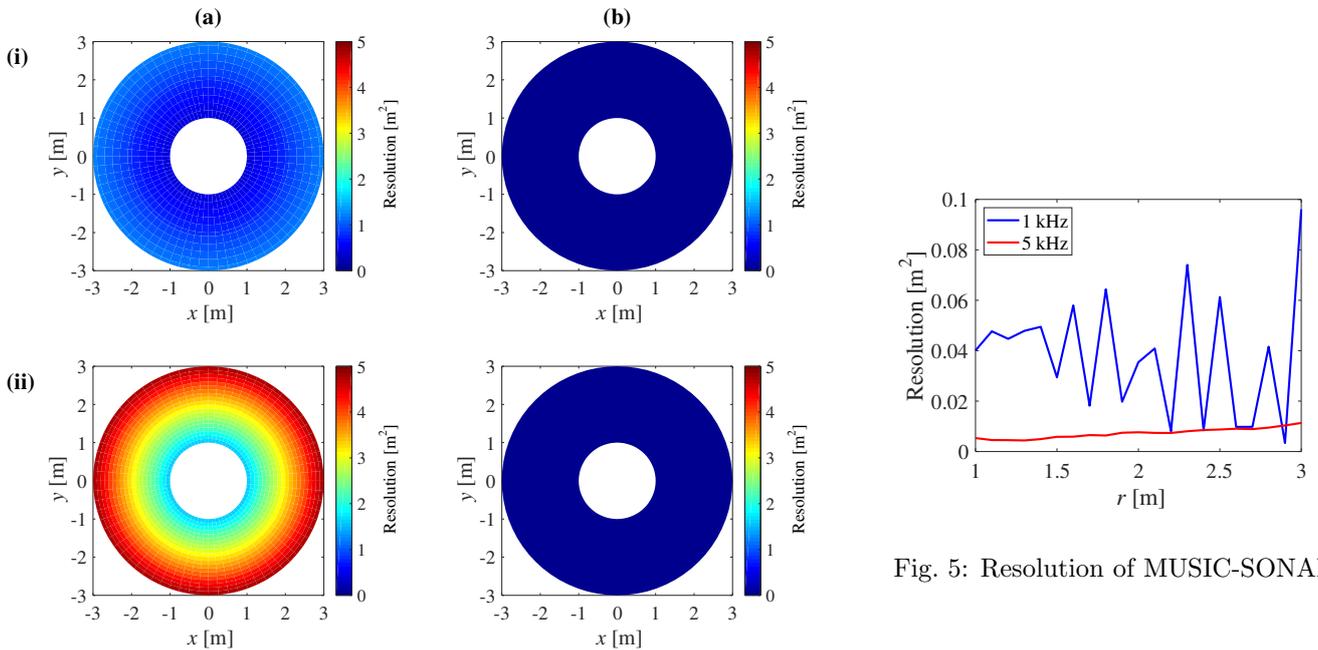


Fig. 4: Resolutions of: (a) synthetic aperture method, (b) MUSIC-SONAR. Transmitted frequency was: (i) 5 kHz, (ii) 1 kHz.

3.3 考察

各手法を用いた場合のターゲット検出性能の評価を行った。評価指標として、計測の分解能を用いた。各条件において、解析結果の最大点の半値以上の面積を分解能として

算出した。算出結果を Fig. 4 に示す。Fig. 3 同様、Fig. 2 に示す座標系に従ってプロットされている。(a) が合成開口法、(b) が MUSIC-SONAR の結果について分解能を算出した結果である。また、(i) は送信周波数が 5 kHz、

は 1 kHz の場合の結果を示している。各点にターゲットがあった場合の分解能をカラーマップにて示している、合成開口法の場合、送信周波数が 1, 5 kHz どちらにおいても分解能が 1 m^2 以上あり、実用には適さないことがわかる。MUSIC-SONAR では、分解能が 1 m^2 以下であり、合成開口法と比べ分解能が高いことがわかる。詳細に確認するため、MUSIC-SONAR における分解能算出結果を θ 方向に平均した結果を Fig. 5 に示す。横軸が r 、縦軸が分解能であり、青線で送信周波数が 1 kHz、赤線が 5 kHz の場合の結果である。MUSIC-SONAR では、方位角方向の分解能が距離によって固定されているため (本実験の場合 1°)、 r が大きくなるほど面積が大きくなっている。また、 r が 1~3 m の範囲において、送信周波数が 5 kHz の場合は 0.01 m^2 以下、1 kHz の場合は 0.1 m^2 以下となっている。つまり、最小検出能がそれぞれ $0.1 \text{ m} \times 0.1 \text{ m}$, $0.3 \text{ m} \times 0.3 \text{ m}$ 程度であることがわかる。一般的に LIDAR の分解能は \pm 数 cm であり、それらと比較すると本実験条件での MUSIC-SONAR は分解能は低いが、障害物検知などといったロボットのための周辺環境認識のためには十分な分解能であると言える。また、各パラメータを最適に設計すればより分解能が向上する可能性がある。以上の結果から、MUSIC-SONAR の有用性を確認することができた。

4 おわりに

本稿では、ロボットなどの自律動作などのための周辺環境認識を目的に、MUSIC 法に基づく空中音響センシング手法についての検討を行った。MUSIC 法では音の到来方向は高分解能で検出できるが、伝搬時間方向の分解能が低いといった問題点があるため、時間方向の分解能を確保できるよう拡張した MUSIC-SONAR を提案した。評価実験により、従来手法である合成開口法では計測ができない状況でも、高分解能な計測が可能ながわかり、有用性が確認できた。しかし、本稿では単純な計算機シミュレーションでのみ評価を行ったため、実環境では精度が低下する可能性がある。今後は、さまざまな状況での評価実験を行い、センサの最適なパラメータを決定したうえで、実機による実測実験を行い、評価を行う予定である。

参考文献

- [1] Y. Zhang, J. Wang, X. Wang, J. M. Dolan, "Road-Segmentation-Based Curb Detection Method for Self-Driving via a 3D-LiDAR Sensor", *IEEE Transactions on Intelligent Transportation Systems (Early Access)*, pp. 1-11, 2018, DOI: 10.1109/TITS.2018.2789462.
- [2] D. Shen, Y. Huang, Y. Wang, C. Zhao, "Research and Implementation of SLAM Based on LIDAR for Four-Wheeled Mobile Robot", *IEEE International Conference of Intelligent Robotic and Control Engineering (IRCE)*, pp. 19-23, 2018, DOI: 10.1109/IRCE.2018.8492968.
- [3] C. Chi, Z. Li, "High-Resolution Real-Time Underwater 3-D Acoustical Imaging Through Designing Ultralarge Ultrasparse Ultra-Wideband 2-D Arrays", *IEEE Transactions on Instrumentation and Measurement*, Vol. 66, pp. 2647-2657, 2017, DOI: 10.1109/TIM.2017.2709519.
- [4] L. Royer, A. Krupa, G. Dardenne, A. L. Bras, E. Marchand, M. Marchal, "Real-time target tracking of soft tissues in 3D ultrasound images based on robust visual information and mechanical simulation", *Medical Image Analysis*, Vol. 35, pp. 582-598, 2016, DOI: 10.1016/j.media.2016.09.004.
- [5] Y. Bando, T. Otsuka, T. Mizumoto, K. Itoyama, M. Konyo, S. Tadokoro, K. Nakadai, H. G. Okuno, "Posture estimation of hose-shaped robot by using active microphone array", *Advanced Robotics*, Vol. 29, pp. 35-49, 2015, DOI: 10.1080/01691864.2014.981291.
- [6] R. O. Schmidt, "Multiple emitter location and signal parameter estimation", *IEEE Transactions on Antennas and Propagation*, Vol. 34, No. 3, pp. 276-280, 1986, DOI: 10.1109/TAP.1986.1143830.
- [7] K. Hoshiba, K. Washizaki, M. Wakabayashi, T. Ishiki, M. Kumon, Y. Bando, D. Gabriel, K. Nakadai, H. G. Okuno, "Design of UAV-Embedded Microphone Array System for Sound Source Localization in Outdoor Environments", *Sensors*, vol. 17, No. 11, pp. 1-16, 2017, DOI: 10.3390/s17112535
- [8] K. Okutani, T. Yoshida, K. Nakamura and K. Nakadai, "Outdoor auditory scene analysis using a moving microphone array embedded in a quadcopter", *Proceedings of the IEEE/RSJ International Conference on Robots and Intelligent Systems (IROS)*, pp. 3288-3293, 2012, DOI: 10.1109/IROS.2012.6385994.
- [9] T. Ohata, K. Nakamura, T. Mizumoto, T. Tezuka and K. Nakadai, "Improvement in outdoor sound source detection using a quadrotor-embedded microphone array", *Proceedings of the IEEE/RSJ*

International Conference on Robots and Intelligent Systems (IROS), pp. 1902-1907, 2014, DOI: 10.1109/IROS.2014.6942813.

- [10] K. Nakadai, T. Takahashi, H. G. Okuno, H. Nakajima, Y. Hasegawa and H. Tsujino, “Design and Implementation of Robot Audition System ‘HARK’ – Open Source Software for Listening to Three Simultaneous Speakers”, *Advanced Robotics*, Vol. 24, No. 5-6, pp. 739-761, 2010, DOI: 10.1163/016918610X493561.