

柔軟索状レスキューロボットのための ロバスト主成分分析を用いた走行雑音抑圧

坂東 宜昭¹ 池宮 由楽¹ 糸山 克寿¹ 昆陽 雅司² 田所 諭² 中臺 一博³ 吉井 和佳¹ 奥乃 博⁴

¹京都大学 大学院情報学研究科

²東北大学 大学院情報科学研究科

³東京工業大学 情報理工学研究所, (株) ホンダ・リサーチ・インスティテュート・ジャパン ⁴早稲田大学 理工学術院

1. はじめに

柔軟索状ロボット [1, 2] 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被災者を探るために開発されている。人間の侵入が困難な災害現場においては、被災者の声を手がかりにしたレスキューロボットによる探索が有用である。柔軟索状ロボットなどの地上走行型ロボットでは、駆動に伴う振動・摩擦音などの走行雑音によって被災者の声が聞き取りづらくなる問題があった。ロボット周辺の音環境を鮮明にリモートオペレータへ提示するために、ロボットの走行雑音の抑圧が不可欠である。柔軟索状ロボットの走行雑音は、周期的に変化し、接地面の材質・形状に依存する。ロボットの動作雑音を抑圧する従来法 [3-6] では、雑音の定常性を仮定するため性能が劣化する、事前学習が必要といった点で本ロボットには不適であった。この問題を解決するため、繰り返し出現する周波数成分を事前情報を用いずに除去することができる Robust Principal Component Analysis (Robust PCA) [7] を用いて走行雑音抑圧を行う。実際にロボットを動作させながら録音した音響信号を用いた実験により、提案法を評価した。

2. 柔軟索状ロボットの走行雑音抑圧

柔軟索状ロボットの走行雑音抑圧システムには、オンライン処理と、壁とマイクロホンの接触時の対処が不可欠である。本システムはリモートオペレータが災害現場でリアルタイムに音環境を確認するために、オンラインで動作しなければならない。また、ロボット上のマイクロホンが壁や床と接触した場合に雑音成分が大きくなり、雑音抑圧性能の劣化を引き起こすため対処が必要である。

柔軟索状ロボットの走行雑音抑圧システムを図 1 に示す。オンライン動作は、従来オフライン処理である Robust PCA を拡張した Online Robust PCA [8] を用いて解決する。壁との接触問題は、ロボット上に配置方向の異なる複数のマイクロホンを配置することで解決する。本稿で扱う走行雑音抑圧の問題設定は以下の通り。

入力	ロボット上の M 個のマイクロホンで録音した音響信号の振幅スペクトル $z_{t1}, \dots, z_{tM} \in \mathbb{R}^F$
出力	走行雑音を抑圧した振幅スペクトル $s_t \in \mathbb{R}^F$

ここで、 F は周波数成分数を表す。時刻 t 、マイクロホン番号 m の振幅スペクトル $z_{tm} = [z_{tm1}, \dots, z_{tmF}]^T$ は録音信号を短時間フーリエ変換し振幅を計算することで得る。

2.1 プロトタイプ：柔軟索状ロボット

図 2 に、マイクロホンアレイを搭載した柔軟索状ロボットの写真を示す。本体は、直径 38 mm のコルゲートチューブからなり、全長 3 m である。 $M = 8$ のマイクロホンがロボット表面に 40 cm 間隔で 90 度ずつ回転させながら装着されている。両端のマイクロホン間の距離

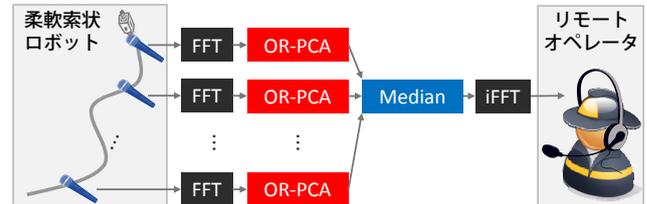


図 1: 走行雑音抑圧システムの概要

は 2.8 m である。本ロボットは、Namari らの Tube-type Active Scope Camera [1] と同様、繊毛と振動モータを用いた駆動で前進する。振動モータはロボット内に 40 cm 間隔で 7 つ直列に装着されている。

2.2 Robust PCA

Robust PCA は、入力行列 $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_T]$ (T はフレーム数) を低ランク行列 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_T]$ とスパース行列 $\mathbf{E} = [\mathbf{e}_1, \dots, \mathbf{e}_T]$ に分解する。行列分解には、以下のコスト関数 $l(\mathbf{X}, \mathbf{E})$ が用いられる。

$$l(\mathbf{X}, \mathbf{E}) = \frac{1}{2} \|\mathbf{Y} - \mathbf{X} - \mathbf{E}\|_F^2 + \lambda_1 \|\mathbf{X}\|_* + \lambda_2 \|\mathbf{E}\|_1 \quad (1)$$

ただし、 $\|\cdot\|_F$ 、 $\|\cdot\|_*$ 、 $\|\cdot\|_1$ および λ_1, λ_2 はそれぞれ Frobenius, 核, L1 ノルムとスケールパラメータを表す。Online Robust PCA では、式 1 の核ノルムを Frobenius ノルムへ式変形することでコスト関数を単純化し、入力ベクトル \mathbf{y}_t を低ランク成分 \mathbf{x}_t とスパース成分 \mathbf{e}_t に逐次的に分解する [8]。

2.3 Robust PCA を用いた走行雑音抑圧

Robust PCA の音響信号への応用では、振幅スペクトル z_{tm} に対して処理を行う。雑音成分を高頻度出現する低ランク成分 $\mathbf{x}_{tm} \in \mathbb{R}^F$ 、音声成分を出現頻度が少ないスパース成分 $\mathbf{e}_{t,m} \in \mathbb{R}^F$ とみなすことで音声成分を分離する [9]。ただし、入力の振幅スペクトル $z_{t,m}$ と雑音の振幅スペクトル $\mathbf{x}_{t,m}$ 、音声の振幅スペクトル $\mathbf{e}_{t,m}$ の間に次の関係が成り立つことを仮定する。

$$z_{tm} = \mathbf{x}_{tm} + \mathbf{e}_{tm}$$

提案法では、各マイクロホンごとに Online Robust PCA を適用し、その結果の中央値を出力とする。まず、各マイクロホンの入力 z_{tm} の各周波数成分 z_{tmf} へ正規化係数 α_f をかけたスペクトル $z'_{t,m}$ に Online Robust PCA を適用する。Robust PCA の学習では各周波数成分が同じ重みで学習される。本稿で扱う柔軟索状ロボットの走行雑音は低周波数成分に強いピークをもつため、経験的に得た以下の正規化係数を用いて低周波数の成分を抑圧し過学習を軽減する。

$$z'_{tm} = [\alpha_1 z_{tm1}, \dots, \alpha_F z_{tmF}]^T, \alpha_f = (\mu_1 e^{-\mu_2 f} + 1)^{-1}$$

ただし、 μ_1 と μ_2 はスケールパラメータである。出力信号 $s_t = [s_{t1}, \dots, s_{tF}]^T$ は、各マイクロホンのスパース成分 \mathbf{e}_{tm} の各周波数成分 e_{tmf} の中央値とする。

$$s_{tf} = \frac{1}{\alpha_f} \text{Median}(e_{t1f}, \dots, e_{tMf})$$

Ego-Motion Noise Suppression based on Robust PCA for a Hose-shaped Rescue Robot: Y. Bando, Y. Ikemiya, K. Itoyama, M. Konyo, S. Tadokoro, K. Nakadai, K. Yoshii and H. G. Okuno

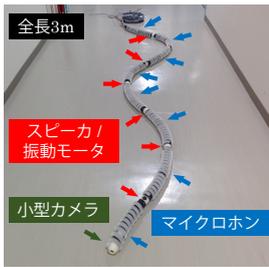


図 2: 柔軟索状ロボット

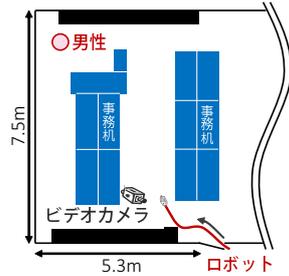


図 3: 実験 1 の配置図

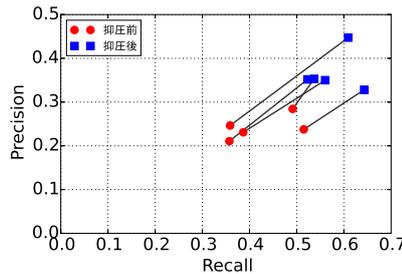


図 4: 実験 1 の被験者毎の VAD 結果

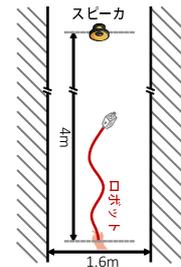


図 5: 実験 2 の配置図

3. 実験

柔軟索状ロボットを用いて、走行雑音と音声を同時に録音した実録音を用いた被験者実験と、これらを別々に録音し SNR を変化させながら混合して音声の分離精度を評価する実験を行った。

3.1 実験 1: 実録音を用いた実験

実際に柔軟索状ロボットを駆動させながら助けを呼ぶ声を録音し、被験者実験により雑音抑圧の有効性を確認する。図 3 のように、室内に助けを呼ぶ男性を配置し、部屋の扉から柔軟索状ロボットを侵入させた。このとき、男性には「助けて」「おーい」などと呼ばせロボット上のマイクロホンアレイで録音した。録音は 24 bit 量子化、16 kHz サンプリングで約 1 分間行った。

5 名の被験者に録音信号を雑音抑圧した場合としない場合の音響信号を無作為な順で聴いてもらい、声がする間ボタンを押してもらった。被験者が音響信号中にボタンを押した時刻と放した時刻を記録し、音声区間検出 (VAD) の精度を計算した。VAD の正解データは図 3 中のビデオカメラの録音信号を用いて手作業で作成した。雑音抑圧法の各パラメータは実験的に与えた。

図 4 に被験者ごとの VAD 結果の Precision と Recall を示す。すべての被験者で雑音抑圧した場合に Precision と Recall が共に向上している。また、F 値は平均で 30.6% から 44.6% に向上した。これらは雑音抑圧により、音声の感知がより容易になったことを示唆している。

3.2 実験 2: シミュレーション混合音を用いた実験

柔軟索状ロボットを用いて音声と走行雑音を個別に録音し、SNR を -25 dB から 10 dB まで 5 dB ずつ変化させながら混合して音声の分離精度を評価した。図 5 に示すように、ロボットと音声を再生するスピーカを配置した。音声は、ノイズを軽減するために、ロボットの初期姿勢時のインパルス応答に男声の音声を畳み込んで作成した。混合した音響信号の長さは 60 秒である。

提案法と、先端のマイクロホンのみを Online Robust PCA をかけた場合 (Mono OR-PCA) と、Histogram-based Recursive Level Estimation (HRLE) 法 [6] との比較を行った。HRLE 法はオンラインで雑音を抑圧する手法の一つで、Online Robust PCA が複数の基底の線型結合で現在の雑音を表現するのに対し、HRLE 法では最頻のスペクトルを雑音とみなす。

図 6 に NSDR による評価結果を示す。NSDR は、目的音の歪みと雑音の歪みが抑圧前からどの程度回復したかを示す指標である。-15 dB 以上の SNR で提案法が最も NSDR が高い。これは、Online Robust PCA をマイクロホンアレイに適用する提案法の有効性を示している。

一方で、-20 dB 以下の SNR では、提案法より先端のマイクにのみ Online Robust PCA を適用した場合の

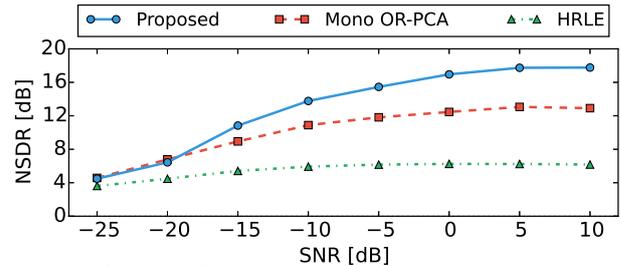


図 6: 実験 2 の結果. 各 SNR での音声の NSDR を示す。

NSDR の方が高い。これは、ロボット上のマイクロホンと音源が直線的に配置されているために、根本のマイクロホンは先端のマイクロホンより音声の音量が小さくなり、目的音の SNR が小さい時に性能が劣化していると考えられる。本問題は、統合に使用するマイクロホンを動的に選択することで解決が期待できる。選択の指針には、調波構造等の音声特有の特徴を用いて各マイクロホンごとに推定した音声含有率を用いる。

4. おわりに

本稿では、柔軟索状ロボットの走行雑音抑圧のために、Online Robust PCA をマイクロホンアレイに対し適用する手法を開発した。実録音を用いた被験者実験にて、雑音抑圧を行わない場合に比べて、音声区間認識精度が向上することを確認し、シミュレーション混合による実験で提案法により NSDR が最大 4.9 dB 向上することを確認した。今後は模擬瓦礫環境での評価、被災者との双方向音声通信システムの構築・評価などを行う。

謝辞 本研究は、科研費基盤 (S) No.24220006 および ImPACT 「タフ・ロボティクス・チャレンジ」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] G. Ince *et al.* Incremental learning for ego noise estimation of a robot. In *IROS-2011*, 131–136.
- [4] T. Tezuka *et al.* Ego-motion noise suppression for robots based on semi-blind infinte non-negative matrix factorization. In *ICRA-2014*, 6293–6298.
- [5] B. Cauchi *et al.* Reduction of non-stationary noise for a robotic living assistant using sparse non-negative matrix factorization. In *SMIAE-2012*, 28–33.
- [6] H. Nakajima *et al.* An easily-configurable robot audition system using histogram-based recursive level estimation. In *IROS-2010*, 958–963.
- [7] E. J. Candès *et al.* Robust principal component analysis? *Journal of the ACM*, 58(3):11, 2011.
- [8] J. Feng *et al.* Online robust pca via stochastic optimization. In *NIPS-2013*, 404–412.
- [9] C. Sun *et al.* Noise reduction based on robust principal component analysis. *Journal of Computational Information Systems*, 10(10):4403–4410, 2014.