

# マイクロホンアレイの位置推定によるホース型ロボットの姿勢推定

坂東 宜昭<sup>†</sup> 大塚 琢馬<sup>†</sup> 糸山 克寿<sup>†</sup> 昆陽 雅司<sup>‡</sup> 田所 諭<sup>‡</sup> 中臺 一博<sup>§</sup> 奥乃 博<sup>†</sup>

<sup>†</sup> 京都大学 大学院情報学研究科

<sup>‡</sup> 東北大学 大学院情報科学研究科

<sup>§</sup> (株) ホンダ・リサーチ・インスティテュート・ジャパン, 東京工業大学 情報理工学研究科

## 1. はじめに

ホース型ロボット [1][2] とは細長い形状が特徴のレスキューロボットで、瓦礫の隙間に挿入し先端のカメラを用いて被災者の捜索を行う。ホース型ロボットの課題に姿勢推定があり、自由に変化するホースの姿勢を制御するために姿勢情報が必要である。既存の内界センサによる姿勢推定法 [3] は積分型計測法であり、累積誤差の問題だけでなく、ひっくり返ったりして姿勢が突然変化した場合に対応できなかった。本研究では非積分型計測法として、ホース型ロボットにマイクロホンアレイと小型スピーカを装着し、小型スピーカから再生する試験音のマイクロホン間の到達時間差を手がかりとした手法を開発する。到達時間差は、現在のマイクロホンと音源の位置だけに依存するので累積誤差問題を回避できる。さらに、マイクロホンアレイを使用して音源定位・分離に応用できるので被災者の位置推定や音声強調が可能となる(図1)。以降では、音による姿勢推定法について述べ、障害物を配置した実験室での評価を報告する。

## 2. 音によるホース型ロボットの姿勢推定

図2に提案法でのマイクロホンと小型スピーカの配置を示す。マイクロホンと小型スピーカは交互にホース上で等間隔 $l$ だけ離して配置する。各マイクロホンと小型スピーカをそれぞれ手元から順に  $\text{mic}_1, \text{src}_1, \dots, \text{src}_N, \text{mic}_M$  とする。ただし、 $M, N$  はそれぞれマイクロホンと小型スピーカの個数を表し、 $N = M - 1$  である。本稿ではロボットの姿勢は二次元平面上での曲線で表現されるとし、 $\text{mic}_i, \text{src}_j$  の各座標は、 $\mathbf{x}_{\text{mic}_i}, \mathbf{x}_{\text{src}_j} \in \mathbb{R}^2$  とする。

本稿で扱うマイクロホンアレイを用いた姿勢推定問題は次の通り。

入力	試験音 $H(\omega)$ を録音した $M$ チャンネル同期音響信号 $\{S_{\text{mic}_1}(\omega), \dots, S_{\text{mic}_M}(\omega)\}$
出力	ロボット姿勢 $\mathbf{x}_{\text{mic}_i}, \mathbf{x}_{\text{src}_j}$
仮定	(1) 再生スピーカの番号 $j$ は既知。 (2) $\mathbf{x}_{\text{mic}_1}$ と $\mathbf{x}_{\text{src}_1}$ は既知。

ここで、 $\omega$  は周波数を表し、試験音とは到達時間差推定のために小型スピーカで再生する元信号である。出力は、マイクロホンと小型スピーカの二次元座標である。仮定 (1) により複数の小型スピーカが区別され、仮定 (2) により推定姿勢の回転と平行移動が制限される。

マイクロホンと音源の位置を同時推定する関連研究に、補助関数法による手法 [4] と EKF-SLAM による手法 [5] がある。前者はオフライン設計でロボットの姿勢推定には適さず、後者はオンライン手法だが運動モデルが既知の単一音源を仮定しており、瓦礫中では使用できない。提案法はホース上に配置された複数個の音源を仮定し、ホースの姿勢を表す状態空間モデルを用いたオンライン推定を実現する。

Posture Estimation of Hose-shaped Robot with Loudspeakers and Microphone Array: Y. Bando, T. Otsuka, K. Itoyama (Kyoto Univ.), M. Konyo, S. Tadokoro (Tohoku Univ.), K. Nakadai (Titech, HRI-JP) and H. G. Okuno (Kyoto Univ.)

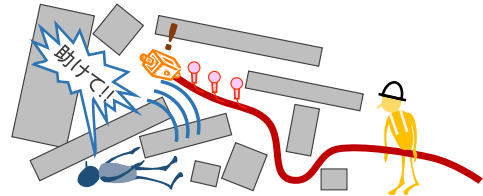


図1: ホース型ロボットの使用例。瓦礫内に進入し被災者を発見する。

### 2.1 手法概要: マイクロホン位置のオンライン推定

提案法では 1) 各小型スピーカから順に試験音を再生し、2) 試験音の各マイクロホンへの到達時間差を推定し、3) 得られた到達時間差から姿勢モデルの事後確率をガウス分布で近似し Unscented Kalman Filter (UKF) [6] を用いて点推定する。

### 2.2 到達時間差推定

到達時間差推定における課題は 1) 外部雑音、2) 残響・反射、3) 推定誤りがある。我々はこれまで 1), 2) を解決したが [7]、本稿では加えて 3) を解決する。

1) 外部雑音: 実環境では常に雑音が存在し、観測信号に混入する。この問題は Time Stretched Pulse (TSP) [8] を用いて解決する。TSP は自己相関が小さく到達時刻推定が容易で、非線形特性の強い小型スピーカでも大きなエネルギーで再生できるため、本問題を解決できる。

2) 残響と反射: ホース型ロボットが進入する狭い空間は試験音の残響や反射が発生する。残響の問題は、残響に頑健な相互相関である GCC-PHAT [9] を用いて解決する。反射の問題は、直接音が反射音より早く到達する性質を利用し、GCC-PHAT の相関係数のうち閾値を超える最初の時刻を到達時刻とすることで解決する。

3) 推定誤り: マイクロホンと小型スピーカを結ぶ直線上に障害物が存在すると直接音の到達時間差が得られない。到達時刻フィルタを設計し信頼できるチャンネル間の到達時間差のみで姿勢推定を行う。本フィルタは隣合うマイクロホン間の到達時間差がホース上のマイクロホンの間隔 ( $2l$ ) より長い場合にそのマイクロホンから末端のマイクロホンまでの到達時刻を棄却するよう設計する。

### 2.3 姿勢と到達時間差の定式化

図3にホースの姿勢モデルを示す。ロボットの姿勢は、マイクロホンと小型スピーカを頂点とする区分線形曲線により近似する。ホースの姿勢を表す状態変数  $\xi_k$  は、各頂点の角度  $\theta_{m,k}$  ( $1 \leq m \leq M+N-2$ ) と隣り合うマイクロホンと小型スピーカ間の距離  $l_{n,k}$  ( $1 \leq n \leq M+N-1$ ) からなる  $2(M+N)-3$  次元ベクトルである

$$\xi_k = [\theta_{1,k}, \dots, \theta_{N+M-2,k}, l_{1,k}, \dots, l_{N+M-1,k}]^T$$

各マイクロホンと小型スピーカの座標は手元側のマイクロホンと小型スピーカの座標  $\mathbf{x}_{\text{mic}_1}, \mathbf{x}_{\text{src}_1}$  を用いて再帰的に計算する。

状態遷移モデル 状態遷移は random walk で表現する

$$p(\xi_k | \xi_{k-1}) \sim \mathcal{N}(\xi_{k-1}, \mathbf{Q})$$

ただし、 $\mathbf{Q}$  は各状態変数の時間変化の大きさを制御する共分散行列である。

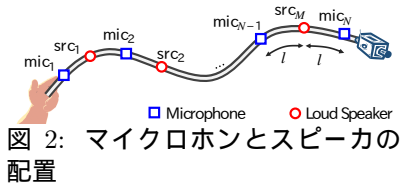


図 2: マイクホンとスピーカ配置

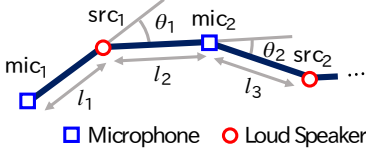


図 3: 姿勢モデル

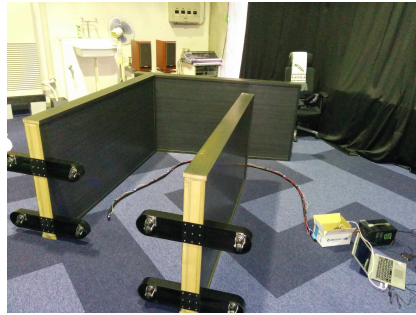


図 4: 実験風景

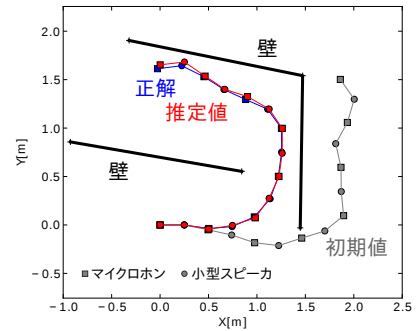


図 5: 正解姿勢 (青) と推定結果の一例 (赤)

観測モデル 観測は  $src_j$  再生時の直接音の各マイクロホン間の到達時間差と隣接するマイクロホンと小型スピーカ間の距離  $l_{n,k}$  とする．本モデルでは各頂点は 2 次元平面上で自由な位置を取りうる．観測として  $l_{n,k}$  を与え  $l_{n,k}$  の存在範囲に制限を加えて推定値の発散を抑制する．正解の  $l_{n,k}$  は得られないが，ホース上の配置間隔  $l$  を観測値と考え推定する．観測変数  $y_t^j \in \mathbb{R}^{(N^2+N)/2+M-1}$  の定義を次に示す．

$$p(y_k^j | \xi_k) = \mathcal{N}(g(\xi_k), \mathbf{R}_k)$$

$$g(\xi_k) = \begin{bmatrix} \tau_{mic_{i_1} \rightarrow i_2}^{src_j} = \frac{D_k^{i_2, j} - D_k^{i_1, j}}{C} | i_1, i_2 \in \mathcal{I} \\ [l_{n,k} | n = 1, \dots, M + N - 2]^T \end{bmatrix}^T$$

$$\mathbf{R}_k = \begin{bmatrix} \sigma_{r_r}^2 \mathbf{I}_{(N^2-N)/2} & \mathbf{0} \\ \mathbf{0} & \text{diag}(\sigma_{r_{l_1, k}}^2, \dots, \sigma_{r_{l_{M+N-1, k}}^2}) \end{bmatrix}$$

ここで， $D_k^{i,j}$  と  $C$  はそれぞれ， $mic_i, src_j$  間の距離，音速を表す．また， $\mathcal{I}$  は到達時刻フィルタで得たマイクロホンの番号である．マイクロホン・小型スピーカ間距離  $l_{n,k}$  の標準偏差  $\sigma_{r_{l_{n,k}}}$  は姿勢が大きく曲がりくねるほど増大すると考えられるので以下のように与える

$$\sigma_{r_{l_{n,k}}} = a \times \frac{|\theta_{n,k}| + |\theta_{n+1,k}|}{2} + b.$$

ただし， $a, b$  は曲げと標準偏差の関係を示すモデルパラメータである．

### 3. 実験

提案法と到達時刻フィルタのない改善前法 [7] とを比較しその有効性を確認する．また，障害物の有無で提案法を比較しその影響を考察する．図 4 のように，高さ約 1[m] の壁を設置した実験室に駆動機構を持たないモックアップロボットを静置し姿勢を推定する．モックアップロボットは MEMS マイクロホンと小型スピーカが交互に  $l=0.25$ [m] の間隔で配置された全長 3.5[m] のコルゲートチューブで，マイクロホンの個数は  $N=8$  個である．実験室の残響時間は  $RT_{60}=800$ [ms] である．試験音には TSP 信号 (512[ms]) を用い，各小型スピーカから計 56 回再生し録音した．録音は 24bit 量子化，16kHz サンプリングで行った．UKF の初期値は， $\theta_{m,0}$  は正解姿勢を平均とする正規分布から標準偏差 15 [deg] で乱数生成し， $l_{n,0}$  はロボットの縮みはなしとする (0.25[m]) ．その他のパラメータは実験的に与えた．提案法 (proposed)，到達時刻フィルタのない改善前法 (w/o filter)，壁を省きオープンスペースで提案法を用いた場合 (openspace) の 3 条件について，それぞれ 64 個の異なる初期値による推定結果について先端位置誤差を評価した．

図 5 に提案法の推定結果の一例を，図 6 に先端位置誤差の平均と標準偏差をそれぞれ折れ線とエラーバーで示す．提案法と改善前法を比較すると，提案法は到達時刻

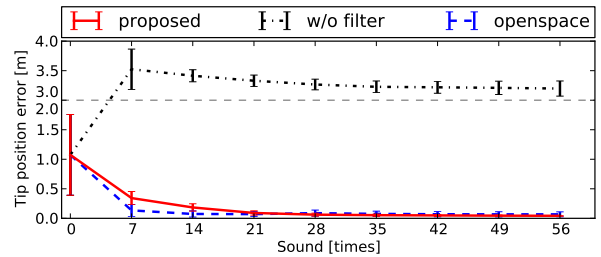


図 6: 推定結果の各更新における先端位置誤差

フィルタにより精度が向上している．また，従来法である加速度・ジャイロセンサが実現した精度は 3.0[m] のホースで先端位置誤差が 0.2[m] であるのに対し，提案法は 56 回目の更新で平均 0.043[m] であり従来法より高い精度を実現した．一方で提案法とオープンスペースでの結果を比較すると提案法の収束が遅い．これは到達時刻フィルタにより，一度の姿勢更新に使用できる情報が少ないためと考えられる．提案法ではスピーカを手元から先端まで順に鳴らしているが，推定値の分散を最小化するように次に鳴らすスピーカを動的に決定することにより収束の高速化が期待できる．

### 4. まとめ

本稿では音によるホース型ロボットの姿勢推定法について述べた．提案法は到達時刻フィルタを導入し改善前法では精度が低下する障害物が存在する環境下でも頑健に動作することを確認した．今後は実環境での評価，内界センサとの統合，音源定位・分離の実現などを行う．謝辞 本研究は科研費基盤 (S) No.24220006 の支援を受けた．

### 参考文献

- [1] H. Namari et al. Tube-type active scope camera with high mobility and practical functionality. In *IROS-2012*, 3679–3686.
- [2] A. Kitagawa et al. Development of small diameter active hose-ii for search and life-prolongation of victims under debris. *Journal of Robotics and Mech.*, 15(5):474–481, 2003.
- [3] M. Ishikura et al. Shape estimation of flexible cable. In *IROS-2012*, 2539–2546.
- [4] N. Ono et al. Blind alignment of asynchronously recorded signals for distributed microphone array. In *WASPAA-2009*, 161–164.
- [5] H. Miura et al. Slam-based online calibration of asynchronous microphone array for robot audition. In *IROS-2011*, 524–529.
- [6] S. Julier et al. New extension of the kalman filter to nonlinear systems. In *AeroSense-1997*, 182–193.
- [7] 坂東 宜昭 et al. ホースの伸び縮みによるマイク位置の変化を許容するマイクロホンアレイを用いたホース型ロボットの姿勢推定. In 第 38 回人工知能学会 AI チャレンジ研究会, 51–56.
- [8] Y. Suzuki et al. An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses. *The Journal of the Acoustical Society of America*, 97:1119.
- [9] C. Zhang et al. Why does phat work well in lownoise, reverberative environments? In *ICASSP-2008.*, 2565–2568.