

チューブ型ロボットの姿勢推定のための EKF-SLAMを用いた可変マイクロホンアレイ位置推定

坂東 宜昭[†]水本 武志[‡]中臺 一博^{‡‡}奥乃 博[‡][†] 京都大学 工学部情報学科[‡] 京都大学大学院 情報学研究科 知能情報学専攻^{‡‡} ホンダ・リサーチ・インスティテュート・ジャパン

1. はじめに

災害現場での被災者発見には、がれき内へ進入可能なチューブ型ロボットが有用である [1]。チューブ型ロボットに搭載される主な環境センサであるビデオカメラの視野角は狭いので、リモートオペレータがロボット周辺の環境を把握するには時間を要す。また、チューブ型ロボットは、姿勢が随時変化するので、それ自身の姿勢を認識するのが難しい [1]。チューブ型ロボットにマイクロホンアレイを組み込み、1) 各マイクロホンの位置同定、2) マイクロホン群による音源定位、3) ロボットの姿勢同定、4) 視聴覚情報統合による環境情報提示、といった機能を実現できれば、同型ロボットの有用性が高まると期待される。特に、マイクロホンの位置同定は、音源定位や姿勢同定の性能を左右する重要な機能である。本稿では、チューブ型ロボットのためのマイクロホン位置推定を行う手法を提案し、数値シミュレーションによりその性能を評価する。

2. チューブ型ロボットのモックアップ

デジタルマイクをロボット表面に配置し、ロボットのモータ位置にブザーを配置したプロトタイプを用いる。図 2 に示すように、マイクロホンは $L=50\text{cm}$ ごとに配置されており、ブザーは各マイクロホン間の中点に位置する。

3. 問題設定

入力	マイクロホンアレイで録音した同期音声
出力	各マイクロホンの位置
前提	ロボット上のマイクと音源の位置は既知

入力を同期マイクロホンアレイで録音された音声、出力を各マイクロホンの相対座標とする問題を解く手法として補助関数法を用いた手法 [3] と EKF-SLAM を用いた手法 [4] が存在する。しかし補助関数法による手法では時間計算量が大きくオフライン用途向けであることが知られており [4]、チューブ型ロボットの姿勢推定には利用できない。また、EKF-SLAM 法では音源の移動に制約があり、瓦礫中のチューブ型ロボットの姿勢推定に用いられない。そこで我々は、マイクロホン位置が与える音源位置の制約仮定し、その仮定を状態空間モデルに組み込むことで、マイクロホン位置の推定を行う。よって本稿で取り上げる問題設定は次のように表される。

4. EKF による姿勢推定

マイクロホンの位置推定にはオンライン推定の可能な EKF (Extended Kalman Filter) [5] を用いる。チューブ型ロボットは十分遅く進行するので、推定期間中のロボッ

EKF based calibration of microphone array for attitude estimation of tube robot: Yoshiaki Bando (Kyoto Univ.), Takeshi Mizumoto (Kyoto Univ.), Kazuhiro Nakadai (HRI-JP) and Hiroshi G. Okuno (Kyoto Univ.)

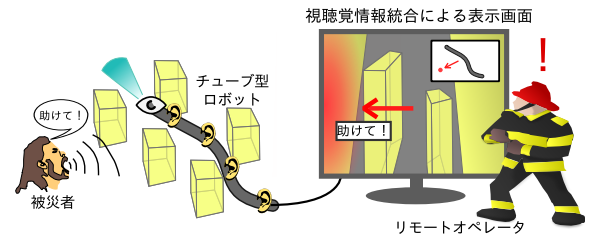


図 1: リモートオペレータによる瓦礫中の被災者の発見

トは静止しているとする。また一定時刻後のマイクロホン位置の変動は微小であるので、一定時刻ごとに EKF により独立したロボットの姿勢推定を行う。ただし EKF は初期値依存性が大きいので、1 時刻前の推定結果を次の推定の初期値に用い推定精度の向上を図る。

複雑な形状に曲がりくねるチューブ型ロボットの形状を細部まで推定することは困難であるので、本稿では以下の簡略化したモデルにより推定を行う。図 3 に示すようにロボットは各マイクロホンを節にした長さ L の線分によってつながれた形状をしており、その中点にブザーが配置されているとする。マイクロホンは端から順に $m_i (i = 1, 2, 3, 4)$ とする。また、 m_i と m_{i+1} に挟まれているブザーを s_i と表す。

4.1 状態遷移モデル

ロボットの自由度は 2 で、図 3 に示すように関節角度は θ_2, θ_3 で表す。したがって、ロボットの姿勢を示す状態ベクトルは次のように表現できる。

$$x_k = [\theta_2^k, \theta_3^k]^T \quad (1)$$

また、推定期間中のロボットは静止しているので、状態遷移モデルは以下のように表すことができる。

$$x_k = x_{k-1} + w_k \quad (2)$$

w_k はモデル誤差で、白色雑音 (平均 0, 分散 σ_p^2) とする。

4.2 観測モデル

観測ベクトルを各ブザーの発音に対する各マイクロホンの到達時間差とする。到達時間差は、(1) 各ブザーから順番に 4kHz の矩形波を再生して同期マイクロホンアレイで録音し、(2) 各チャンネル毎に、シフト幅 1 の短時間フーリエ変換の 4kHz 帯におけるパワーのピーク時刻を求め、(3) それらの時間差を求めることで算出する。

この時、観測モデルは次のように表すことができる。

$$z_k = [\tau_{1,3}^1 \tau_{1,4}^1 \tau_{2,1}^2 \tau_{2,4}^2 \tau_{3,1}^3 \tau_{3,2}^3]^T + v_k \quad (3)$$

時間差 $\tau_{j,k}^i$ は s_i が発した音の m_j と m_k が録音した時刻の差である。 $D_{i,j}, D_{i,k}$ をそれぞれ s_i と m_j, m_k との距離とすると、 $\tau_{j,k}^i$ は

$$\tau_{j,k}^i = \frac{D_{i,j} - D_{i,k}}{c} \quad (4)$$

となる。 $\tau_{1,2}^1$ は定数であり、 $\tau_{3,4}^1$ は観測ベクトルより計算できるように、任意の $\tau_{j,k}^i$ は観測ベクトルから計算可能である。 v_k は観測誤差で、白色雑音 (平均 0, 分散 σ_c^2) とする。

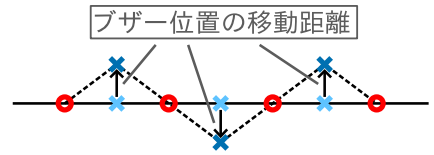
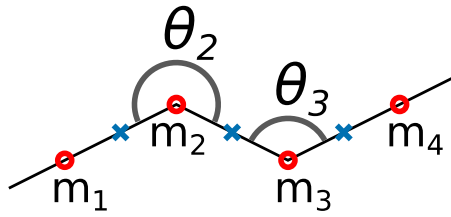
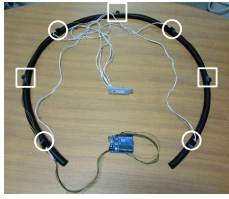


図 2: チューブ型ロボットのモックアップ 図 3: ロボットのモデル化 (●:マイクロホン, ○:ブザー) 図 4: モデルから逸脱したブザー位置 (●:マイクロホン, ×:ブザー)

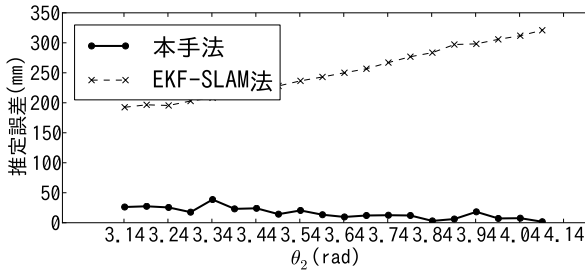


図 5: θ_2 の変動に対する推定誤差

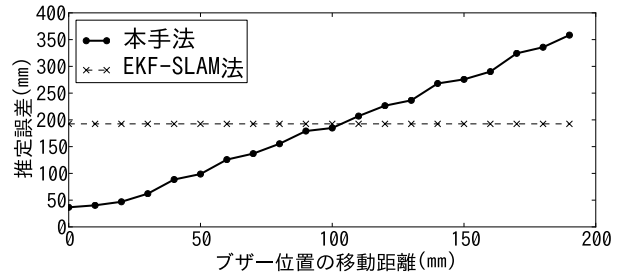


図 6: ロボットの形状に対する推定誤差

4.3 EKF の更新式の導出

本手法は EKF を用い、予測、観測、更新の 3 ステップを繰り返す事によってマイクロホン位置推定を行う。予測ステップでは k 回目のマイクロホン位置の平均 \hat{x}_k^- と分散 P_k^- は式 (2) から次のように更新される。

$$\hat{x}_k^- = \hat{x}_{k-1} \quad (5)$$

$$P_k^- = P_{k-1} + Q \quad (6)$$

ここで、 Q は $Q = \text{diag}(\sigma_p^2, \sigma_p^2, \sigma_p^2)$ で表される共分散行列である。

観測ステップでは、各マイクロホンでの観測時刻から式 (3) に従って以下を得る。

$$h(\hat{x}_k^-) = [\hat{r}_{1,3}^1 \ \hat{r}_{1,4}^1 \ \hat{r}_{2,1}^2 \ \hat{r}_{2,4}^2 \ \hat{r}_{3,1}^3 \ \hat{r}_{3,2}^3]^T \quad (7)$$

更新ステップでは $h(\hat{x}_k^-)$ と z_k の差を最小にするようにカルマンゲイン K_k を導出する。

$$K_k = P_k^- H_k^T (H_k P_k^- H_k + R) \quad (8)$$

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - h(\hat{x}_k^-)) \quad (9)$$

$$P_k = (I - K_k H_k) P_k^- \quad (10)$$

ここで、 $H_k = \frac{\partial h(x_k)}{\partial x_k} |_{x_k = \hat{x}_k^-}$ はヤコビアンであり、 R は $R = \text{diag}(\sigma_c^2, \sigma_c^2, \sigma_c^2)$ で定義される共分散行列である。

5. 評価実験

数値シミュレーションによって本手法と EKF-SLAM 法を比較する。1 回の試行には 40 個の観測データを用い、10 試行の平均を評価する。

5.1 ロボットの初期値に対する頑健性

固定した初期値に対するロボットの姿勢変動に対する頑健性を評価する。 θ_2 の変動に対する推定値の誤差を計算した。 $\theta_2 = \theta_3 = \pi(\text{rad})$ を初期値とし、 θ_2 を $\pi(\text{rad})$ から $\pi + 0.5(\text{rad})$ の範囲で変動させ推定を行った。

図 5 に結果を示す。この結果よりロボットの首振りに対しては本手法がより頑健といえる。角度差が大きい時に精度が良くなっているのは、 H_k を計算する際に用いる三角関数が非線形であるためと考えられる。

5.2 モデルからの逸脱に対する頑健性

実環境中のチューブ型ロボットは曲線的に折れ曲がるため、ブザーはマイクロホン同士の中点上に位置しない。ブザー位置をモデルの仮定する位置から移動させた場合の影響を評価する。図 4 のようにブザー位置を垂直にずらし推定を行った。初期値は真値とする。ただし、ブザーとマイクロホンは同じチューブで繋がっているので L をマイクロホンとブザー間の折れ線の長さとする。

図 6 に結果を示す。ブザー位置の移動距離が増大するにつれて推定値の誤差も増大している。実環境中ではロボットが大きく曲がることもあるが、本結果よりそのような場合には精度が大きく下がってしまうことが分かる。

6. おわりに

本稿ではチューブ型ロボットに音源定位・分離・認識機能のための姿勢推定モデルを開発した。物理的な形状制約を利用して EKF による姿勢推定を実現した。チューブ型ロボットを折れ線でモデル化した場合は従来手法より良い性能を得た。一方、実環境でのロボットの大きな折れ曲がりには対応できないことも判明した。今後はマイクロホン間距離とブザー位置誤差を推定するモデルを検討し、より高精度な姿勢推定を行う予定である。

謝辞 本研究の一部は科研費基盤 (S) No.24220006 の支援を受けた。チューブ型ロボットのサンプルを提供いただいた東北大学大学院田所教授、昆陽准教授らに感謝する。

参考文献

- [1] K. Hatazaki, *et al.*: "Active scope camera for urban search and rescue", 2596-1602, IROS-2012
- [2] T. Mizumoto, *et al.*: "Design and Implementation of Selectable Sound Separation on Texai Telepresence System using HARK", ICRA-2011, pp.2130-2137
- [3] N. Ono, *et al.*: "Blind alignment of asynchronously recorded signals for distributed microphone array," pp.161-164, WAS-PAA 2009
- [4] H. Miura, *et al.*: "SLAM-based online calibration of asynchronous microphone array for robot audition," pp.524-529, IROS 2011
- [5] Greg Welch, *et al.*: "An Introduction to the Kalman Filter," U of North Carolina (UNC), Department of Computer Science, 1995