

低サイドローブ設計 64ch 球形マイクロホンアレイの開発

Design and Implementation of Omni-Directional Ball Microphone Array

○ 佐々木洋子*, 椋澤光隆*, Simon THOMPSON*, 加賀美聡*, 尾路京一†

Yoko SASAKI*, Mitsutaka KABASAWA*, Simon THOMPSON*, Satoshi KAGAMI*, Kyoichi ORO†

* 産業技術総合研究所 デジタルヒューマン工学研究センター

*National Institute of Advanced Industrial Science and Technology

† 関西電力(株)

†Kansai Electric Power Co., Inc.

y-sasaki@aist.go.jp

Abstract

This paper presents a microphone array design and the evaluation result of the developed microphone array. We propose an evaluation index of directional characteristic of Delay and Sum BeamForming to optimize microphone array design. Using beamforming simulation, we obtain a microphone arrangement which minimizes sidelobes, and improves the basic performance of beamforming. It has 64 microphones in a 350mm diameter ball designed to mount on a mobile robot and omni-directional directivity in azimuth and elevation. The performance of the proposed microphone array is verified in different real environments. Experimental results of sound localization show the effectiveness of the array in some challenging environment and its robustness for different pressure sound sources to cover larger areas.

1 はじめに

「どこから何の音がするか」周囲の音を捉える機能は、特に移動ロボットの環境知覚機能のひとつとして重要である。この機能を実現する方法として、複数のマイクロホンをロボットに搭載したマイクロホンアレイによるアプローチが一般的で、これまでに多くのシステムが提案されている [1, 3, 4, 6]。一方、様々な条件が想定される実環境中では、未知の環境条件、音圧差・距離差の異なる複数音の扱いなど、まだ課題も多い。後段の信号処理部分の信頼性を増すためにも、マイクロホンアレイの基本性能の向上は重要な要素と言える。

遅延和ビームフォーミング (Delay and Sum BeamForming, DSBF) は、マイクロホンアレイによる音源定位・分

離の最も簡単な手法である。環境の伝達関数など事前情報を必要とせず、計算も簡単なため移動ロボットに適した手法と言える。DSBF の性能は一般的にマイクロホン数に依存し、大規模なシステムほど高い性能が得られる。壁面状に配置した 1020ch アレイなど、大規模なシステムの有効性が示されている [7, 9]。一方、DSBF はサイドローブが多く、鋭い指向性が得られない、という欠点も知られている。特に小規模なシステムでは、マイクロホン配置によってその特性が異なり、性能に大きく影響する。

DSBF 以外にも様々なアレイ信号処理手法がある [5]。Griffith-Jim 型に代表される適応型ビームフォーミングでは適応的に死角を形成することで高い SN 比が得られる。ただしステアリングベクトルの推定間違いが性能を劣化させるため、移動ロボットのような動的な条件では扱いが難しい。近年広く用いられている音源定位手法として、Multiple Signal Classification (MUSIC) が挙げられる。事前に環境の伝達関数および音源数を与える必要があるが、鋭い指向性が得られ比較的少ないマイクロホン数でも高精度な音源定位が実現可能である。ただし、弱い音や距離の離れた音といった指向性の弱い信号の定位は難しい。

本稿では、ロボットに搭載可能な数十チャンネルの小規模なシステムを対象とした、DSBF に適したマイクロホンアレイの設計について述べる。ビームフォーミングの基本性能を向上させるため、ビームフォーミング時の指向特性を定量評価するための指標を定義し、方位角・仰角の全方位に高感度な特性を得られるマイクロホンアレイの設計を行う。後半では開発したマイクロホンアレイによる屋外での音源定位実験の結果について述べる。

2 遅延和ビームフォーミングの指向特性

本節では、まず DSBF の基礎式について整理し、マイクロホンアレイの性能を定量評価するための指向特性の評価指標を定義する。

2.1 遅延和ビームフォーミング

マイクロホンアレイの中心を原点とする極座標系で焦点 C の座標を (l, θ, ϕ) とおく。 C から i 番目のマイクロホン ($i = 1, 2, \dots, M$) までの距離を L_i とすると、 i 番目のマイクロホン入力に与える遅延は次式で表わされる。

$$\tau_i(l, \theta, \phi) = \frac{l - L_i(l, \theta, \phi)}{V_s} \quad (1)$$

ただし V_s は音速である。

各マイクロホン入力の位相を揃えて加算することで目的方向の信号が得られる。時刻 t における i 番目のマイクロホン入力を $x_i(t)$ とおくと、焦点 C に対する DSBF 出力は式 (2) となる。

$$s_c(t) = \frac{1}{M} \sum_{i=1}^M x_i(t + \tau_i) \quad (2)$$

焦点を全方位にスキャンさせることで、「空間スペクトル」と言われる各方向に対する音圧分布を示す曲線（二次元定位の場合は曲面）が得られる。ひとつの点音源に対する空間スペクトルが指向特性を表わす。

2.2 指向特性の評価指標

対象音源の高精度な音源定位・分離を行うためには、メインローブは鋭く、サイドローブは低く抑えることが重要となる。そこで、DSBF 時の指向特性を定量評価するため、以下の 2 つの指標を定義する。

- IM_s : 球の表面積に対するメインローブエリアの面積の割合
- IS_{max} : サイドローブゲインの最大値

メインローブエリアとは焦点方向のゲインに対して -12dB までの範囲とし、サイドローブゲインとはそれ以外の範囲のピークゲインとする。Figure 1 に焦点方向を 0dB とした空間スペクトルの例を示す。中央の \times が焦点、ハッチング部分がメインローブエリアとなる。

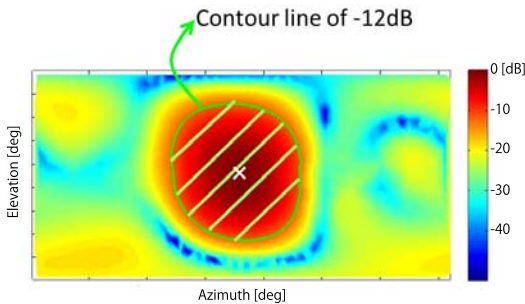


Figure 1: Directivity pattern of DSBF

上記指標をもとに、評価関数を次式のように定義する。

$$Eva = \alpha \frac{1}{a} \frac{1}{N} \sum_{\omega=\omega_L}^{\omega_H} IM_{s\omega} - \frac{1}{b} \frac{1}{N} \sum_{\omega=\omega_L}^{\omega_H} IS_{max\omega} \quad (3)$$

ただし ω は周波数、 ω_L, ω_H は対象とする下限・上限の周波数である。また N は ω_L から ω_H までの周波数ビンの数を表わす。 α は重み付けのための係数である。 a, b は各項の重みを等しくするための正規化係数である。本研究では、一様にランダムに生成したマイクロホンアレイに対する指標を用い、 a を IM_s の母平均、 b を IS_{max} の母平均とする。

3 マイクロホン配置の設計

3.1 低サイドローブ特性を得る配置

方位角・仰角の全方位に同一の指向性を持たせるため、マイクロホンの配置は同心球状とし、各球面上で偏りのないように分布させる。また低サイドローブ特性を得る方法として提案している、球の中心からの距離 r の関数としてマイクロホンの密度関数 [2] を用いる。この密度関数の基本的な考え方は、DSBF で各マイクロホン入力を同位相化することを仮想的にマイクロホンを音源から等距離の球面への射影することと捉えられるため、射影時に仮想球面上で均一に分布させるものである。つまり、マイクロホン密度 ρ は $\rho \propto 1/\cos(r)$ と定義できる。

式 (4) に 4 次関数として求めた密度関数を示す。

$$\rho(r) = \frac{0.328}{\cos(\frac{r}{R})^4} + \frac{0.117}{\cos(\frac{r}{R})^3} - \frac{0.496}{\cos(\frac{r}{R})^2} + \frac{0.117}{\cos(\frac{r}{R})} - 0.122 \quad (4)$$

各項の係数は射影後のマイクロホンが均一分布になるよう繰り返し計算により求めた。

式 (4) を基に以下の手順でマイクロホン配置を生成する。

1. アレイサイズおよびマイクロホン数を与える。
2. マイクロホンを配置する同心球の数を与え、各球の半径と配置するマイクロホン数のパターンを生成する。
3. 2. で生成したパターンごとに、各球面上で偏りなく分布させた配置を生成する。

有限個のマイクロホンに離散化すると、一つの条件（アレイサイズ/マイクロホン数）に対して式 (4) を満たす配置が多数考えられる。ここで生成した複数の配置を式 (3) で定量評価し、最終的な配置を決定する。

3.2 マイクロホン数・アレイサイズの検討

DSBF では大規模なシステムほど性能がよく、マイクロホン数が多いほど高い SN 比が得られ、アレイサイズが大きいほどメインローブ幅が細くなる傾向にある。

まず必要なマイクロホン数およびアレイサイズを検討するために、シミュレーション上でマイクロホン数およびアレイサイズと性能の関係を評価する。

以下の条件を対象として、 6×8 種類のそれぞれの条件について式 (4) を満たす 100 通りのマイクロホン配置を生成した。

- アレイ直径 : 0.2, 0.3, 0.4, 0.6, 0.8, 1.0 [m]
- マイクロホン数 : 30, 50, 70, 90, 110, 130, 150, 170

マイクロホン同士が干渉し合うものを除き, 合計 3842 通りを用いて評価した.

なお, 評価においては特定方向へ指向性が特化することを防ぐため, 20 の異なる焦点に対して式 (3) の Eva を求め, 平均値を評価値とした. 次節以降のシミュレーションについても同様である.

Figure 2 に結果を示す. 左から a) メインローブサイズ, b) サイドローブゲイン, c) 両者の等高線表示となっており, それぞれ縦軸がアレイサイズ, 横軸がマイクロホン数となっている. グラフの右下の値がないのは, マイクロホン同士の干渉により十分な数の配置パターンが得られなかったためである. a) では上にいくほど値が小さくなっており, アレイサイズが大きいほど鋭いメインローブとなることが分かる. b) では左右に値が変化しており, マイクロホン数が SN 比に影響することが分かる. またマイクロホン間の距離が小さくなる右下ほどサイドローブゲインが下がっている.

Figure 2 の結果を踏まえ, 我々の移動ロボットに搭載するマイクロホンアレイとして, 本稿ではマイクロホン数: $M = 64$, アレイ直径を 360 mm と設定する. ロボットの詳細については 4.1 節で述べる. 目標性能は, $(IMs, ISmax) = (15\%, -14dB)$ とする.

3.3 配置シミュレーション

マイクロホン数を $M = 64$, 最大アレイ直径を 360 mm とし, シミュレーション上で指向特性の評価を行った. 周波数は 500 ~ 3000Hz まで 100Hz おきに計算した. Figure 3 に 70000 通りのマイクロホン配置に対する評価値の分布を示す. 縦軸がサイドローブゲイン, 横軸がメインローブサイズとなっており, 左下ほど性能が高いことを示している. 横軸の下段の値は参考値であり, メインローブの形状が円であると仮定したときのメインローブ幅を表わしている. Figure 3 において, マゼンタ及び緑の点はそれぞれ $\alpha = 1.5, 1.7$ での評価値の上位 10 点を示している.

この結果から得られた最良配置の候補として, 図中マゼンタで示した点では, すべて直径 350mm の球面上に 51 のマイクロホンを持つ配置であった. ここでシミュレーションから得られた配置をそのまま実装するのは構造が複雑になるため, 以上の結果を踏まえ, 以下のように実装のための制約条件を設定する. 最外周のマイクロホンを 350mm の球面上に 50 個とし, 高さを制限し均一に分布させるため, C60 フラーレンの頂点上に配置する. 頂点は高さの異なる 8 層に分解できるため, 下側 2 層を除いた 6 層上の計 50 点をマイクロホンの位置とする. その内側に配置する残りの 14 個のマイクロホンについては, 同

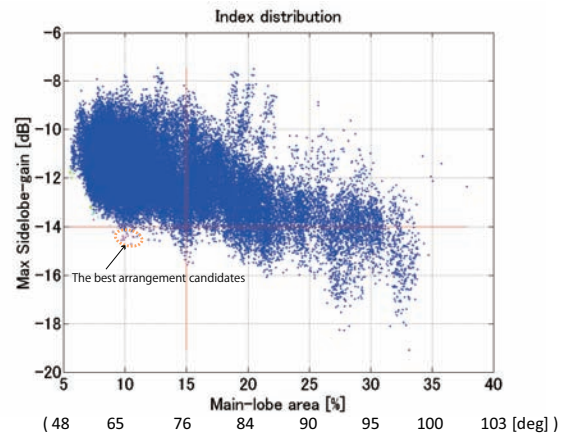


Figure 3: Evaluation Index Distribution (No Constraint)

様の手順でシミュレーションによる指向特性評価を行い, 決定する.

Figure 4 に 3000 通りのマイクロホン配置に対する評価値の分布を示す. 最良配置は, 直径 350mm の球面上に 50 個, 直径 150mm の球面上に 12 個, さらに内側に 2 個のマイクロホンがあり, 評価値は $(IMs, ISmax) = (9.1\%, -15.1dB)$ となった. 配置の詳細については 4.1 節で述べる.

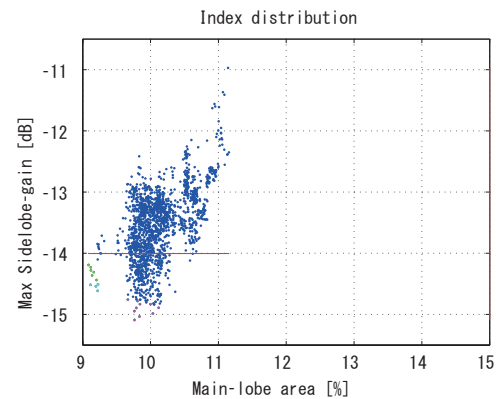


Figure 4: Evaluation Index Distribution (Fixed outside)

Figure 5 に設計したマイクロホンアレイの指向特性を示す. $(\theta, \phi) = (180^\circ, 0^\circ)$ が焦点である. それぞれの周波数における左右は同じデータを三次元表示したものと平面に展開したものである. 平面展開では上下が広がっているが, 全体的にサイドローブが低く抑えられていることが確認できる.

4 実装

前節で設計したマイクロホン配置をもとに, 球形マイクロホンアレイを実装した. 本節では, 開発したマイクロホンアレイと, 屋外での音源定位実験について述べる.

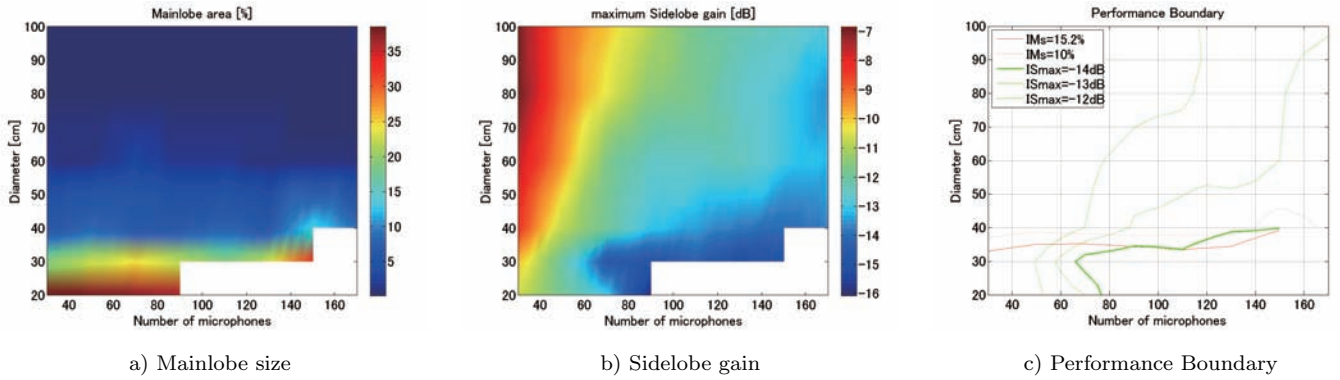


Figure 2: Consideration of Array Size and Number of Microphones

方向を $\phi = 0^\circ$, 真上方向を $\phi = 90^\circ$ と設定する .

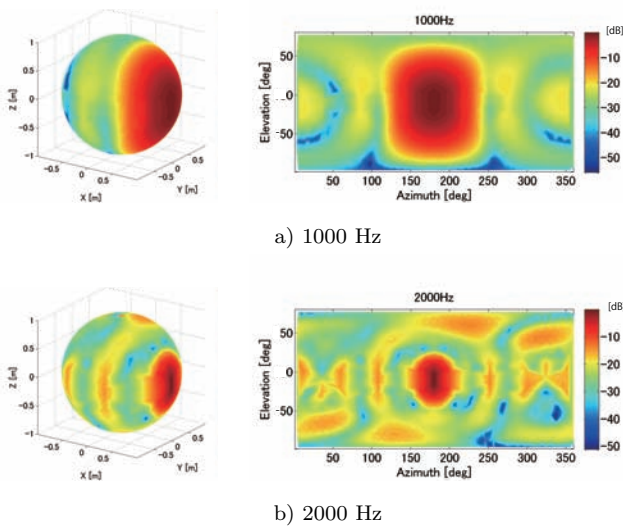


Figure 5: Directivity Pattern of Proposed Microphone Array

4.1 64 チャンネル球形マイクロホンアレイ

Figure 6 に開発したマイクロホンアレイを示す . 上段が CAD 図面のスナップショットで a) にマイクロホン基板 , b) にマイクロホン基板とジグを表示している . ひとつのマイクロホン基板のサイズは 30×20 mm で , 底部に設置したコントロール基板に接続されている . すべてのマイクロホンは上向きに設置している . ジグの設計においては構造物による音波の回り込みの影響を抑えられるよう考慮し , -20° から 90° (真上方向) までの範囲で直接波を捉えられるような設計になっている . c) , d) はマイクロホンアレイおよび搭載したロボットの写真である . ロボットは Segway の RMP200 ATV をベースに天板等を改良したもので , 2 輪駆動の倒立振り子機構である . 変電所内を自律走行し , 各種機器を点検できるように設計されている . マイクロホンアレイはロボットの上面に設置した .

本節以降では , 座標系として $\theta = 0^\circ$ をロボットの正面 , $\theta = 90^\circ$ をロボットの左方向とし , 仰角については , 水平

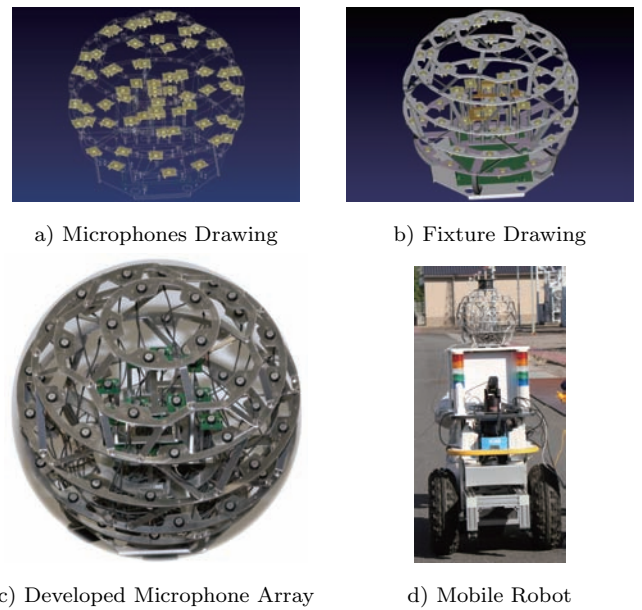


Figure 6: Developed Microphone Array

開発したマイクロホンアレイの仕様を Table 1 にまとめる . ゲインおよびサンプリング周波数はソフトウェア上で変更可能である . 次節の実験は , すべて 16kHz サンプリングで行う .

Table 1: Spec. of The Microphone Array Board

Microphone	Primo EM100PT
Num. of Channels	64
Sampling Frequency	8, 16, 32, 48 [kHz]
Resolution	16 [bit]
Amplifier	AK4563A (Programmable Gain Amp.)
Interface	USB 2.0
Power Supply	+5 [v]

4.2 音源定位実験

ロボットに搭載したマイクロホンアレイを用いて、音源定位実験を行った。音源定位にはDSBFの後段処理としてFBS(Frequency Band Selection)を併用した手法 [8]を用いる。一回の計算に用いるデータ長は1024点(64msec)とした。

4.2.1 静止時の音源定位精度評価

まずスピーカを音源とし定位角度の精度評価を行った。音源には、男声/女声の連続発話およびクラシック音楽を用い、スピーカ(YAMAHA101III)から再生した。背景雑音に対する音源のSNRは約10dBである。音源までの距離を3m, 9mとし、仰角を0, 15, 30, 45degと変化した場合の平均角度誤差をFigure 7, 8に示す。それぞれ、赤が角度誤差、緑/青が方位角/仰角成分を示している。最大でも3.6degと高精度に定位できている。誤差は主に仰角成分であり、仰角が大きいほど誤差が小さくなっている。これはマイクロホンを上向きに設置しているため、各素子の仰角方向の指向性が異なることと、ジグによる回り込みの影響が考えられる。

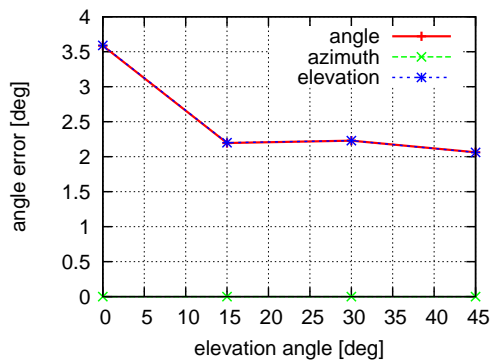


Figure 7: Average Sound Localization Error in Static Condition (distance=3m)

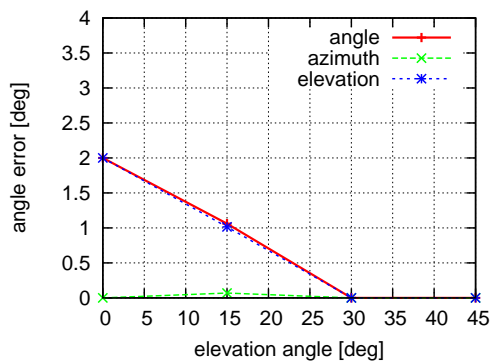


Figure 8: Average Sound Localization Error in Static Condition (distance=9m)

4.2.2 変電所内の音源定位

次に変電所の機械音を対象として音源定位実験を行った。定位計算の条件は前節と同様である。Figure 9に実験を行った変電所内の配置図を示す。オレンジの丸で囲った2カ所が主な音源である。図の左方向が北となっており、南側の円は分路リアクトルの稼働音で直径約2[m]の大口径ファンが主な音源、西側の楕円は変圧器の動作音で、人には敷地内全域で聞こえる低周波数帯域の音である。変圧器は6ブロックあり、各ブロックに3基ずつ並んでいる。これらの機械音を除くと比較的静かな環境である。水色の下向き矢印が静止時のマイクロホンアレイの位置を表わす。位置はGPS(Garmin Geko301)で取得した。マイクロホンアレイの向きは、図左方向が0[deg]、反時計回りに方位角正方向となっている。また変圧器前の赤線はロボットの軌跡を表わす。位置はNTPでロボットと時刻を同期させた測量用のレーザ測距装置で測定した。ロボットの走行速度は約1.3m/sであった。

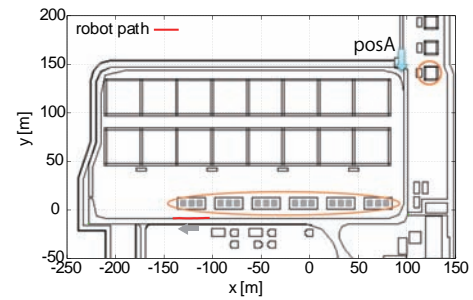


Figure 9: Robot Path in the Power Substation

Figure 10, 11に音源定位の結果を示す。それぞれ上段が仰角、下段が方位角の結果となっている。Figure 10はFigure 9で示したposAで静止させた状態での結果である。 $\theta = 180^\circ$ 方向10mの位置に分路リアクトルのファンがあり、 3×6 基並んだ変圧器の中心が、およそ $\theta = 45^\circ$ 方向に180m離れた位置となっている。グラフではそれぞれの音源を定位できていることが確認できる。

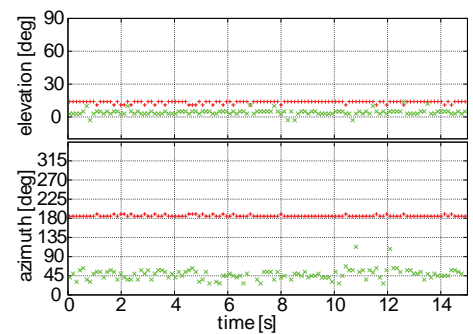


Figure 10: Sound Localization Result in the Power Substation: static condition at posA

Figure 11は変圧器前を走行したときの結果である。通過

した3つの変圧器をそれぞれ定位できていることがわかる。また仰角の値を見ると、真横を通過する際に ($\theta = 270^\circ$) 最も高くなっている。

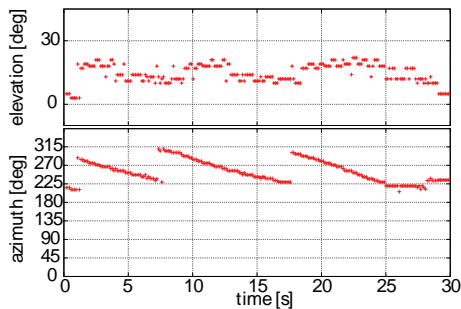


Figure 11: Sound Localization Result in the Power Substation: from moving robot

5 おわりに

本稿では、DSBFの基本性能を向上させるためのマイクロホンアレイの設計について述べた。指向特性を定量評価するための評価関数を定義し、様々な配置に対してDSBFの性能を定量評価することで、全方位に高感度な特性を持つ球形マイクロホンアレイを設計した。開発したマイクロホンアレイは、サイドローブを低減させたことで音圧差のある複数の音源を定位可能である。ロボットに搭載したアレイによる屋外での音源定位実験では、近くの音源とともに指向性の弱い離れた音源を定位可能で、ロボットの走行中も有効であることを確認した。

参考文献

- [1] Hideki Asoh, Isao Hara, and Futoshi Asano. Tracking human speech events using a particle filter. In *Proceedings of IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP 2005)*, pp. MSP-P2.6, Philadelphia, USA, 2005.
- [2] Tomoaki Fujihara, Yoko Sasaki, Satoshi Kagami, and Hiroshi Mizoguchi. Arrangement optimization for narrow directivity and high s/n ratio beam forming microphone array. In *Proceedings of the 7th Annual IEEE Conference on SENSORS (IEEE SENSORS 2008)*, pp. 450–453, Lecce, Italy, October 2008.
- [3] Carlos Toshinori Ishi, Shigeki Matsuda, Takayuki Kanda, Takatoshi Jitsuhiro, Hiroshi Ishiguro, Satoshi Nakamura, and Norihiro Hagita. Robust speech recognition system for communication robots in real environments. In *Proceedings of IEEE-RAS International Conference on Humanoid Robots (HUMANOIDS2006)*, pp. 340–345, Genova, Italy, December 2006.
- [4] Hyun-Don Kim, Jong-Suk Choi, and Munsang Kim. Speaker localization among multi-faces in noisy environment by audio-visual integration. In *Proceedings of IEEE-RAS International Conference on Robots and Automation (ICRA2006)*, pp. 1305–1310, Orlando, Florida, May 2006.
- [5] Nikolaos Mitianoudis and Mike E. Davies. Audio source separation: Solutions and problems. *International Journal of Adaptive Control and Signal Processing*, Vol. 18, No. 3, pp. 299–314, March 2003.
- [6] Kazuhiro Nakadai, Hirofumi Nakajima, Masamitsu Murase, Satoshi Kaijiri, Kentaro Yamada, Yuji Hasegawa, Hiroshi G. Okuno, and Hiroshi Tsujino. Real-time tracking of multiple sound sources by integration of in-room and robot-embedded microphone arrays. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-2006)*, pp. 852–859, Beijing, China, September 2006.
- [7] Harvey F. Silverman, William R. Patterson III, and Joshua Sachar. Factors affecting the performance of large-aperture microphone array. *The Journal of the Acoustical Society of America*, Vol. 111, No. 1, pp. 2144–2157, May 2002.
- [8] Yuki Tamai, Yoko Sasaki, Satoshi Kagami, and Hiroshi Mizoguchi. Three ring microphone array for 3d sound localization and separation for mobile robot audition. In *Proceedings of 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2005)*, pp. 903–908, Edmonton, Canada, August 2005.
- [9] E. Weinstein, K. Steele, A. Agarwal, and J. Glass. Loud: A 1020-node modular microphone array and beamformer for intelligent computing spaces. Technical Report MIT-LCS-TM-642, MIT/LCS Technical Memo, April 2004.