

ロボットによる物体認識法の開発

—人工知能・機械学習技術を用いてセンサデータを巧みに操る—

電子制御工学科・田中 大介

ロボットが様々な環境で自律的に動作するためには、ロボットのセンサから得られる視覚情報（カメラ）や聴覚情報（マイク）を基にして周囲の環境を認識する必要があります。例えば、自分はどのような場所にいるのか（人が大勢いたりモノが多数あるのか）、足元はどのような材質の床なのか（どのように動いたり歩いたりすれば良いのか）、目の前にあるモノはどのような形でどれくらい硬いのか（どれくらい力を加えても良いのか）といった情報です。環境情報があることによって、ロボットはより適切な動作を行うことができます。我々の研究グループでは、物体認識（ここでは「目の前にある物体の特性を推定する」こととします）問題を例題として研究を行なっています。

このような物体認識問題を解くにあたって、次のような問題が考えられます。

1. カメラやマイクそれぞれのセンサデータを独立に用いるのではなく、それらの情報を相補的に利用できるようになれば、頑健な認識が可能となります。例えば、モノの特徴的な部分がカメラの視界から消えてしまっても、音響情報や触覚情報から補える可能性があります。しかし、使うセンサデータを増やせば増やすほど、同時に処理しなければならない情報量が増えるため計算量が膨大となり、リアルタイムでの認識が困難になってしまう恐れがあります。
2. センサデータは、ロボットが何らかの動作を行って得られるものです。すなわち、動作中時々刻々と変化するセンサデータをサンプリングすると、1動作により1個の時系列信号が得られます。この時系列信号を認識に利用するにあたり、平均を取るなどして時間情報を圧縮することも考えられますが、動作中の時間変化の情報を利活用するためには、時系列信号を“時系列信号らしく”取り扱う必要があります。

我々の研究室では、まず2.の問題に挑んでおり、『コンテナの中に入ったワッシャーの数を推定する』というタスクに対して、音響情報に基づき推定する手法を開発しました [1]。ワッシャー入りのコンテナを振ったときの音は、振り方によって変わります。提案法はマイクから得られた音響情報とコンテナを振るときの行動情報を融合(図 1(a))し、振ったときにコンテナの中のワッシャーがコンテナの内部にあたり跳ね返る音の特徴を抽出(図 1(b))します。この提案手法の有効性を実機ロボットを用いて検証(図 1(c))したところ、行動情報を入れない場合と比較して約 6% 認識率が向上しました。

現在は 1. の問題にも取り組んでおり、画像や触覚情報等も含めた認識技術の開発に取り組んでいます [2]。

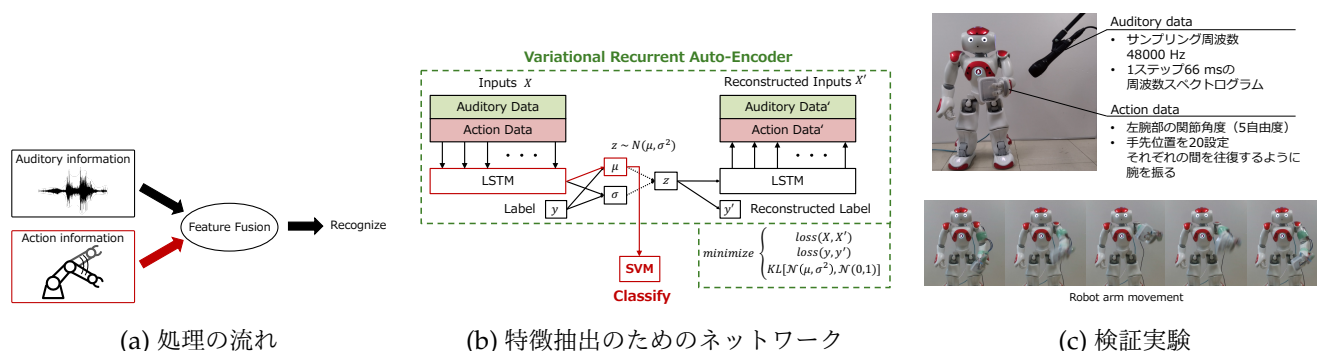


図 1: 提案手法 | 音響情報と行動情報に基づく物体認識法 [1]

発表論文

- [1] 林良優, 田中大介: “Variational Recurrent Auto-Encoder を用いた音響・行動情報に基づく物体認識” 第 19 回 計測自動制御学会 システムインテグレーション部門講演会 (SI2018), 3A1-07 (USB), 2018/12/13-15
- [2] Tomoya Yamanokuchi, Ryoya Hayashi, and Daisuke Tanaka: “Dimensionality Reduction Method for Gaussian Process Regression and Its Application to Object Recognition Tasks” SICE Annual Conference 2018, pp. 646-649, Nara, Japan, September 11-14, 2018